Targeted Social and Emotional Learning

Toolkit technical report

Ciara Keenan, Jaimee S. Mallion, Richard Nugent, Sophie Hedges, Iman Idrees, Lauren Murphy, Karen Allen, Ofelia Torres, Jessica Hinks, Jenny Towers

March 2026

About National Children's Bureau

This report has been produced by the National Children's Bureau on behalf of the Youth Endowment Fund. The National Children's Bureau works collaboratively across the issues affecting children to influence policy and get services working together to deliver a better childhood. They were commissioned by the Youth Endowment Fund (YEF) as their Toolkit Partner 2023-2026. The Toolkit partnership is managed on behalf of NCB by Dr Ciara Keenan. Any queries relating to the methods should be directed to ckeenan@ncb.org.uk

About Youth Endowment Fund

The Youth Endowment Fund's mission is to prevent children and young people becoming involved in violence. They do this by finding out what works and building a movement to put this knowledge into practice. The fund was established in March 2019 by children's charity Impetus, with a £200m endowment and ten-year mandate from the Home Office. For more information, please visit www.youthendowmentfund.org.uk.

Acknowledgements

Abstract/Plain Language Summary

The objective of this report is to review the evidence on the effectiveness of targeted social and emotional learning (Targeted SEL) in reducing and/or preventing violence and offending involving children and young people. SEL involves the acquisition of the knowledge and skills to develop healthy identities, manage emotions, feel and show empathy for others, communicate effectively, and establish and maintain supportive relationships. Targeted interventions focus on young people at heightened risk of, or already showing, social and emotional problems, or engagement in crime, violence, or antisocial behaviour.

Key findings

- Targeted SEL interventions are associated with a **42.7%** reduction in violence outcomes, based on a meta-analysis of 15 outcomes across nine studies. There is **Low confidence** in this finding. Although the evidence base included one cluster randomised controlled trial and eight randomised controlled trials that provided causal evidence, the overall body of evidence did not meet our threshold for a moderate confidence rating.
- Targeted SEL interventions are associated with a **31.1%** decrease in crime and offending, based on a meta-analysis of seven measured outcomes across three studies. There is **Very Low confidence** in this finding. One cluster RCT and two RCTs provided causal evidence (two high quality, one low quality).
- Gender was found to significantly moderate the impact of Targeted SEL interventions, with the largest and most consistent effect seen when boys make up most or all of the sample.
- The effectiveness of Targeted SEL does not differ meaningfully by the ethnic composition of study samples. Only studies in which the sample's ethnicity was not stated or unclear showed a significant effect, possibly due to the large number of studies within this group. This demonstrates that under-reporting of participants' ethnicity remains a major issue.
- Age was found to significantly moderate the impact of Targeted SEL interventions. For younger children (0–9 years), effects were moderate

although this result was not statistically significant, while for early adolescents (10–14 years), estimated impacts were high but also not statistically significant. For older adolescents (15–17 years), the effect size was very large, but this estimate is based on a single study and should be interpreted with caution. Studies which target a broad range of ages showed a statistically significant and high impact.

- Successful implementation relies on high fidelity to the original intervention design, underpinned by clear protocols, substantial teacher training and ongoing support, sufficient staffing for needs-focused one-to-one or small-group delivery, and deliberate planning to ensure focused coverage of core SEL content drawing on developmentally appropriate formats (e.g., play-based learning for younger children).

- Two studies provided data on the costs associated with setting up and running Targeted SEL interventions. Costs included paying for building space, initial and ongoing training, provisions of handouts and materials and supplies for programme leaders. However, research is needed to explore the cost-effectiveness of Targeted SEL, including potential long-term savings to the youth justice system, education, health and children's services.

- Among the RCTs included in our review, only one reported qualitative data on participants' perceptions of engaging in a Targeted SEL intervention, with positive reviews of their experience of the group, group content, activities and group leadership. However, more research is needed to develop our understanding of children and young people's experiences of Targeted SEL further.

Conclusion

Our meta-analysis finds that Targeted SEL programmes substantially reduce children and young people's involvement in violence, and improve behaviour, social skills, and school engagement, with promising but uncertain evidence of reduced crime and offending. Targeted SEL interventions should be implemented using one-to-one or small-group formats to match children and young people's needs, supported by clear protocols, monitoring systems to ensure high fidelity to

core SEL content and original intervention design. Implementing Targeted SEL programmes also requires sustained teacher training and development.

Table of Contents

List of tables

List of figures

Preface on Terminology

This review draws on evidence spanning over half a century, during which language around personal characteristics has evolved significantly. At times, we may have to reproduce original terminology used in studies which we recognise today as being outdated and unacceptable offensive terms. This only occurs when the terminology is used in direct quotations or refers to an outcome that the author measured that remains relevant to our analysis. The wider narrative will adhere to current inclusive-language standards guided by the National Children's Bureau, Youth Endowment Fund and Race Equality Foundation. These guiding principles include using capitalisation to acknowledge shared identities (e.g., Black, Asian), whilst not capitalising white due to its association with white supremacy. The review also avoids deficit framing and respects individuals' self-identification. Person-first language will generally be used when referring to children and young people, except for Deaf and autistic communities, who widely prefer identity-first language. The team acknowledges limitations in terminology and strives for respectful and precise representation throughout. The full preface on terminology can be accessed here.

Objective and Approach

The objective of this report is to review the evidence on the effectiveness of **Targeted SEL** in reducing and/or preventing violence and offending involving children and young people. Social and emotional learning (SEL) involves the acquisition of the knowledge and skills to develop healthy identities, manage emotions, feel and show empathy for others, communicate effectively, and establish and maintain supportive relationships. SEL is generally recognised to comprise five key competencies: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (CASEL, n.d.).

Previous reviews within the youth justice field have focused more specifically on 'social skills training', a subset of SEL which focuses primarily on self-control, perspective-taking, and the inhibition of antisocial behaviour. These reviews have shown that children who take part in social skills training are less likely to become involved in crime and violence (Beelmann & Lösel, 2021; Gaffney et al., 2021; Piquero et al., 2016a). With social skills training, the presumed causal mechanism is that developing self-control, anger management, and problem-solving skills encourages children to think before they act, reduce impulsivity, and avoid aggression, leading to a reduced risk of involvement in antisocial behaviour, crime and violence (Gaffney et al., 2021).

The current review extends the focus to social and emotional learning more broadly, recognising that skills such as empathy, emotional regulation, and communication, are also important in supporting children to develop positive, prosocial identities, healthy, supportive relationships, and positive coping strategies, all of which help children to avoid aggression and antisocial behaviour.

The review included any studies which provided outcomes for children and young people aged 0 to 17. Included programmes were those in which the primary active component was the structured development of social and emotional skills through a documented SEL curriculum (manual, lesson plans, defined modules), with at least 30% of contact time devoted to SEL skill-building[1]. Examples include

---

[1] The 30%–50% threshold was agreed at the scoping stage as a pragmatic way of defining a substantively SEL-focused intervention and reducing overlap with other strands.

PATHS, Second Step, and emotional literacy classes. The intervention must also seek to reduce crime, offending, or violence, or externalising behaviours associated with violence such as aggression, bullying or substance misuse, or school exclusions and disciplinary actions. Only studies with a treatment as usual comparison group were included; studies with an alternative treatment group were excluded.

Social and emotional learning programmes can be targeted or universal. The focus of this review is on **Targeted SEL** only. Targeted programmes focus on young people at heightened risk of or already showing social and emotional difficulties, or engagement in crime, violence, or antisocial behaviour. While the causal mechanisms are the same for universal and targeted programmes, children and young people in targeted programmes may have more complex circumstances or needs that may affect the ways in which they access or engage with programmes. Targeted programmes often use small-group or individual formats, sometimes with more tailored support, while universal programmes are more likely to offer a class-based, structured curriculum.

This technical report draws on a comprehensive systematic review methodology and includes:

- **Twenty-six effectiveness studies** from the USA, Canada, the Netherlands, and Turkey, providing outcomes data for meta-analysis: 15 violence outcomes across 9 studies; 7 crime and offending outcomes across three studies; and 135 outcomes across the whole of the YEF outcomes framework.
- **A meta-review consisting of four systematic reviews** providing implementation insights on what factors influenced the effectiveness of Targeted SEL interventions. Findings are reported according to the following implementation factors: intervention curriculum, delivery method, duration/dosage, fidelity, intervention providers and training, and parental involvement.

The remainder of this report is structured as follows: First, the **Description of the Intervention** outlines the key components of Targeted SEL and its intended implementation. Second, **How Effective is the Intervention?** presents findings from our meta-analysis on crime reduction and broader social outcomes. Third,

Who Does it Work For? examines evidence on the populations that benefit most from Targeted SEL. Fourth, **What Factors Affect Implementation?** explores key factors which influence the effectiveness of Targeted SEL interventions. Fifth, **How Much Does It Cost?** reviews available cost data. Finally, the **Conclusion and Takeaway Messages** summarise key findings and recommendations, followed by **Appendices** detailing the systematic review methodology and characteristics of included research.

Description of the Intervention

In the following section details are provided on the interventions which inform this report, noting their key components, any equipment, materials, supplies or training required, the duration and intensity of interventions, who delivered the interventions, and where and how the interventions were delivered.

Features of the approach

Social and emotional learning supports children to think before they act, understand other people's perspectives, communicate effectively, and use strategies for managing impulsiveness or aggression. Programmes can vary from a single session to multi-year interventions.

Targeted SEL is aimed at children and young people with an identified need such as conduct problems, behavioural difficulties or aggression.

One of the included studies in the review is from England (Spence & Marzillier, 1981). Two are from Canada (Dong et al., 1979; Schneider & Byrne, 1987), one from the Netherlands (Stouwe et al., 2016), one from Turkey (Çelik et al., 2016) and the remainder from the USA.

Key components of Targeted SEL

All the Targeted SEL programmes included in this review are manualised approaches to supporting children and young people to develop social and emotional competencies.

Programmes are typically highly structured, with session plans, activities, and in some cases scripts. Children and young people are supported to recognise emotions, develop empathy, display pro-social behaviour and problem-solve. Structured techniques are presented, such as problem-solving through identifying difficult situations and planning and evaluating potential solutions and their consequences (Block, 1978; Camp et al., 1977; Dupper & Krishef, 1993; Edleson & Rose, 1982; Feindler et al., 1984; Kazdin et al., 1987; Lochmann et al., 1993; Mannarino et al., 1982; Prinz et al., 1994; Schneider & Byrne, 1987; Smokowski et al., 2004; Spence & Marzillier, 1981).

Participants were identified in a number of ways, typically the selection is driven specifically by an identified behavioural need or elevated risk level, rather than demographic traits. Many programs use structured assessments and checklists to identify specific challenging behaviours (Hudley et al., 1998); some programs identify participants by reviewing existing school administrative records to find children with a history of behavioural issues (Collier, 2002); while some participants may also be identified through voluntary or mandatory referrals from external bodies, such as children's services or the youth justice system (Burke & Loeber, 2015)

Where interventions directly targeted existing anger, delinquency or offending behaviour (Burke & Loeber, 2015; Collier, 2002; Conduct Problems Prevention Research Group, 2010a; Feindler et al., 1984; Hudley et al., 1998; Hudley & Graham, 1993; Kazdin et al., 1987; Lochmann et al., 1993; Quinn, 2002; Smokowski et al., 2004; Spence & Marzillier, 1981; Splett et al., 2015; Stouwe et al., 2016; Tanner & Holliman, 1988), these typically taught similar problem solving strategies, but also incorporated anger management techniques, advice on how to calm down, deal with the police or keep out of fights, and might also address impulsivity or attributional biases.

A small number of studies had a primary focus on attributions or assertiveness. Hudley et al. (1998) and Hudley & Graham (1993) described an attributional intervention (named as The BrainPower Program in Hudley et al., 1998), which encouraged participants not to infer hostile intent. Participants were supported to understand and identify non-hostile intent, make attributions and generate rules about how best to respond. Assertion training (Dong et al., 1979) and assertiveness social-skills training (Tanner & Holliman, 1988) focused on increasing assertive behaviours to replace aggression. Children were taught to identify non-assertive, assertive and aggressive behaviours, consider their own responses and practice calming techniques such as breathing exercises.

One study had a slightly different focus. The Children of Divorce Intervention Project (Pedro-Carroll, 1983) worked to help children understand their feelings about their parents' divorce and reduce feelings of stigma and isolation by emphasising children's shared experiences. The programme supported development of children's self-control and problem-solving skills to help children

recognise that while the divorce was out of their control, they can control their behaviour.

Interventions typically lasted less than a year. An exception to this is the Fast Track Program (Conduct Problems Prevention Research Group, 2010a), which aimed to prevent serious behavioural difficulties through a ten-year intervention incorporating social skills training, support with peer relationships and parenting practices. 'Peer-pairing' sessions were also offered to support peer relationships and social skills. In the first year of the intervention children and families were all offered the same level of support, however in subsequent years support was tailored to individual need. For instance, the programme adjusted the intensity of tutoring, home visits, and peer coaching based on children's and families' assessed needs. In addition to these interventions, an adapted version of Promoting Alternative Thinking Strategies (PATHS, a universal social skills programme) was delivered by teachers in the classroom.

All programmes had an element of instruction or learning, though used various other techniques to deliver their aims such as:

- Group discussions to support problem-solving and provide feedback (Block, 1978; Burke & Loeber, 2015; Camp et al., 1977; Collier, 2002; Dong et al., 1979; Dupper & Krishef, 1993; Edleson & Rose, 1982; Hudley et al., 1998; Hudley & Graham, 1993; Kazdin et al., 1987; Prinz et al., 1994; Quinn, 2002; Spence & Marzillier, 1981; Splett et al., 2015; Webster-Stratton et al., 2001).

- Role play/observations to model and practice positive behaviours. This could include identifying and expressing emotions, handling confrontation, reflecting on individual behaviour and considering alternative responses to challenging situations (Block, 1978; Burke & Loeber, 2015; Collier, 2002; Dong et al., 1979; Dupper & Krishef, 1993; Edleson & Rose, 1982; Feindler et al., 1984; Hudley et al., 1998; Hudley & Graham, 1993; Kazdin et al., 1987; Lochmann et al., 1993; Mannarino et al., 1982; Prinz et al., 1994; Quinn, 2002; Schneider & Byrne, 1987; Spence & Marzillier, 1981; Splett et al., 2015; Webster-Stratton et al., 2001).

- Homework or journalling, which included tasks such as reflective exercises to encourage the examination of feelings or learn new skills (Block, 1978;

Çelik et al., 2016; Collier, 2002; Dupper & Krishef, 1993; Feindler et al., 1984; Kazdin et al., 1987; Spence & Marzillier, 1981; Splett et al., 2015; Stouwe et al., 2016).

- Games, rewards and praise were central elements to several interventions. Some of these incorporated individual or group games where children could win prize tokens for positive behaviour and lose tokens for negative behaviour (Camp et al., 1977; Çelik et al., 2016; Edleson & Rose, 1982; Feindler et al., 1984; Hudley et al., 1998; Hudley & Graham, 1993; Kazdin et al., 1987; Prinz et al., 1994; Schneider & Byrne, 1987; Shelton et al., 2000; Splett et al., 2015; Stouwe et al., 2016; Webster-Stratton et al., 2001).

- Parents and carers were supported to enforce appropriate discipline and positive communication (Burke & Loeber, 2015; Conduct Problems Prevention Research Group, 2010a; Splett et al., 2015; Stouwe et al., 2016).

- Calming/relaxation techniques such as breathing exercises could support children and young people to move through negative feelings (Feindler et al., 1984; Quinn, 2002; Tanner & Holliman, 1988).

Other techniques mentioned included family therapy (Burke & Loeber, 2015) and academic support (Burke & Loeber, 2015; Conduct Problems Prevention Research Group, 2010a; Shelton et al., 2000).

Equipment, materials or supplies

As per the eligibility criteria, all included interventions were manualised and had a detailed programme manual outlining session content and activities. However, Quinn (2002) noted that teachers were required to create the necessary materials.

Çelik et al. (2016) and Mannarino et al. (1982) reported that assessments were conducted to identify those at risk of certain behaviours and therefore eligible for the intervention. Meanwhile, Conduct Problems Prevention Research Group (2010) utilised assessments to determine the duration and intensity of the intervention and develop personalised intervention plans.

Several interventions used video tapes, either to present new information, enhance learning or to record and feedback on children's behaviour and progress (Block, 1978; Hudley & Graham, 1993; Lochmann et al., 1993; Spence & Marzillier, 1981; Tanner & Holliman, 1988). Pedro-Carroll (1983) also reported that children were shown a video titled "My parents are getting a divorce".

Materials and props to facilitate learning were also reported. These included photographs of people showing specific emotions, puppets, flash cards, colouring books, toys and a chalkboard (Feindler et al., 1984; Mannarino et al., 1982; Tanner & Holliman, 1988; Webster-Stratton et al., 2001).

Other materials required for delivery included various cards for the First Step to Success Early Intervention Program (Çelik et al., 2016). These included red-green cards for the class module, forms to record children's progress and cards for the home modules to support parents to teach the weekly skills. Social skills training (Edleson & Rose, 1982) also required assignment cards where children recorded and reflected on their handling of problems they had encountered that week. Tanner and Holliman (1988) also reported that children were transported home from the sessions on a bus.

Who delivers Targeted SEL

A variety of professionals were involved in the delivery of the social emotional learning programmes, with some delivered by multiple experts from different fields:

- Eleven interventions were delivered by mental health professionals such as counsellors, therapists or psychologists (Çelik et al., 2016; Collier, 2002; Dong et al., 1979; Feindler et al., 1984; Kazdin et al., 1987; Lochmann et al., 1993; Pedro-Carroll, 1983; Prinz et al., 1994; Spence & Marzillier, 1981; Splett et al., 2015; Webster-Stratton et al., 2001). The therapists delivering Social Skills Training (Spence & Marzillier, 1981), The Children of Divorce Intervention Project (Pedro-Carroll, 1983) and Peer Coping-Skills (PCS) Training (Prinz et al., 1994) were supported by professionals such as teachers, doctoral students and social workers.

- Seven were delivered by teachers, either regular classroom teachers or external professionals (Camp et al., 1977; Çelik et al., 2016; Conduct Problems Prevention Research Group, 2010a; Hudley & Graham, 1993; Quinn, 2002; Shelton et al., 2000; Tanner & Holliman, 1988).

- Smokowski et al. (2004) and Stouwe et al. (2016) reported that the interventions were delivered by social workers.

- Others involved in delivering the interventions included parents or carers (Çelik et al., 2016), the principal researcher (Dupper & Krishef, 1993), family coordinators (Conduct Problems Prevention Research Group, 2010a), SNAP service providers (Burke & Loeber, 2015) and psychology students (Mannarino et al., 1982; Schneider & Byrne, 1987).

- Block (1978) referred to "leaders" delivering the intervention, but it is unclear who filled this role. It was noted that these leaders were "trained at the post-master's level with 8 years of professional experience each". Similarly, Edleson & Rose (1982) and Hudley et al. (1998) referred to "group leaders"/"instructors".

How was the intervention delivered

Most interventions were delivered face-to-face in groups (Block, 1978; Burke & Loeber, 2015; Çelik et al., 2016; Collier, 2002; Conduct Problems Prevention Research Group, 2010a; Dong et al., 1979; Dupper & Krishef, 1993; Edleson & Rose, 1982; Feindler et al., 1984; Hudley et al., 1998; Hudley & Graham, 1993; Lochmann et al., 1993; Mannarino et al., 1982; Pedro-Carroll, 1983; Prinz et al., 1994; Quinn, 2002; Schneider & Byrne, 1987; Shelton et al., 2000; Smokowski et al., 2004; Spence & Marzillier, 1981; Splett et al., 2015; Tanner & Holliman, 1988; Webster-Stratton et al., 2001).

Camp et al. (1977), Kazdin et al. (1987) and Stouwe et al. (2016) held solely individual face-to-face sessions, while Çelik et al. (2016), Conduct Problems Prevention Research Group (2010) and Lochmann et al. (1993) held a combination of group and individual sessions.

Splett et al. (2015) reported that parent training occurring both remotely by telephone and in person.

Where Targeted SEL is delivered

Most of the interventions were wholly or partially delivered in schools (Block, 1978; Camp et al., 1977; Çelik et al., 2016; Conduct Problems Prevention Research Group, 2010a; Dong et al., 1979; Dupper & Krishef, 1993; Edleson & Rose, 1982; Feindler et al., 1984; Hudley et al., 1998; Hudley & Graham, 1993; Lochmann et al., 1993; Mannarino et al., 1982; Pedro-Carroll, 1983; Prinz et al., 1994; Quinn, 2002; Shelton et al., 2000; Smokowski et al., 2004; Splett et al., 2015; Tanner & Holliman, 1988).

As well as school delivery, the Conduct Problems Prevention Research Group (2010) and Çelik et al. (2016) had additional parent training and home visiting elements which were delivered at home.

The studies reported by Burke & Loeber (2015) and Stouwe et al. (2016) were conducted in the community while Webster-Stratton et al.'s (2001) Incredible Years Dinosaur Social Skills and Problem Solving Curriculum was delivered in a clinic. Problem-solving skills training (Kazdin et al., 1987) was delivered in a psychiatric facility while the social skills training reported by Spence and Marzillier (1981) was delivered in a care home.

Training for the providers of Targeted SEL

Just over half of the studies reported on training for the providers of Targeted SEL and many of these described substantial training efforts.

Many involved weekly supervision throughout the implementation of the programmes, with feedback and further training involved in some of these sessions (Conduct Problems Prevention Research Group, 2010a; Hudley & Graham, 1993; Kazdin et al., 1987; Lochmann et al., 1993; Mannarino et al., 1982; Pedro-Carroll, 1983; Schneider & Byrne, 1987; Splett et al., 2015; Stouwe et al., 2016).

Some studies reported the length of training. Lochmann et al. (1993) reported two weeks of training while Block (1978), Mannarino et al. (1982), Pedro-Carroll (1983), Splett et al. (2015) and Stouwe et al. (2016) described several training sessions.

Meanwhile Hudley and Graham (1993) reported 16 hours of training while Çelik et al. (2016) reported a full day training session. Çelik et al. (2016) also reported that facilitators watched the programme being modelled for five days before

beginning delivery, and that training and support was provided for parents' delivery of the home module, holding weekly meetings with parents to explain the week's content.

Camp et al. (1977), Hudley and Graham (1993) and Kazdin et al. (1987) reported that providers conducted rehearsals by seeing training cases/groups or conducting training role plays.

Feindler et al. (1984) noted that the programme was delivered by a "trained" therapist, although did not elaborate on this, while Smokowski et al. (2004) reported that little training is necessary because the programme is fully manualised, however also noted that the provider was "trained by the research project".

Duration of Targeted SEL

Where duration was reported, interventions were mostly implemented for between one and three months (Burke & Loeber, 2015; Camp et al., 1977; Collier, 2002; Dong et al., 1979; Dupper & Krishef, 1993; Edleson & Rose, 1982; Feindler et al., 1984; Hudley et al., 1998; Hudley & Graham, 1993; Kazdin et al., 1987; Pedro-Carroll, 1983; Quinn, 2002; Smokowski et al., 2004; Spence & Marzillier, 1981; Splett et al., 2015). Social Skills Training reported by Mannarino et al. (1982) was slightly longer at 14 weeks, while Tools4U was the shortest intervention lasting less than a month (Stouwe et al., 2016). Other interventions were implemented for six months to a year (Lochmann et al., 1993; Schneider & Byrne, 1987; Shelton et al., 2000; Tanner & Holliman, 1988). The one exception is the Fast Track intervention (Conduct Problems Prevention Research Group, 2010a) which continued for 10 years.

Where intensity was reported, interventions were mostly implemented weekly (Burke & Loeber, 2015; Dong et al., 1979; Dupper & Krishef, 1993; Edleson & Rose, 1982; Lochmann et al., 1993; Mannarino et al., 1982; Pedro-Carroll, 1983; Prinz et al., 1994; Schneider & Byrne, 1987; Splett et al., 2015; Stouwe et al., 2016; Tanner & Holliman, 1988; Webster-Stratton et al., 2001).

The programmes reported by Collier (2002), Hudley et al. (1998), Hudley and Graham (1993) and Spence and Marzillier (1981) held sessions twice weekly, while those reported by Block (1978), Camp et al. (1977), Çelik et al. (2016) and Quinn

(2002) were delivered daily. The self-control training programme (Feindler et al., 1984) was delivered biweekly while problem solving skills training (Kazdin et al., 1987) was delivered 2 to 4 times a week.

Intensity of the Fast Track intervention (Conduct Problems Prevention Research Group, 2010a) ranged from weekly in the first year, biweekly in the second to monthly in subsequent years. Additional individualised support was provided at home for parents and through academic tutoring which varied but could be as frequently as twice a week.

How Effective is the Intervention?

This section examines the effectiveness of Targeted SEL in reducing violence, crime and offending, and other related outcomes through a systematic review and meta-analysis, to provide a robust and objective summary of existing evidence, incorporating advanced statistical techniques, including robust variance estimators (Pustejovsky & Tipton, 2022), for improved accuracy.

Quantitative data from **26 effectiveness studies** (with Schneider and Byrne (1987) providing two independent clusters based on a multi-arm trial) provided information across a variety of **135 outcomes** related to the impact of Targeted SEL for children and young people.

These studies, included in the meta-analysis, assessed the effectiveness of Targeted SEL across a total of **2,679 children and young people.** These 26 effectiveness studies employed a range of study designs, including:

- Quasi-Experimental Designs (QED): (n = 1, 3.8%)
- Randomised Controlled Trials (RCT): (n = 22, 84.6%)
- Cluster Randomised Controlled Trials (cRCT): (n = 3, 11.5%)

The 26 effectiveness studies varied considerably in their methodological design and reporting characteristics (see Appendix 3 for individual study details).

Most studies used a prospective design and assigned participants at the start of the study (n = 20, 76.9%), while five studies used pre-existing (natural) differences to create comparison groups (n = 5, 19.2%). One study did not state how participants were allocated. Most studies (n = 23, 88.5%) allocated at the level of individual participants, while three studies (11.5%) allocated groupings or clusters of individuals, e.g. classrooms or schools. All but one study (n = 25, 96.2%) used random allocation, while one study used quasi-random procedures.

The effectiveness studies spanned multiple decades, with the earliest conducted by Camp et al. (1977) and the most recent by Çelik et al. (2016). The studies were conducted in four different countries including:

- USA: (n = 21, 80.8%)
- Canada: (n = 3, 11.5%)

- Netherlands: (n = 1, 3.9%)

- Turkey: (n = 1, 3.9%)

Fuller descriptions of the included studies are available in Appendix 3. Studies were assessed for methodological quality using the YEF-EQA critical appraisal tool and were rated as follows:

- High: (n = 2, 7.7%)

- Moderate: (n = 17, 65.4%)

- Low: (n = 7, 26.9%)

Nine studies did not report their funding source (34.6%). Where reported, funding mainly came from national institute or government grants (n = 15, 57.7%), with two studies being funded through university research grants (7.7%). Most studies did not report any conflict of interest (n = 24, 92.3%). One study explicitly declared no conflict of interest and one study stated that the researchers have a financial interest in the intervention (Conduct Problems Prevention Research Group, 2010).

As outlined earlier, the intensity of interventions varied from daily (n = 4, 15.4%) or a few times per week (n = 6, 23.1%) to weekly (n = 14, 53.8%); two studies did not state how frequently they were delivered. Interventions also differed in their duration, with the majority lasting 1 – 3 months (n = 18, 69.2%). One intervention lasted less than a month (3.8%), two lasted 3 – 6 months (7.7%), four lasted 6 – 12 months (15.4%), and one lasted more than five years (3.8%). One study did not state how long the intervention lasted (3.8%).

Interventions were predominantly delivered in schools (n = 16, 61.5%) or in schools and at home (n = 3, 11.5%), though two were conducted in clinic or inpatient settings (7.7%) and two in the community (7.7%). Three studies did not specify where the intervention took place. Interventions were delivered by a range of professionals, with the majority delivered by a counsellor, therapist, psychologist or other health professional (n = 12, 46.2%) or a teacher or other education professional (n = 7, 26.9%). Three were delivered by a service provider or leader (11.5%), two by a social worker (7.7%) and one by a researcher (3.8%). One study did not report who delivered the intervention.

In terms of demographic representation, most studies had mixed sex samples (n = 10, 38.5%), though seven (26.9) had majority male and six (23.1) had male-only samples. One study had a female-only sample, and two studies did not report the gender balance of the sample. In terms of ethnicity, eight studies (30.8%) reported diverse or balanced participant groups, four studies (15.4%) reported some diversity, and three (11.5%) reported a majority white (≥85% participants of white ethnicity) sample. Eleven studies (42.3%) did not report the ethnicity of the sample. Six studies (23.1%) reported that the majority of families in the sample were on low incomes, while three (11.5%) reported that *some* families were on low incomes. Seventeen studies (65.4%) did not provide information on the socioeconomic status of participants.

Measured outcomes

Across the 26 effectiveness studies, 135 outcomes were identified and extracted, spanning **ten outcome categories** within the YEF Outcomes Framework:[2]

- Behavioural difficulties (k=50, n=23[3])
- Building and maintaining relationships (social-emotional) (k=33, n=12)
- General mental health (k=14, n=6)
- Regulating and managing emotions (k=5, n=3)
- Ability to resolve conflicts (k=3, n=2)
- Crime and offending (k=7, n=3)
- Helping others (prosocial behaviour) (k=4, n=3)
- School engagement (k=14, n=7)
- Self-esteem (k=1, n=1)
- Bullying (k=4, n=1)

To inform our violence estimate, outcomes must be a measure of violent crime or violent offending, or a measure of overt or relational aggression that was directed at another person. Measures of bullying can inform the violence estimate where

---

[2] The YEF Outcomes Framework identifies specific outcomes linked to reducing the risk of children and young people becoming involved in crime and violence, providing a structured approach for measuring the impact of interventions
[3] In this context, k = number of outcomes measured, and n = number of studies.

the study reports bullying as repeated physical violence or repeated verbal harassment.

Across the nine studies that met this criteria, we identified 15 violence outcomes. Thirteen of these were related to behavioural difficulties and two related to crime and offending. Collectively, these measures assess general, physical, verbal, and relational aggression, including overt harm and harm that occurs through social relationships, as well as behavioural proxies to capture both the form and severity of aggressive behaviour across youth populations. Outcomes were derived in a number of ways, including from administrative school records (n=8, 6%), observations (n=44, 33%), self-reported survey measures(n=22, 16%), and survey measures reported by teachers(n=38, 28%), parents(n=18, 13%), and peers(n=4, 3%), with one outcome (1%) not having the data source specified.

We present summary results from two separate multivariate meta-analyses on violence, and crime and offending below (Table 1).

Table 1[4]: *Summary of findings on violence and crime outcomes*

Outcome	SMD	CI (95%)	*P*	% reduction	Impact rating	Number of studies	Evidence rating
Violence	-0.35	-0.59 to -0.12	0.0215	42.7%	High	9	2
Crime and Offending	-0.26	-0.5 to -0.02	0.179	31.1%	High	3	1

Meta-analysis of violence outcomes in Targeted SEL interventions

Targeted SEL is associated with a high impact on violence, corresponding with a 43% reduction in violence, based on 15 measured outcomes across 9studies.

---

[4] SMD = effect size; 95% CI = uncertainty around the effect; p value = likelihood the result is due to chance; % reduction = % decrease in the outcome; impact rating = importance of effect; evidence rating = strength of the evidence. See Appendix 1 for full methods

The primary focus in the initial analysis is the reduction and prevention of violence, as defined by YEF. Violence is a broad construct that incorporates incidents/behaviours as well as convictable offences. Violence may be of a physical, verbal, psychological, or sexual nature (YEF, 2023: p.12).

The team identified 15 outcomes measuring **violence** specifically across nine studies: 5 Type C[5], 4 Type D, giving an overall security rating of **Level 2**. Most violence outcomes measured 'overt aggression' (Smokowski et al., 2004), with two outcomes measuring violent offending (both from Feindler et al., 1984). A total of $k = 15$ outcomes were included in the analysis. The estimated average effect based on the random-effects model was $\hat{\mu} = -0.35$, (95% CI: -0.58, -0.12).

This estimate was statistically significantly different from zero $t = -2.98$, $p < 0.01$, the effect remained the same when adjusting for clustering across studies using robust variance estimation, however the p-value increased, but remaining statistically significant at the 0.05 level (t = -3.02, df = 6.4, p = 0.022). The estimate remained the same suggesting the original multivariate meta-analysis was stable.

Table 2: *RVE Output for meta-analysis on violence outcomes*

	Estimate	SE	t-stat	d.f (Satt)	p-val (Satt)	Sig	YEF impact rating	YEF evidence security rating
Intercept	−0.35	0.12	−3.02	6.4	0.022	*	High	2

In practical terms, this indicates that the intervention had a meaningful impact on violence-related outcomes across nine studies. Based on YEF impact

---

[5] Each primary study is rated as either:
- **Type A:** High-quality randomised controlled trial (RCT)
- **Type B:** High-quality quasi-experimental study
- **Type C:** Moderate-quality RCT or quasi-experimental study with minor limitations
- **Type D:** Low-quality RCT or quasi-experimental study with major limitations

categorisation, the effect size (g = -0.35) corresponds to a "high impact" rating, with good certainty around the adjusted estimate.

According to the Q-test, moderate heterogeneity was present; I^2 = 31.4%, Q(df = 14) = 20.42, p = 0.117. This level of variability suggests that the study results are fairly consistent across the nine studies.

The SMD of -0.35 corresponds to a relative risk reduction of 42.7% in the risk of perpetrating violence, with an absolute risk reduction of 6.8%. Among those who participate in Targeted SEL, 9% go on to be involved in violence, compared to 16% of those who receive business-as-usual (Figure 1).
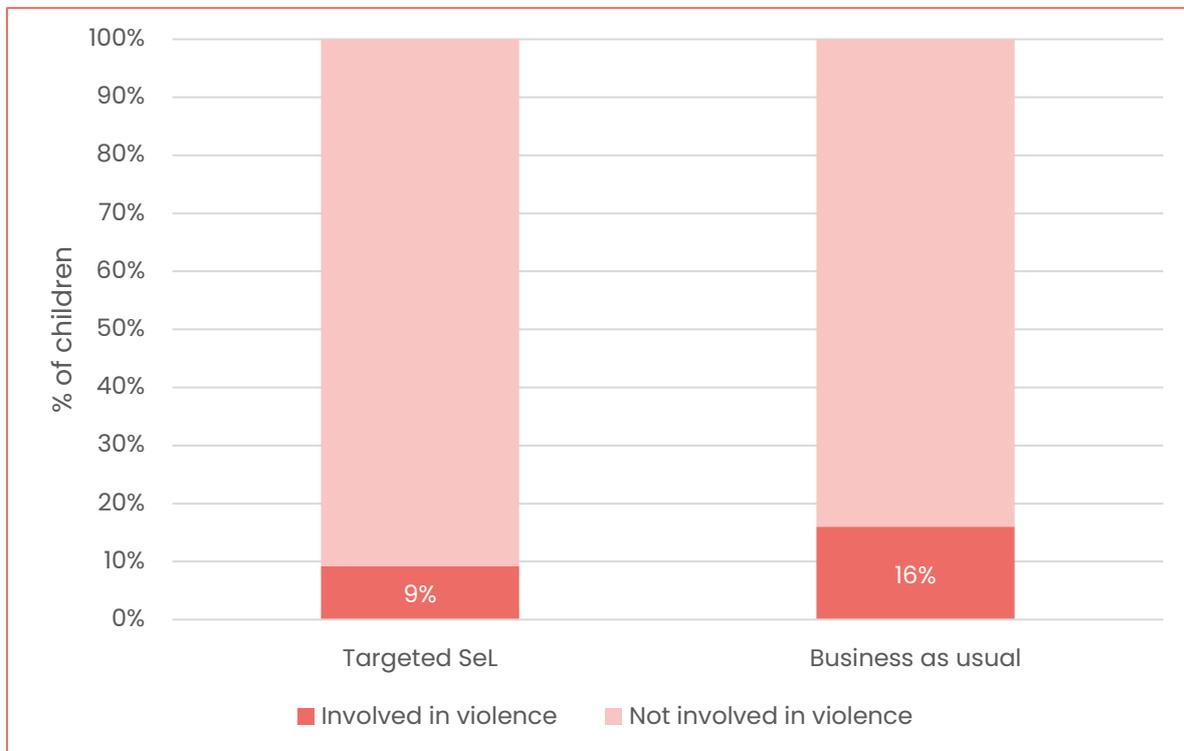
Figure 1: *Absolute risk reduction in violence associated with Targeted SEL interventions compared to business as usual*

Meta-analysis of crime and offending outcomes in Targeted SEL interventions

Targeted SEL is associated with a high impact, corresponding with a 31% decrease in crime and offending, based on seven measured outcomes across three studies.

The team identified seven outcomes measuring **crime and offending** specifically across three studies: 2 Type A, 1 Type D giving an overall security rating of **Level 1**. Crime outcomes included having charges against the individual and breaking rules (Burke & Loeber, 2015), a juvenile arrest index and self-reported offence index (Conduct Problems Prevention Research Group, 2010b), and violent offending (Feindler et al., 1984).

A total of $k = 7$ outcomes were included in the analysis. The estimated average effect based on the multilevel random-effects model was $\hat{\mu} = -0.26$, (95% CI: -0.52, -0.00).

This estimate was statistically significantly different from zero at the conventional level, t = -1.99, p = 0.046, although the effect was not statistically significant when adjusting for clustering across studies using robust variance estimation (t = -2.18, df = 1.74, p = 0.179).

Although the estimate remained the same (g = -0.26), the increases in the p-value and standard error suggest that within-study correlation was underestimated in the original model.

Table 3: *RVE Output for meta-analysis on crime and offending outcomes*

	Estimate	SE	t-stat	d.f (Satt)	p-val (Satt)	Sig	YEF impact rating	YEF evidence security rating
Intercept	−0.26	0.12	−2.18	1.74	0.179	N.S.	High	1

In practical terms, this indicates that the intervention had a high impact on crime and offending-related outcomes across studies. Based on YEF impact categorisation, the effect size (g = -0.26) corresponds to a "high impact" rating, and this effect was statistically significant. After adjusting for the clustering of outcomes within the three studies using robust variance estimation, the result was no longer statistically significant (p = 0.179). Although the direction and magnitude of the effect are promising, the limited number of studies and low evidence security rating (Level 1) mean that strong conclusions cannot be drawn.

According to the Q-test, moderate heterogeneity was present I^2 = 49.6%, Q(df = 6) = 11.91, p = 0.064. This level of variability suggests that the study results are fairly consistent across the three studies.

The SMD of -0.26 corresponds to a relative risk reduction of 31.1% in the risk of being involved in crime and offending, with an absolute risk reduction of 7.8%. Among those who participate in Targeted SEL, 17% go on to be involved in crime or offending, compared to 25% of those who receive business-as-usual (Figure 2).
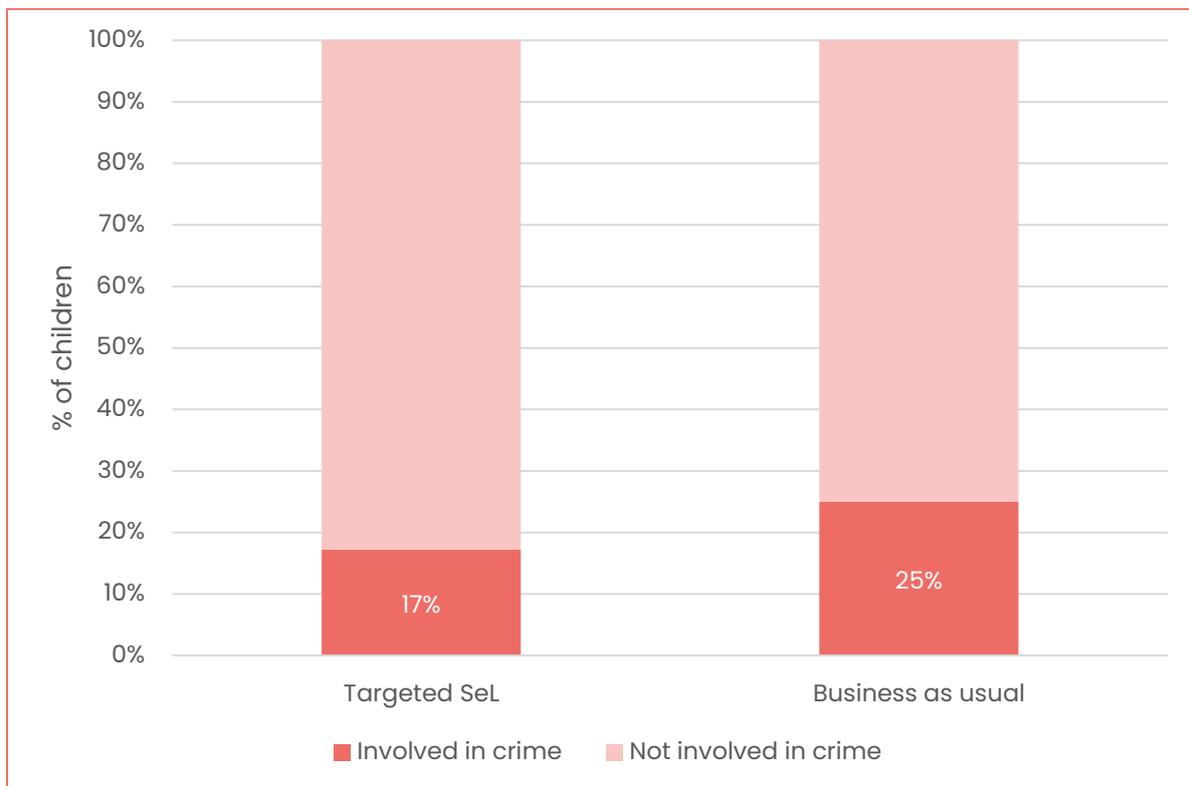
Figure 2: *Absolute risk reduction in crime and offending associated with Targeted SEL interventions compared to business as usual*

Meta-analysis of all related outcomes in Targeted SEL interventions

Including the violence, crime, and offending outcomes described above, the team extracted a total of 135 measured outcomes from 26 effectiveness studies that align with the YEF outcomes framework. These studies included 2 Type A, 17 Type C, 7 Type D, giving an overall security rating of **Level 3**.

A total of $k = 135$ outcomes were included in the analysis. The estimated average effect based on the random-effects model was $\hat{\mu} = -0.49$, (95% CI: -0.72, -0.27).

This estimate was statistically significantly different from zero, *t* = -4.30, *p* <0.001, and remained large and statistically significant when adjusting for clustering across studies using robust variance estimation (t = -4.3, df = 24.3, p <0.001). The estimate remained the same (g = -0.49) suggesting the original meta-analysis was stable.

Table 4: *RVE Output for meta-analysis of all related outcomes*

	Estimate	SE	t-stat	d.f (Satt)	p-val (Satt)	Sig	YEF impact rating	YEF evidence security rating
Intercept	-0.49	0.12	-4.3	24.3	<0.001	***	High	3

In practical terms, these findings indicate that the intervention had a high and statistically robust impact across multiple relevant YEF outcomes including violence, crime, and offending. The pooled effect size (g = -0.49) corresponds to a "high impact" rating under the YEF categorisation framework. This effect was statistically significant in the multivariate random-effects meta-analysis and remained large and statistically significant after adjusting for clustering of outcomes within studies using robust variance estimation (p < .001), indicating that the result is stable and unlikely to be driven by modelling assumptions.

According to the Q-test, the true outcomes appear to be extremely heterogeneous; Q(134) = 783.09, p < 0.001, I^2 = 82.9%. Such significant variability suggests that the differences in effect sizes are not solely due to random chance but may be influenced by specific study characteristics or contexts. To explore potential sources of this heterogeneity, conducting moderator analyses is recommended.

Subgroup analysis

Outcome type

To explore whether programme effectiveness varied by outcome category, subgroup meta-analyses were conducted (Table 5).

Table 5: *Subgroup analysis on outcome domains*

Outcome Domain	k	n	SMD (robust SE)	p-value	YEF impact rating
Ability to resolve conflicts	3	2	–0.37 (0.51)	0.606	High
Behavioural difficulties	50	23	–0.37 (0.13)	0.010*	High
Building & maintaining relationships (social-emotional)	33	12	–0.30 (0.09)	0.021*	High
Bullying	4	1	–0.25 (0.02)	n/a	
Crime and offending	7	3	–0.35 (0.13)	0.119	High
General mental health	14	6	–0.49 (0.20)	0.109	High
Helping others (prosocial behaviour)	4	3	–0.59 (0.10)	0.034*	High
Regulating and managing emotions	5	3	–0.17 (0.28)	0.611	Moderate
School engagement	14	7	–1.20 (0.45)	0.047*	High
Self-esteem	1	1	+0.13 (0.00)	n/a	

Between-group heterogeneity (Q_n test); 10 outcome domains; $\chi^2(9) = 64.1$, p < .001 ***

Note. Estimates computed using robust variance estimation (RVE) with CR2 adjustment for dependent effect sizes (Tipton, 2015). Subgroup-specific p-values are based on robust t-tests with Satterthwaite small-sample degrees of freedom. The omnibus test of differences across outcome domains was conducted using a CR2-robust Wald chi-square test. Significance levels: p < .05 (*), p < .01 (**), p < .001 (***).

Targeted SEL interventions demonstrate statistically significant beneficial effects in several outcome domains. For example, in the domain of school engagement (SMD = −1.20, k = 14, p <0.05), Targeted SEL interventions showed a very large improvement, meaning the intervention is associated with substantially better school engagement compared with the business-as-usual control condition.

In building and maintaining relationships (k = 33, n = 12), Targeted SEL programmes showed a significant improvement (SMD = −0.30, p <0.05).

Similarly, statistically significant improvements favouring Targeted SEL interventions were observed for behavioural difficulties (SMD = −0.37, k = 50, p <0.05) and helping others (prosocial behaviour; SMD = −0.59, k = 4, p <0.05). These negative SMD values indicate that young people receiving Targeted SEL support showed lower levels of behavioural difficulties and stronger positive outcomes in these domains compared to control groups.

Finally, a Q-test for between-group heterogeneity indicated that the impact of Targeted SEL interventions varies significantly across outcome domains. A robust Wald chi-square test indicated significant between-group heterogeneity across outcome domains ($\chi^2(9) = 64.1$, p < .001), providing strong evidence that effect sizes differ by type of outcome measured. In practical terms, this means the effectiveness of Targeted SEL interventions is not uniform: these programmes appear to be more effective in improving certain outcomes (such as social relationships, behaviour, prosociality, and school engagement) than others.

Country

To investigate potential differences in the effectiveness of Targeted SEL programmes, subgroup analyses were conducted by country (Table 6).

Table 6: *Subgroup analysis on country*

Country	k	n	SMD (robust SE)	p-value	YEF impact rating
USA	116	21	-0.47 (0.12)	0.002 **	High
Netherlands	10	1	-0.07 (0.00)		n/a
Canada	7	3	-0.58 (0.04)	0.004 **	High
Turkey	2	1	-0.08 (0.00)		n/a
Between-group heterogeneity (Q_n test) 4 countries; $\chi^2(3) = 29k$ p < 0.001 ***					

Significant variation in programme effects emerged across countries. Targeted SEL interventions showed the strongest average effects in Canada (SMD = −0.58, p < 0.005) and the USA (SMD = −0.47, p < 0.005), reflecting high improvements in outcomes relative to comparison groups. Smaller but still statistically significant effects were observed for the Netherlands (SMD = −0.07, p < 0.001) and Turkey (SMD = −0.08, p < 0.001). However, these estimates each came from a single study.

The omnibus test indicated strong evidence of between-country heterogeneity ($\chi^2(3) = 29k$, p < .001), meaning effects differed substantially across national contexts. In practical terms, Targeted SEL interventions appear to yield larger benefits in North American samples than in the single-study European or Turkish samples included in the dataset.

Publication year

Subgroup meta-analyses were conducted by decade to explore variation in effect sizes over time, given the wide range of publication years across included studies (Table 7).

Table 7: *Subgroup analysis on publication year*

Publication period	k	n	SMD (robust SE)	p-value	YEF impact rating
1970–1979	10	3	−1.89 (1.18)	0.281	High
1980–1989	41	8	−0.55 (0.10)	0.002 **	High
1990–1999	22	5	−0.25 (0.12)	0.202	High
2000–2009	29	5	−0.29 (0.28)	0.393	High
2010–2019	33	5	−0.29 (0.16)	0.202	High
Between-group heterogeneity (Q_n test); 5 periods; $\chi^2(4) = 1.56$; $p = 0.182$					

Across publication-year subgroups, only one period, the 1980–1989 cohort, showed a statistically significant effect (SMD = −0.55, p < 0.01), indicating strong benefits of Targeted SEL interventions for studies published in this decade. Effects for other publication periods were not statistically significant, with wide confidence intervals especially for earlier decades (e.g., 1970–1979), reflecting limited data (k = 10, n = 3 studies).

The omnibus heterogeneity test was not statistically significant ($\chi^2(4) = 1.56$, p = 0.182), suggesting no strong evidence that intervention effects differ systematically by publication year. In practical terms, this means that although

the 1980s subgroup shows a statistically detectable effect, overall, there is no robust pattern indicating that Targeted SEL intervention effectiveness has changed meaningfully across decades.

Age band

Subgroup meta-analyses were conducted on the age of the intervention participants was conducted to evaluate if age impacted effectiveness. To examine whether programme effectiveness varied by age, studies were grouped into three age-band categories: 0–9 years (Pre-school and Early Childhood/Early Primary), 10–14 years (Middle Childhood/Early Adolescence), and 15–17 years (Adolescence/Middle to High School). Studies that included participants spanning more than one age band were classified as Mixed age groups (Table 8).

Table 8: *Subgroup analysis on developmental age band*

Age band	k	n	SMD (robust SE)	p-value	YEF impact rating
Mixed	61	13	-0.49 (0.11)	0.006 **	High
0–9 (Pre-school & Early Childhood / Early Primary)	54	9	−0.21 (0.11)	0.130	Moderate
10–14 (Middle Childhood / Early Adolescence)	14	3	-0.34 (0.21)	0.281	High
15–17 (Adolescence / Middle/High-school)	6	1	-3.23 (0.43)	n/a	
Between-group heterogeneity (Q_n test); 4 age bands; $\chi^2 (3) = 15.7$; $p < .001$ ***					

Targeted SEL interventions showed varying effects across different age bands. For children aged 0–9 (k = 54 effects; n = 9 studies), impact was moderate and not statistically significant (SMD = −0.21, p = .130). Similarly, for young people aged 10–14 (k = 14; n = 3), impact was high but not statistically significant (SMD = −0.34, p = 0.281).

For older adolescents aged 15–17 (k = 6; n = 1), the estimated effect size was very large (SMD = −3.23), but the p-value was not statistically significant (0.084). This estimate comes from a single study and was very large and so this result should be interpreted with considerable caution.

Studies including mixed-age samples (k = 61; n = 13) showed high impact according to the YEF framework (SMD = −0.49, p < 0.01).

The Wald χ^2 test indicated significant heterogeneity across age bands ($\chi^2(3)$ = 15.7, p < 0.001), meaning that the effectiveness of Targeted SEL interventions differs meaningfully depending on the age group targeted.

Ethnicity

To explore whether programme effectiveness varied by ethnicity, studies were grouped into three categories: Majority white (>85% white participants), Some diversity (15–49% Black and Global Majority), and Diverse/balanced (≥50% Black and Global Majority) (Table 9).

Table 9: *Subgroup analysis on ethnicity*

Ethnicity	k	n	SMD (robust SE)	p-value	YEF impact rating
Not stated/unclear	51	11	−0.38 (0.10)	0.014 *	High
Diverse/balanced	35	8	−0.78 (0.32)	0.102	High

Some diversity	30	4	−0.06 (0.12)	0.644	Small
Majority white	19	3	−0.50 (0.22)	0.156	High
Between-group heterogeneity (Q_n test); 4 subgroups; χ^2 (3) = 2.56; p = 0.053					

Ethnic composition of the sample did not emerge as a strong moderator of Targeted SEL intervention effects. All subgroups showed beneficial effects favouring Targeted SEL programmes, however, studies in which the sample's ethnicity was not stated or unclear (k = 51 effects; 11 studies) showed a significant average effect (SMD = −0.38, p < 0.05). Effects in the other categories were not statistically significant, likely reflecting small degrees of freedom for this type of robust test.

The omnibus heterogeneity test was non-significant across the four ethnicity categories, χ^2(3) = 2.56, p = 0.053. In practical terms, this suggests that intervention effectiveness likely does not differ meaningfully by the ethnic composition of the study samples, but underreporting remains an issue.

Gender

To investigate whether the effectiveness of Targeted SEL programmes varied according to the gender composition of the study samples, subgroup analyses were conducted across five categories: studies with only male participants, majority male samples, only female participants, a gender balance, and those with no reported gender information (Table 10).

Table 10: *Subgroup analysis on gender*

Sample gender composition	k	n	SMD (robust SE)	p-value	YEF impact rating
Mixed sex	56	10	−0.33 (0.22)	0.190	High
Boys only	26	6	−0.59 (0.13)	0.056	High

Majority boys	27	7	−0.60 (0.11)	0.005 **	High
Not stated/unclear	22	2	−0.30 (0.19)	0.354	High
Girls only	4	1	−0.25 (0.02)	n/a	
Between-group heterogeneity (Q_n test); 5 gender groupings; $\chi^2 (4) = 4.07$; p = 0.003 **					

Gender composition significantly moderated the impact of Targeted SEL interventions. Significant beneficial effects were observed in the majority-boys' samples (SMD = −0.60, p < 0.005) and the girls-only sample (SMD = −0.25, p < 0.05), however, the finding with girls was a smaller effect and based on just one study.

In the category of boys-only groups, effects were sizable but non-significant at the 0.05 level (SMD = −0.59, p = 0.056). No significant effects were observed for mixed-sex samples or for studies without clear gender reporting. In simple terms, effects are largest and most consistent when boys make up most or all of the sample. Although the girls-only sample was also significant, this shows a much smaller effect (−0.25) and the evidence base is much thinner (being based on only one study).

A robust Wald chi-square test indicated significant between-group heterogeneity across gender categories ($\chi^2(4) = 4.07$, p < 0.01), meaning the effectiveness of Targeted SEL interventions differs meaningfully depending on the gender composition of participating children and young people.

Multiple meta-regressions (models 1-5)

Following these five subgroup analyses, we ran five separate random-effects meta-regression models with robust variance estimation to examine pre-specified moderators: (1) quality, (2) setting, (3) intervention features, (4) population characteristics, and (5) outcome characteristics. See Appendix 1 for an overview of the methods used in this section and Appendix 3 for a list of the studies that provided data for these analyses.

Moderator Analysis 1 – Quality Moderators

The characteristics in model 1 include:

- Study design
- Study timing
- Unit of allocation into group
- Method of allocation into group
- Quality appraisal as assessed by the YEF-EQA tool

***Table 11:** Meta-regression (with robust inference) on quality moderators*

Moderator	β (estimate)	SE	t	df	p	95% CI
Intercept	−0.47	0.15	-3.15	1.00	0.20	−2.38, 1.43
Design = individual RCT (vs reference)	−0.11	0.23	-0.48	1.10	0.71	−2.48, 2.26
Design = QED (non-randomised)	0.45	0.22	2.06	1.06	0.28	−1.98, 2.88
Study timing = prospective	−0.05	0.25	−0.20	1.07	0.87	−2.85, 2.74
Study quality (YEF EQA) = High	0.36	0.24	1.51	1.12	0.35	−1.99, 2.70
Study quality (YEF EQA) = Low	0.24	0.22	1.08	12.07	0.30	−0.25, 0.73

Note. Values are from a multivariate random-effects meta-regression estimated using REML (k = 135), clustering on study ID to account for non-independence of multiple effect sizes within studies. Statistical inference is based on CR2-robust standard errors with Satterthwaite small-sample t-tests.

The estimated between-study variance was τ^2 = 0.36 (SD = 0.60), indicating substantial heterogeneity across studies. Residual

heterogeneity remained highly significant (QE(df = 129) = 749.91, p < .0001). The omnibus test of moderators was not significant (QM(df = 5) = 1.74, p = 0.884), indicating that study design, study timing, and YEF EQA quality ratings did not meaningfully explain heterogeneity in effect sizes in this model (Table 11).

Moderator Analysis 2 – Setting-level Moderators

The characteristics in model 2 include:

- Country of intervention
- Setting of intervention

Table 12: *Meta-regression (with robust inference) on setting-level moderators*

Moderator	β (estimate)	SE	t	df	p	95% CI
Intercept	−0.49	0.33	−1.49	2.40	0.25	−1.70, 0.72
Country = Netherlands (vs reference)	0.10	6.54	0.02	2.25	0.99	−25.20, 25.40
Country = Turkey (vs reference)	−0.10	0.24	−0.40	2.40	0.72	−0.97, 0.78
Country = USA (vs reference)	−0.17	0.23	−0.73	1.40	0.57	−1.71, 1.37
Intervention setting = Community (vs reference)	0.32	6.54	0.05	1.86	0.97	−29.96, 30.59
Intervention setting = Not specified (vs reference)	−0.17	0.32	−0.52	2.24	0.65	−1.41, 1.08
Intervention setting = School (vs reference)	0.10	0.30	0.34	1.31	0.78	−2.13, 2.34

Intervention setting = School and home (vs reference)	0.50	0.24	2.10	2.00	0.17	−0.53, 1.53

The estimated between-study variance component was τ^2 = 0.39 (SD = 0.63), indicating substantial variability in effect sizes across studies. Residual heterogeneity remained highly significant (QE(df = 127) = 709.03, p < .0001), and the omnibus test of moderators was non-significant (QM(df = 7) = 2.09, p = 0.955).

In simple terms, neither country nor intervention setting meaningfully explained heterogeneity in effect sizes in this model. The large robust standard errors and very small Satterthwaite degrees of freedom for several moderators reflect limited information within clusters, suggesting that these moderator effects are estimated with considerable uncertainty.

Moderator Analysis 3 – Intervention-level Moderators

The characteristics in model 3 include:

- Intervention intensity
- Intervention duration
- Intervention provider
- Training for intervention providers
- Delivery method
- Group or individual intervention

Table 13: *Meta-regression (with robust inference) on intervention-level moderators*

Moderator	β (estimate)	SE	t	df	p	95% CI
Intercept	−1.56	0.74	−2.12	3.18	0.12	−3.83, 0.71
Duration > 5 years	−0.86	1.26	−0.68	3.33	0.54	−4.65, 2.93

Duration = 1–3 months	0.63	0.83	0.76	3.62	0.49	−1.77, 3.04
Duration = 3–6 months	0.82	1.04	0.79	4.07	0.47	−2.05, 3.69
Duration = 6–12 months	−0.11	0.83	−0.13	2.99	0.90	−2.77, 2.54
Duration = unclear	0.53	4.83	0.11	4.51	0.92	−12.29, 13.35
Intensity = daily	−0.05	0.61	−0.08	3.54	0.94	−1.82, 1.73
Intensity = not stated/ unclear	0.61	0.83	0.73	2.36	0.53	−2.48, 3.69
Intensity = weekly	1.00	0.40	2.45	2.82	0.10	−0.34, 2.30
Implementer training = yes (pre-intervention)	−1.51	1.00	−1.51	3.71	0.21	−4.38, 1.35
Implementer training = yes (manual/ qualification)	0.64	0.83	0.78	2.35	0.51	−2.45, 3.74
Implementer training = yes (ongoing supervision)	−0.95	0.65	−1.47	4.29	0.21	−2.70, 0.80
Delivery method = face-to-face and phone	−1.47	1.06	−1.39	3.44	0.25	−4.60, 1.66

Delivery method = not specified	0.04	0.30	0.13	3.92	0.91	−0.79, 0.87
Provider = not stated	0.58	0.48	1.23	2.88	0.31	−0.96, 2.13
Provider = researcher	0.34	0.74	0.47	2.51	0.68	−2.29, 2.97
Provider = service provider/unspecified leader	−0.29	0.52	−0.55	2.59	0.63	−2.10, 1.53
Provider = social worker	−0.49	0.65	−0.75	2.58	0.51	−2.74, 1.77
Provider = teacher/ education professional	0.51	0.46	1.11	2.71	0.36	−1.04, 2.05
Group & individual delivery	2.16	1.22	1.77	2.94	0.18	−1.76, 6.08
Individual delivery only	1.24	0.79	1.57	3.03	0.21	−1.26, 3.73
Group/individual = not specified	−0.38	0.77	−0.49	1.90	0.68	−3.88, 3.13

The estimated between-study variance component was $\tau^2 = 0.22$ (SD = 0.47), indicating substantial heterogeneity across studies (Table 13). Residual heterogeneity remained statistically significant ($Q_E(df = 113) = 487.75$, $p < .0001$). The omnibus test of moderators was not statistically significant ($Q_M(df = 21) = 30.61$, $p = 0.080$), suggesting that intervention duration, intensity, implementer training, delivery method, provider type, and delivery format do not collectively explain heterogeneity in effect sizes. At the individual level, weekly

intensity showed a marginal association with effect size, but this did not remain statistically significant under robust inference.

Moderator Analysis 4 – Population-level Moderators

The characteristics in model 4 include:

- Age band
- Ethnicity
- Gender
- Socio-economic status

Table 14: *Meta-regression (with robust inference) on population-level moderators*

Moderator	β (estimate)	SE	t	df	p	95% CI
Intercept	−0.46	0.16	−2.83	4.42	0.04	−0.89, −0.03
Age band = 10–14 (vs reference)	−0.003	0.47	−0.01	3.72	1.00	−1.36, 1.35
Age band = 15–17 (vs reference)	−3.04	1.13	−2.70	6.15	0.04	−5.78, −0.30
Age band = Mixed	−0.26	0.11	−2.30	6.84	0.06	−0.53, 0.01
Ethnicity = Majority white	−0.08	0.36	−0.22	4.64	0.84	−1.01, 0.86
Ethnicity = Not stated/unclear	−0.11	0.32	−0.34	3.85	0.76	−1.01, 0.79
Ethnicity = Some diversity	0.02	0.40	0.05	4.95	0.97	−0.99, 1.03

Moderator	β (estimate)	SE	t	df	p	95% CI
Gender = Girls only	0.45	0.71	0.64	6.12	0.55	−1.28, 2.18
Gender = Majority boys	0.11	0.27	0.40	4.56	0.71	−0.61, 0.83
Gender = Mixed sex	0.39	0.25	1.56	3.48	0.20	−0.35, 1.13
Gender = Not stated/unclear	0.22	0.23	0.96	3.76	0.40	−0.43, 0.87
SES = Not stated/unclear	0.10	0.27	0.37	2.75	0.74	−0.80, 1.00
SES = Some low-income families	−0.23	0.46	−0.50	2.78	0.66	−1.74, 1.29

The estimated between-study variance was τ^2 = 0.10 (SD = 0.31), indicating moderate heterogeneity. Residual heterogeneity remained statistically significant (QE(df = 122) = 557.46, p < .0001). The omnibus test of moderators was significant (QM(df = 12) = 49.33, p < .0001), driven primarily by a large negative association for studies conducted with adolescents aged 15–17. All other age, ethnicity, gender, and socioeconomic status moderators were not statistically significant, suggesting limited explanatory power for heterogeneity beyond this age group effect (Table 14).

Moderator Analysis 5. Outcome Moderators

The characteristics in model 5 include:

- YEF's outcome category
- Methods used to collect the outcome data

Table 15: *Meta-regression (with robust inference) on outcome-level moderators*

Moderator	β (estimate)	SE	t	df	p	95% CI

Intercept	−0.96	0.17	−5.76	6.16	0.001	−1.36, −0.55
Outcome: Regulating and managing emotions	−0.01	0.31	−0.04	1.69	0.97	−1.60, 1.57
Outcome: Ability to resolve conflicts	0.05	0.14	0.40	1.03	0.76	−1.55, 1.66
Outcome: Behavioural difficulties	0.18	0.10	1.76	4.85	0.14	−0.09, 0.45
Outcome: Bullying	0.32	0.48	0.66	1.85	0.59	−1.93, 2.57
Outcome: Crime and offending	0.37	0.12	2.96	2.04	0.10	−0.15, 0.88
Outcome: General mental health	0.04	0.11	0.35	4.16	0.74	−0.26, 0.34
Outcome: Prosocial behaviour	−0.21	0.15	−1.34	1.87	0.32	−0.92, 0.50
Outcome: School engagement	−0.04	0.15	−0.30	3.79	0.78	−0.46, 0.37
Outcome: Self-esteem	0.26	0.04	6.11	1.29	0.07	−0.06, 0.59
Data source: Not stated	1.03	0.28	3.70	3.59	0.03	0.22, 1.84
Data source: Observation	0.26	0.16	1.69	2.81	0.20	−0.25, 0.78

Data source: Parent report	0.34	0.23	1.46	1.34	0.34	−1.32, 2.00
Data source: Peer report	0.72	0.20	3.60	3.49	0.03	0.13, 1.31
Data source: Self-completion	0.68	0.52	1.33	1.66	0.34	−2.02, 3.39
Data source: Self report	0.22	0.13	1.73	1.80	0.24	−0.39, 0.83
Data source: Teacher report	0.51	0.20	2.60	4.12	0.06	−0.03, 1.05

The estimated between-study variance was $\tau^2 = 0.28$ (SD = 0.52). Residual heterogeneity remained substantial and statistically significant ($QE(df = 118) = 634.91$, $p < .0001$). The omnibus test of moderators was significant ($QM(df = 16) = 42.78$, $p = .0003$), indicating that outcome type and data source jointly explain some heterogeneity, although most individual moderators were not statistically significant after robust adjustment (Table 15).

Publication bias

A funnel plot of the effect size estimates is shown in Figure 3. The rank correlation test (Begg's test) indicated significant funnel plot asymmetry (Kendall's $\tau = -0.31$, $p < .0001$), and the regression test (Egger-type test) likewise provided strong evidence of asymmetry ($z = -7.44$, $p < .0001$). Together, these findings suggest the presence of small-study effects consistent with potential publication bias.

Funnel Plot (All outcomes)
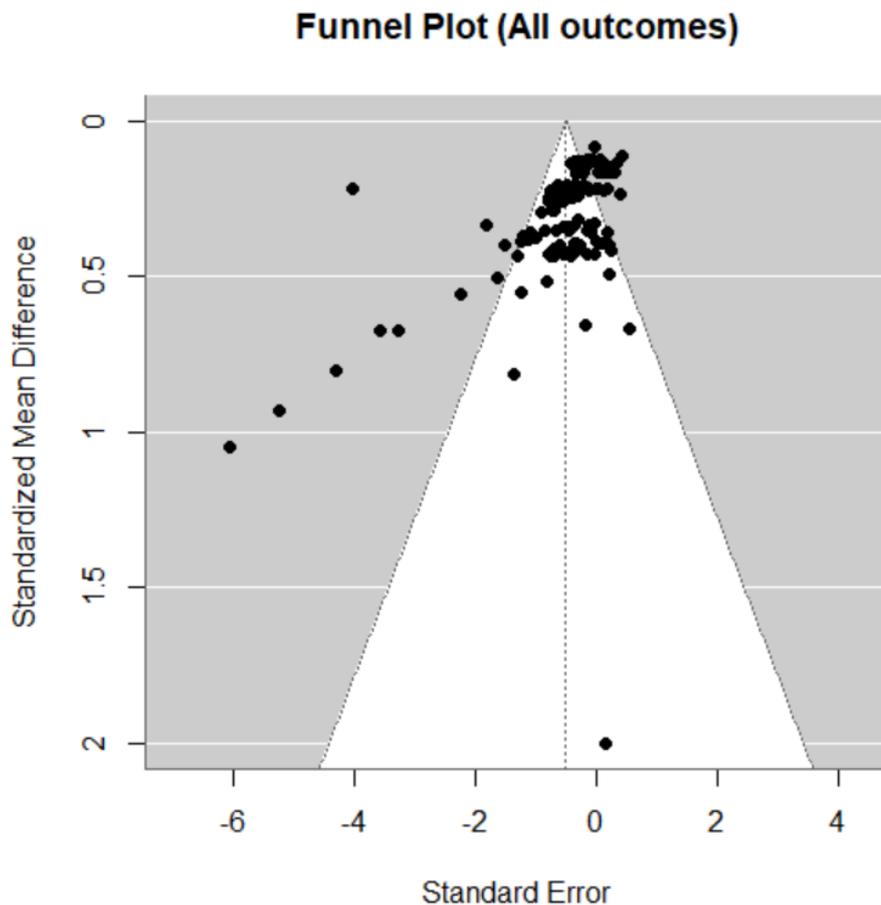
Figure 3: *Funnel plot of all outcomes*

To further examine the impact of publication bias on the pooled effect, a trim-and-fill analysis was conducted. The procedure estimated no missing studies on the right side of the funnel plot (estimated number missing = 0, SE = 5.91), indicating that asymmetry was not attributable to an absence of small studies with effects in the opposite direction. The adjusted random-effects model therefore yielded results identical to the original model.

Under the trim-and-fill model (k = 135; REML estimator), the pooled effect size remained statistically significant (estimate = −0.43, SE = 0.06, z = −7.82, p < .0001, 95% CI [−0.54, −0.32]). Substantial heterogeneity was observed (τ^2 = 0.32, SE = 0.05; I^2 = 87.83%), and the test for residual heterogeneity was significant (Q(134) = 783.09, p < .0001).

Overall, although the funnel plot looked uneven at first, additional checks found no clear evidence that publication bias had distorted the overall result. The uneven pattern is more likely to reflect differences between the studies than missing evidence.

Sensitivity analyses

An examination of the studentized residuals revealed that several studies had values larger than ±3.5603 and may be potential outliers in the context of this model. Similarly, according to the Cook's distances, several studies could be considered to be overly influential (Figure 4).

Figure 4: *Influence diagnostics for individual studies*

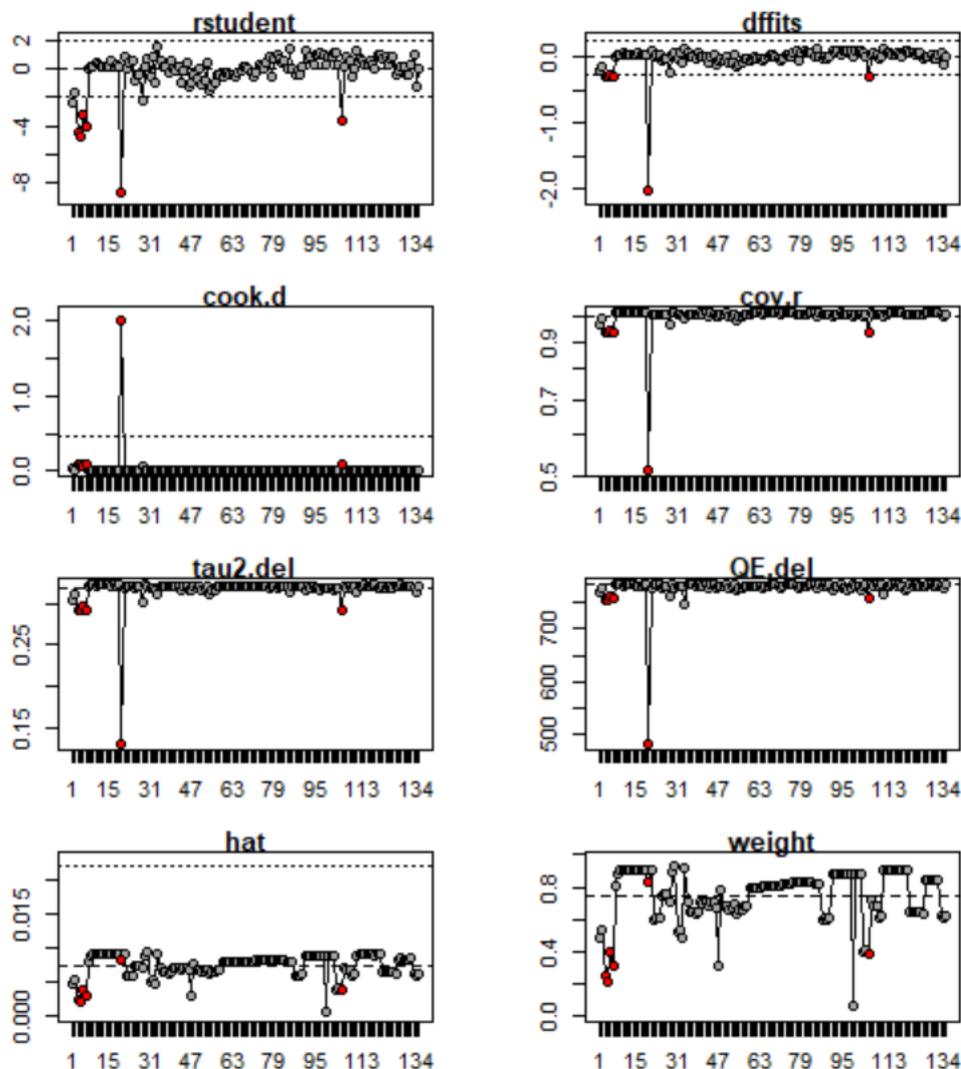To assess the robustness of the overall effect estimate to these observations, a series of sensitivity analyses was conducted. First, analyses were repeated after excluding studies rated as low quality. The overall pooled effect remained statistically significant and of similar magnitude to the main analysis (SMD = -0.49, SE = 0.12, p < .001), indicating that the primary findings were not driven by lower-quality studies.

Second, the analysis was restricted to randomized controlled trials (RCTs and cRCTs) only, excluding quasi-experimental designs. The pooled effect estimate in this subset was nearly identical to the main result and remained statistically significant (SMD = -0.51, SE = 0.12, p < .001), suggesting that study design did not materially influence the overall conclusion.

Third, separate analyses were conducted for post-test and follow-up outcomes. At post-test, the intervention effect remained statistically significant and comparable in size to the overall estimate (SMD = -0.47, SE = 0.10, p < .001). In contrast, the pooled effect at follow-up was larger in magnitude but imprecisely estimated and not statistically significant (SMD = -1.13, SE = 0.61, p = .163), likely reflecting the small number of follow-up studies and reduced statistical power.

Finally, a leave-one-study-out analysis was performed to evaluate the influence of individual studies on the pooled effect (Figure 5). Removal of each study in turn produced pooled effect estimates that ranged from approximately -0.38 to -0.53, with all confidence intervals excluding zero. No single study materially altered the magnitude or statistical significance of the overall effect, indicating that the findings were not unduly driven by any individual study.
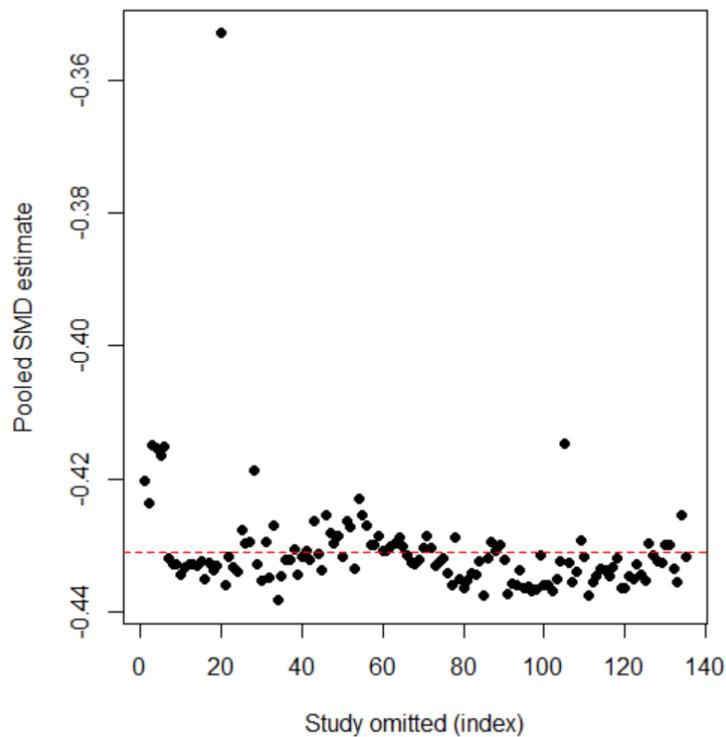
Figure 5: *Leave-one-out sensitivity analysis*

Overall, these sensitivity analyses demonstrate that the main results are robust to the exclusion of potentially influential or lower-quality studies, restrictions to RCTs, outcome timing, and the removal of individual studies.

How Secure is the Evidence?

Violence outcomes

Our confidence in the findings on the prevention and reduction of violence is **Low**[6]. The meta-analysis included 15 violence-related outcomes drawn from nine studies that assessed the impact of Targeted SEL on children and young people.

The studies included:

- One cluster RCT rated as moderate quality **(Type C)**
- Eight RCTs: of these, four were rated as moderate quality **(Type C)** and four as low quality **(Type D)**

As a result, a **Level 2** evidence security rating was applied.

Heterogeneity was moderate (I^2 = 31.4%), meaning the **Level 2** evidence security rating was maintained.

Crime and offending outcomes

Our confidence in the findings on crime and offending is **Very Low**. The meta-analysis included seven crime and offending related outcomes drawn from three studies that assessed the impact of Targeted SEL on children and young people.

The studies included:

- One cluster RCT rated as high quality **(Type A)**
- Two RCTs: of these, one was rated as high quality **(Type A)** and one as low quality **(Type D)**

Due to the limited number of studies, the evidence-base is thin. As a result, a **Level 1** evidence security rating was applied. Heterogeneity was moderate (I^2 = 49.6%), meaning the **Level 1** evidence security rating was maintained.

---

[6] The process used to determine evidence security ratings is set out in the technical guide.

All outcomes

Our confidence in the findings on all outcomes is **Moderate**. The meta-analysis included 135 outcomes drawn from 26 studies that assessed the impact of Targeted SEL on children and young people.

Study quality, as assessed by the YEF-EQA, ranged from low to high. The studies included:

- Three cluster RCTs: of these, one was rated as high quality **(Type A)** and two as moderate quality **(Type C)**
- 22 RCTs: of these, one was rated as high quality **(Type A),** 14 as moderate quality **(Type C)**, and seven as low quality **(Type D).**
- One QED rated as moderate quality **(Type C)**

As a result, a **Level 3** evidence security rating was applied.

Heterogeneity was high (I^2 = 82.9%). However, subsequent moderator analyses indicate that outcome-level characteristics collectively account for a significant proportion of the observed heterogeneity in effect sizes, indicating that this variation is systematic rather than unexplained. Results remain consistent and robust across sensitivity analyses, robust variance estimation, and methodological differences (e.g., study design, timing of follow-up, and data source), reinforcing confidence in the overall statistically significant effects of Targeted SEL programmes. As such, a **Level 3** evidence security rating was maintained.

What factors affect implementation?

This section summarises implementation-relevant insights from previous systematic reviews, including factors associated with stronger effects in prior syntheses. These findings are presented as contextual evidence and may differ from the moderator analyses in our effectiveness meta-analysis due to differences in included studies, outcome definitions, and analytic approach.

Four systematic reviews were included (see Appendix 5 for details of the reviews). Using AMSTAR 2, two were rated as Low quality (Blewitt et al., 2019; Murano et al., 2020) and two as Critically Low (Goldberg et al., 2019; Yang et al., 2019).

Implementation Processes and Delivery Capacity

Fidelity

Three systematic reviews highlight fidelity, defined as the extent to which an intervention is delivered as intended, as a critical implementation factor in Targeted SEL interventions. Blewitt et al. (2019) found that most studies reported embedding fidelity checks into Targeted SEL interventions by using structured approaches such as observations and checklists, with strong fidelity associated with clear training, standardised protocols and regular monitoring. In contrast, weak or absent fidelity reporting often reflected limited supervision of intervention providers and poor documentation of how interventions were implemented, reducing confidence that interventions were delivered as intended. Murano et al. (2020) similarly identified substantial gaps in fidelity reporting across Targeted SEL studies, noting that insufficient data prevented analysis of fidelity as an outcome moderator. This highlights a key implementation challenge: without systematic monitoring and reporting, it is difficult to determine whether interventions are implemented as intended or if there is variability in delivery.

Yang et al. (2019) explored the contribution of intervention fidelity to the effectiveness of social-emotional interventions. The authors found that interventions with documented good fidelity produced substantially stronger effects on reducing negative behaviours than those with unknown fidelity. This indicates that consistent, high quality implementation is a key mechanism driving

success. However, Yang et al.'s (2019) findings must be interpreted with caution; whilst 15 studies were clearly Targeted SEL interventions, 14 aimed to develop social emotional competencies in children and young people but lacked a specific affective component.

Taken together, these findings highlight the importance of embedding fidelity supports, such as training, monitoring and clear delivery guidance, within intervention planning and implementation. This helps ensure that Targeted SEL interventions are delivered consistently and achieve their intended benefits.

Intervention Providers and Training

Implementation quality can be influenced by how Targeted SEL interventions are delivered in practice, including the preparedness, training and support of those responsible for delivery. Blewitt et al.'s (2019) review of 19 Targeted SEL interventions highlights important implementation variation in *who* delivers programmes and *how* delivery is supported. Most interventions were implemented by classroom teachers, with others led by trained facilitators or researchers, but there was wide inconsistency in the training, preparation and support provided to teachers responsible for delivery. Teacher training ranged from minimal or one-off instruction to structured professional development with ongoing coaching, indicating substantial differences in implementation support across studies.

Blewitt et al. (2019) highlight that Targeted SEL interventions where teachers received substantial training (e.g., structured professional development and ongoing coaching) and ongoing support, showed the strongest and most consistent improvements in children and young people's social skills, behavioural regulation, and teacher–child relationship quality. This emphasises that implementation of Targeted SEL interventions requires deliberate investment in training, coaching and ongoing support to enable consistent, high-quality delivery.

Family Engagement and Involvement

Family engagement/involvement shapes how an intervention is implemented, supported, and sustained across settings. Decisions about whether and how families are involved affect the contexts in which skills are practised and the

coordination required between school and home, all of which influence fidelity, feasibility and consistency of delivery.

The extent to which family members are involved in delivering or supporting the implementation of Targeted SEL interventions differs across reviews. Murano et al. (2020) found that 11 of the 15 Targeted SEL interventions reviewed were implemented primarily through parent-led, home-based models, with parents acting as the main intervention providers following structured training. Training provided to parents focused on how to deliver new interaction and behaviour-support strategies directly to their children in the home-environment. In contrast, Blewitt et al. (2019) found that parental involvement was rarely embedded in Targeted SEL interventions, with only three of 19 studies reporting any form of parent engagement. Where included, parents were supported through workshops, home visits and take-home materials, this appeared to support generalisation of skills beyond a school environment. However, evidence on the added benefit of family involvement was limited and largely descriptive.

Taken together, these findings highlight that family involvement in the implementation of Targeted SEL intervention is variable, but where embedded it requires deliberate planning and support. For practice, this suggests that when parents are expected to contribute meaningfully to the delivery of Targeted SEL interventions, particularly in home-based models, implementation should include structured training, guidance and ongoing support to enable parents to implement strategies consistently and with fidelity.

Delivery Characteristics

Delivery Method

Delivery method refers to *how* an intervention's content is delivered and experienced in practice, shaping participant engagement, skill practice and consistency of delivery. Consequently, decisions about delivery method are central to implementation planning, as they determine whether an intervention can be implemented consistently, at sufficient quality, and in a way that maximises its intended impact.

Blewitt et al. (2019) found that Targeted SEL interventions were most commonly implemented through one-to-one and small-group delivery formats, allowing programmes to be tailored to the specific social, emotional and behavioural needs of participating children and young people. Small-group approaches supported peer interaction, role-play and collaborative learning, while still enabling intervention providers to provide individualised support, making them a flexible and developmentally appropriate delivery method for young children.

Indeed, Goldberg et al. (2019) found that SEL interventions which provided additional needs-focused and targeted support to children and young people experiencing emotional or behavioural difficulties produced stronger social-emotional and behavioural adjustment effects than interventions which only used whole-school approaches. This highlights the importance of implementing needs-focused provision in one-to-one or small group delivery formats for children and young people at elevated risk of emotional or behavioural difficulties.

From an implementation perspective, these findings suggest that delivery format is a key consideration when planning Targeted SEL interventions. One-to-one and small-group models require sufficient staffing and training, but they enhance engagement and facilitate responsive, needs-led delivery. This highlights the importance of aligning delivery formats with intervention aims and child needs when implementing Targeted SEL interventions in practice.

Duration/Dosage

Dosage typically refers to the amount of content or number of sessions delivered (e.g., sessions completed, hours of exposure), while duration refers to the overall length of time across which the intervention is implemented (e.g. weeks or months). Duration and dosage influence the fidelity of interventions as they determine how much of the intended intervention content is actually delivered in practice, influencing participants' exposure to and practice of core components.

Yang et al. (2019) identified intervention duration as an important structural implementation factor in targeted interventions which aimed to improve social emotional competencies of children and young people, distinguishing between shorter interventions (less than one school year) and longer interventions (one

school year or more). Yang et al. (2019) found that shorter, more focused Targeted SEL interventions (less than one school year) were associated with stronger improvements in positive social-emotional outcomes and reductions in negative behaviours than longer interventions. However, this advantage appears closely linked to the curriculum focus, meaning findings must be treated with extreme caution, as shorter programmes were more clearly SEL-focused while longer interventions lacked affective content (i.e., had lower fidelity to SEL curriculum).

Overall, implementation decisions about duration interact closely with curriculum focus and clarity. For practice, these findings highlight the importance of planning interventions with a realistic timeframe that enables delivery of core SEL curriculum components.

Intervention Curriculum

The curriculum of an intervention shapes what content is delivered and the level of complexity involved during real-world delivery. How the curriculum is implemented in practice, including what is emphasised, simplified, or omitted, affects participants' exposure to core mechanisms of change.

Blewitt et al. (2019) highlighted that Targeted SEL interventions reviewed typically focused on implementing curriculum components aiming to build relationship-based strategies (e.g., strengthening teacher–child interactions) with explicit skill instruction delivered through structured lessons, modelling, role-play, guided practice and reinforcement. Targeted SEL interventions that integrated these components into the curriculum design also tended to use developmentally appropriate formats, such as play-based activities, dyadic or peer-mediated interactions, and embedding SEL skills within everyday academic tasks.

From an implementation perspective, these findings suggest that Targeted SEL interventions are most feasibly and consistently delivered when curricula blend relational strategies with clearly structured teaching methods. Play-based and integrated formats appear to support engagement, enjoyment and acceptability of Targeted SEL interventions, particularly for young children. However, more research is needed to understand how best to implement Targeted SEL curriculum with older children.

Summary of Key Findings from Meta-Review

To briefly summarise, key findings from this section include:

- Fidelity is a critical implementation factor in Targeted SEL interventions, with strong delivery supports (e.g., training, monitoring and clear protocols) linked to more consistent implementation, while poor reporting limits confidence in whether interventions are delivered as intended.
- Most Targeted SEL interventions are implemented by classroom teachers, but there is wide inconsistency in the training, preparation and support provided to those responsible for delivery.
- Providing teachers with substantial training (e.g., structured professional development and coaching) and ongoing support, improve outcomes for children and young people.
- Targeted SEL interventions are most commonly implemented using one-to-one or small-group formats, supporting tailored, responsive and needs-led delivery, but requiring sufficient staffing and training to implement appropriately.
- Ensuring focused delivery of core SEL content matter is covered in the implementation of Targeted SEL interventions appears to be more important than programme length alone.
- Targeted SEL curricula are most feasibly implemented and engaging for younger children when they combine relational strategies with clearly structured skill instruction using developmentally appropriate, play-based and integrated formats.

- Two systematic reviews were rated as low quality and two were rated as critically low quality, meaning the conclusions above should be interpreted with care. Experiences of children and young people

One moderate quality pilot RCT examined the experiences of children and young people who participated in the Growing Interpersonal Relationships through Learning and Systematic Supports programme (GIRLLS; Splett et al., 2015). This 10-week, school-based Targeted SEL intervention was designed for female students (aged 12-15 years) who were identified by teachers as having elevated levels of relational aggression, with additional workshops provided for their parents.

Females receiving the intervention demonstrated statistically and clinically significant reductions in aggressive behaviours compared to the control.

To capture the experiences of children and young people, social validity measures were collected, with findings indicating that the female students generally perceived session topics as understandable, relevant, and important across most sessions. However, lower ratings were observed for sessions which involved more complex discussion formats and individual application tasks, with these perceived as less helpful and more difficult to engage with. Overall, participants reported good to very good perceptions of their experience in the group (Mean = 3.85; range 1-5 where 1 is the worst and 5 is the best), group content (M = 4.10), activities (M = 3.88), and group leadership (M = 4.2). Journal prompts (M = 3.0) and weekly goal sheets (M = 3.55) were rated less favourably, likely reflecting limited engagement with these components outside of lessons. While participants generally valued the group-based format and facilitation, the findings underscore the importance of ensuring that session activities are developmentally appropriate, clearly structured and adequately scaffolded to support understanding and engagement throughout the intervention to promote sustainable long-term effectiveness.

This study provides valuable but limited evidence on the views and experiences of children and young people participating in a Targeted SEL intervention. Specifically, the evidence base is restricted to a single pilot RCT involving only adolescent girls which limits the generalisability of findings. Further research is needed across a broader range of populations, interventions, and settings to strengthen understanding of how different groups of children and young people perceive and engage with Targeted SEL provision.

How much does it cost?

Two studies provided data on the cost of Targeted SEL interventions, of which one was from the US (Conduct Problems Prevention Research Group, 2010) and one was from the UK (Department for Education, 2016). Using the YEF-EQA, one paper was rated as very low quality (Department for Education, 2016). Costs associated with the Targeted SEL Fast Track preventative intervention were provided online and not reported directly in the primary research, meaning that the YEF-EQA could not be used to assess quality of cost-data (Conduct Problems Prevention Research Group, 2010).

The most recent study, conducted in the UK, evaluated the costs of the Better by Design (BBD) project, which aimed to build the social and emotional skills of young people in, or on the edge of care, who presented with challenging behaviours (Department for Education, 2016). The cost analysis demonstrated that over the 10-month evaluation period (April 2015 – February 2016), costs of providing the project, including paying for a residential hub, exceeded financial savings from avoided out-of-borough residential care by £369,000. Authors highlight that this was impacted by high and fixed costs associated with the residential setting, with lower than anticipated numbers of young people accessing the service (three vs. planned nine).

The Conduct Problems Prevention Research Group (2010) examined the costs associated with the Fast Track preventative intervention; a multi-site, multi-component intervention program targeting children and young people at the highest risk for life-course persistent conduct problems. Whilst the authors note that Fast Track was cost-effective due to reductions in young people needing to utilise mental health services, the actual cost-savings were not provided. The largest costs incurred were associated with initial and ongoing training ($4000 for two-day initial course, $2000 per booster day, plus travel costs), planning to discuss programme implementation with administrators ($1000 per day), and ongoing support from certified trainers ($100 per hour for phone consultation, $2000 per day for in-person support). Additional costs incurred included provisions of handouts, games and materials to young people and parents, and supplies needed for programme leaders (e.g., manuals).

Whilst both programmes outlined their set-up and ongoing costs, neither provided data on the actual cost-effectiveness of children and young people provided with Targeted SEL. For example, Targeted SEL interventions could lead to reduced costs associated with youth justice involvement, health care, social care, and education, leading to net savings in the long-term, but the current evidence base does not explore this.

Conclusion and Takeaway Messages

Our systematic review and meta-analysis demonstrate that Targeted SEL programmes have a substantial impact on reducing violence, crime and offending outcomes among children and young people, with much stronger effects than those observed in our review of universal SEL programmes (Keenan et al., 2026).

We also found evidence of improvements across a range of other outcomes of importance to YEF, such as behavioural difficulties, building and maintaining relationships, school engagement and helping others. Consistent with prior research, the effectiveness of Targeted SEL interventions may vary by programme design, population risk level, implementation quality, and study context, reflected in the considerable heterogeneity observed across studies.

Violence

Targeted SEL interventions were associated with a large and statistically significant reduction in violence, with an estimated 42.7% relative risk reduction across the included studies. Our meta-analysis of 15 violence-related outcomes (nine studies) suggests that the Targeted SEL programmes included in our analysis consistently and meaningfully reduced aggressive, violent, and disruptive behaviours across children and young people with identified needs or elevated risk. However, our confidence in this violence finding is low (evidence security Level 2) given the lower number and quality of randomised trials included.

Our finding on violence is consistent with the wider existing evidence base, Beelmann & Lösel (2021) reported that social skills training can reduce aggression by 28% (d=0.23), and oppositional and/or disruptive behaviours by 32% (d=0.27). Similarly a review by Piquero et al. (2016) found that self-control and social

competence interventions were associated with substantial improvements to self-control behaviours[7] (d=0.32).

Crime and Offending

Our meta-analysis also examined crime and offending outcomes across seven outcomes from three studies. Targeted SEL programmes result in an estimated 31.1% relative reduction in risk and an 8% absolute decrease in crime and offending behaviours amongst children and young people (based on baseline prevalence figures, this translates to 25% of youth becoming involved in crime who did not participate in Targeted SEL interventions to 17% with Targeted SEL). Caution is warranted with this finding as after adjusting for the clustering of outcomes within the three studies using robust variance estimation, the result became non-significant (p = 0.179), indicating that within-study correlation was underestimated in the original model. In simple terms, although the direction and magnitude of the effect are promising, the limited number of studies and very-low evidence security rating (Level 1) mean that strong conclusions cannot be drawn.

Nonetheless, these findings are aligned with previous systematic reviews examining targeted prevention and skills-based interventions for at-risk children and young people. For example, Lipsey & Cullen (2007) found that cognitive–behavioural and skills-based programmes were among the most effective approaches for reducing children and young people's engagement in crime and offending, reporting average reductions in reoffending of approximately 20–30% compared to control groups. Importantly, their review also highlights that programme effects tend to be larger when interventions are targeted at participants with existing behavioural difficulties or deficits in self-regulation, reinforcing the rationale for Targeted SEL approaches.

---

[7] 'Self-control' draws from Gottfredson & Hirschi's (1990) General Theory of Crime and is defined as the capacity to regulate impulses, delay gratification, manage emotions, and consider the consequences of one's actions. Low self-control is consistently associated with higher levels of aggressive and violent behaviour.

All outcomes

Beyond violence, crime and offending outcomes, our meta-analysis encompassed 135 outcomes across 10 YEF Outcome domains (e.g., behavioural difficulties, relationships, mental health, school engagement). We found a large overall effect of Targeted SEL programmes on this broader set of youth outcomes (g=−0.49, p < .001). In practical terms, children and young people who participated in Targeted SEL interventions showed significant improvements across a wide range of personal, social, and school-related outcomes compared to those who did not. These findings are drawn from 26 studies that assessed the impact of Targeted SEL interventions on children and young people and, as a result, we feel moderately confident in the findings and a Level 3 evidence security rating is applied.

Not every outcome category showed a statistically significant change via subgroup analysis, but several did, most notably behavioural difficulties, social and relationship skills, prosocial behaviours, and school engagement all improved for the participants receiving the Targeted SEL programme. Each of these statistically significant findings had a high impact according to the YEF criteria.

Our Universal SEL synthesis (Keenan et al., 2026) similarly found benefits across a broad set of outcome domains, yielding a moderate overall effect (g = -0.21, p < .001) based on 56 studies. The larger overall effect observed for Targeted SEL (g=−0.49, p < .001) is consistent with targeted programmes typically serving children and young people with higher baseline need or risk, leaving greater scope for measurable improvement than in universal samples.

What Works?

Fidelity emerged as a central implementation factor in the meta-review, with consistent delivery of Targeted SEL interventions supported by clear intervention protocols, structured training and ongoing monitoring. Where fidelity processes are weak or poorly documented, it becomes difficult to determine whether outcomes for children and young people are attributable to the intervention itself or to variability in delivery, underscoring the need to embed fidelity supports

within implementation planning from the outset. Importantly, ensuring focused delivery of core SEL content is critical in ensuring the best outcomes for children and young people, highlighting the need for high fidelity to the original intervention design.

Implementation capacity is further shaped by who delivers Targeted SEL and the level of preparation they receive. Most Targeted SEL interventions are delivered by classroom teachers, yet there is considerable inconsistency in training, preparation and ongoing support. The findings indicate that teachers require substantial professional development, including coaching and continued implementation support, to deliver Targeted SEL with confidence and consistency. Without this investment, delivery is more likely to vary across settings and over time, limiting fidelity and sustainability.

One-to-one and small-group formats support tailored, needs-led delivery, whilst requiring sufficient staffing and training to be implemented to a high-quality. For younger children, curricula that integrate relational strategies with structured, developmentally appropriate and play-based instruction are most feasible and engaging in practice. Together, these findings reinforce that the implementation of Targeted SEL interventions depends on deliberate planning, adequate resourcing and high fidelity to the original intervention design.

Who Benefits Most?

Gender

Gender was found to significantly moderate the impact of Targeted SEL interventions, with the largest and most consistent effect seen when boys make up most or all of the sample. Whilst Targeted SEL interventions had a small and significant effect on outcomes for girls, this was based on a single study only.

Importantly, past research has found gender differences in levels of social and emotional functioning amongst children and young people. Theorists suggest this is underpinned by differences in biological (e.g., puberty) and contextual developments (e.g., gender socialisation, behaviour norms; Ross et al., 2019). Boys, on average, display higher levels of externalising behaviours and weaker emotion regulation during key developmental periods (Maguire et al., 2016), which may

create more 'room for improvement', leading to larger observable gains from engaging in Targeted SEL interventions. Girls, by contrast, often show stronger baseline social-emotional competencies (Ross et al., 2019), but may benefit more when SEL is delivered in girls-only settings that align with social norms encouraging emotional expression, reflection, and relational engagement. Furthermore, in mixed-sex groups, peer impression management, and stereotype-based behaviour may affect children and young people's participation in interventions (Herrmann et al., 2024), particularly around vulnerability and emotion sharing, which feature in many Targeted SEL programmes. Together, these mechanisms help explain why Targeted SEL interventions show stronger effects when boys make up most or all of the sample.

Ethnicity

Our findings indicate that the effectiveness of Targeted SEL does not differ meaningfully by the ethnic composition of study samples. Only studies in which the sample's ethnicity was not stated or unclear showed a significant effect, possibly due to the large number of studies within this group. This demonstrates that under-reporting of participants' ethnicity remains a major issue and must be recorded in future research to fully understand who Targeted SEL interventions work best for. According to the YEF impact ratings, Targeted SEL interventions had a high impact on diverse/balanced samples and those that were majority white, whilst the impact on samples with some diversity was harmful. However, these findings were not significant at the $p < .05$ level.

Past research on SEL interventions has highlighted the importance of adapting SEL programmes to meet the cultural needs of children and young people (Li et al., 2024), with strategies including embedding metaphors, content or concepts that are widely used in the targeted populations' culture. Culturally adapted SEL programmes demonstrate more positive intervention effects on outcomes for children and young people (e.g., SEL skills, attitudes towards the self, school and others, positive social behaviour and conduct problems; Im, 2023). As such, interventions which have only some diversity amongst participants (between 15–49% of participants identified as Black and Global Majority) may be less likely to have embedded cultural adaptations sufficiently, compared to studies with a

diverse/balanced population, leading to the harmful impact rating. However, given the findings were not significant overall, they should be interpreted with caution.

Age

Age was found to significantly moderate the impact of Targeted SEL interventions. For younger children (0–9 years), effects were moderate yet non-significant, while for early adolescents (10–14 years), estimated impacts were high but still non-significant. For older adolescents (15–17 years), the effect size was very large; however, this estimate is based on a single study and should be interpreted with caution. In contrast, studies which target a broad range of ages showed a statistically significant and high impact according to the YEF framework. Differences in effects by age may reflect both developmental factors and features of the evidence base. Targeted SEL may have larger observable impacts during adolescence, when self-regulation and risk-taking are particularly salient, and when studies often involve higher-risk samples. However, limited study numbers in this analysis, especially for older adolescents, mean that age-related differences in this moderator analysis should be interpreted cautiously.

Limitations

Challenges with the studies measuring effectiveness

A key challenge in synthesising the evidence on Targeted SEL interventions relates to limitations in the reporting, availability, and consistency of quantitative outcome data across the relatively small body of eligible effectiveness studies.

We also had to exclude one otherwise relevant paper. Spence & Marzillier (1981) did not report the minimum statistical information required to calculate effect sizes, such as group-level means, measures of variance, or sufficient test statistics. The authors were contacted to request additional data, but further information was not made available.

The number of studies contributing to key outcomes, particularly violence and crime and offending, was small, limiting statistical power and increasing uncertainty (reducing confidence) around effect size estimates in those analyses. In addition, and relatedly, the strength of the moderator and meta-regression

analyses was limited by incomplete demographic reporting across many evaluations, which restricted our ability to draw firm conclusions about who Targeted SEL interventions are most effective for.

Challenges with the studies measuring implementation

Regarding the delivery method and specific curricula components, no statistical analyses were conducted, limiting our understanding of the impact of these factors on the effectiveness of Targeted SEL. For cost-data, only two studies were available and neither provided data on the actual cost-effectiveness of children and young people provided with Targeted SEL, instead focusing on outlining set-up and ongoing costs. Given the positive outcomes associated with participating in Targeted SEL interventions, theoretically this could lead to reduced costs across the youth justice, health care, social care and education systems, but research examining this is lacking.

Within the meta-review, two studies were rated as Low quality and two as Critically Low quality using the AMSTAR 2 tool. Low and critically low-quality reviews reduce our confidence in the findings, with examples of methodological issues including lacking a pre-specified protocol, incomplete reporting of excluded studies, inadequate or missing risk-of-bias assessments. This demonstrates the need for higher-quality evidence on the implementation of Targeted SEL interventions.

Overlap of primary studies is a recognised and largely unavoidable methodological consideration in meta-reviews, particularly when reviews address related research questions and draw on overlapping evidence bases. In line with recommendations by Pieper et al. (2014), overlap was systematically assessed using a citation matrix and quantified using the Corrected Covered Area (CCA). In this meta-review of Targeted SEL interventions, overlap across the four included reviews was low (CCA = 3%), which corresponds to a slight overlap according to established thresholds (Pieper et al., 2014)[8]. This finding indicates minimal redundancy of primary studies and suggests that the synthesis is not disproportionately influenced by repeated inclusion of the same evidence.

---

[8] Categorisations for overlap in reviews are Slight = 0-5%, Moderate = 6-10%, High = 11-15%, Very High = >15%

Consequently, the low level of overlap supports the robustness of our findings and strengthens confidence that the conclusions are informed by a largely distinct body of evidence.

Final Thoughts and Recommendations

Targeted SEL interventions are associated with a large reduction in violence amongst children and young people, with an estimated 43% reduction among recipients of these interventions. Targeted SEL programmes also show promising reductions in crime and offending, with an estimated 31% decrease. However, due to the number and quality of studies included in our analysis, we have low confidence in findings related to violence and very low confidence in findings related to crime and offending. In addition, the findings for crime and offending become statistically uncertain after adjusting for clustering. Taken together, these findings should be interpreted with caution.

Based on the current available evidence, we make the following key recommendations:

- Where reducing aggressive, violent, or disruptive behaviour is a key objective, use Targeted SEL programmes with children and young people who have identified needs or elevated risk, as these approaches can achieve substantially greater impacts than Universal SEL programmes (Keenan et al., 2026).
- As with Universal SEL programmes (Keenan et al., 2026), evidence on crime and offending outcomes is more limited and less certain. We recommend more high-quality trials that include validated measures of crime and offending (including administrative outcomes where feasible), alongside longer follow-up periods, to continue to build the evidence base.
- There are persistent limitations in the reporting of demographic variables amongst primary studies, including gender, age and ethnicity. Evaluations need to consistently record and report on these to enable us to develop our understanding of who benefits most.
- Targeted SEL should appropriately match the needs of children and young people, with one-to-one and small group delivery supporting the implementation of this.

- Embed fidelity supports from the outset by using clear intervention protocols, structured monitoring systems, and documentation processes to ensure Targeted SEL is delivered as intended.
- Prioritise focused delivery of core SEL content, maintaining high fidelity to the original intervention design rather than adapting or extending programmes in ways that dilute essential components.
- For teachers facilitating Targeted SEL interventions, we recommend embedding ongoing opportunities for training, including structured professional development, coaching, opportunities to facilitate alongside intervention specialists and manuals to guide delivery.
- Use developmentally appropriate curricula, particularly for younger children, by integrating relational strategies with structured, play-based and experiential learning formats.
- Future research is needed to explore the cost-effectiveness of Targeted SEL interventions and to develop our understanding of the most appropriate intervention duration/dosage that should be provided.

References

Studies included in effectiveness and implementation sections, are marked with an asterisk (*) at the beginning of the reference.

Beelmann, A., & Lösel, F. (2021). A Comprehensive Meta-Analysis of Randomized Evaluations of the Effect of Child Social Skills Training on Antisocial Development. *Journal of Developmental and Life-Course Criminology*, 7(1), 41–65. https://doi.org/10.1007/s40865-020-00142-8

* Blewitt, C., O'Connor, A., Morris, H., May, T., Mousa, A., Bergmeier, H., Nolan, A., Jackson, K., Barrett, H., & Skouteris, H. (2019). A systematic review of targeted social and emotional learning interventions in early childhood education and care settings. *Early Child Development and Care*, 191(14), 2159–2187. https://doi.org/10.1080/03004430.2019.1702037

* Block, J. (1978). Effects of a rational–emotive mental health program on poorly achieving, disruptive high school students. *Journal of Counseling Psychology,* 25(1), 61–65. https://doi.org/10.1037/0022-0167.25.1.61

* Burke, J. D., & Loeber, R. (2015). The effectiveness of the Stop Now and Plan (SNAP) program for boys at risk for violence and delinquency. *Prevention Science*, 16(2), 242–253. https://doi.org/10.1007/s11121-014-0490-2

* Camp, B., Blom, G., Hebert, F., & van Doorninck, W. (1977). 'Think aloud': A program for developing self-control in young aggressive boys. *J Abnorm Child Psychol,* 5(2), 157–169. https://doi.org/10.1007/BF00913091

CASEL. (n.d.). Fundamentals of SEL. Collaborative for Academic, Social, and Emotional Learning. Retrieved 10 December 2025, from https://casel.org/fundamentals-of-sel/

* Çelik, S., Diken, I., Çolak, A., Arıkan, A., Aksoy, F., & Tomris, G. (2016). Effectiveness of the Preschool Version of the First Step to Success Early Intervention Program for Preventing Antisocial Behaviors. *Educational Sciences: Theory & Practice*, 16(2), 511–535. https://doi.org/10.12738/estp.2016.2.0317

* Collier, C. E. (2002). *The effect of pro-social skills training on the problem behavior of selected African American children in the District of Columbia public schools: Recommendations for change.* Dissertation Abstracts International: Section B: The Sciences and Engineering.

* Conduct Problems Prevention Research Group. (2010a). Fast Track intervention effects on youth arrests and delinquency. *Journal of Experimental Criminology*, 6(2), 131–157. https://doi.org/10.1007/s11292-010-9091-7

Conduct Problems Prevention Research Group. (2010b). Fast track intervention effects on youth arrests and delinquency: Erratum. *Journal of Experimental Criminology*, 9, 131–157. http://dx.doi.org/10.1007/s11292-010-9114-4

* Department for Education. (2016). *Evaluation of Better by Design: A report for Achieving for Children.* Department for Education. https://assets.publishing.service.gov.uk/media/5a81b467ed915d74e33ffa4e/Achieving_for_Children_-_Better_by_Design.pdf

* Dong, Y., Hallberg, E. T., & Hassard, H. (1979). Effects of assertion training on aggressive behavior of adolescents. *Journal of Counseling Psychology*, 26(5), 459–461. https://doi.org/10.1037/0022-0167.26.5.459

* Dupper, D. R., & Krishef, C. H. (1993). School-based social-cognitive skills training for middle school students with school behavior problems. *Children and Youth Services Review*, 15(2), 131–142. https://doi.org/10.1016/0190-7409(93)90040-g

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (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

* Edleson, Dr. J. L., & Rose, S. D. (1982). Investigations Into the Efficacy of Short-Term Group Social Skills Training for Socially Isolated Children. *Child Behavior Therapy,* 3(2–3), 1–16. https://doi.org/10.1300/J473v03n02_01

* Feindler, E. L., Marriott, S., & Iwata, M. (1984). Group anger control training for junior high school delinquents. *Cognitive Therapy and Research*, 8(3), 299–311. https://doi.org/10.1007/BF01173000

Gaffney, H., Farrington, D. P., & White, H. (2021). *Social Skills Training: Toolkit Technical Report.* Youth Endowment Fund. https://youthendowmentfund.org.uk/wp-content/uploads/2021/06/Social-Skills-technical-report.pdf

* Goldberg, J. M., Sklad, M., Elfrink, T. R., Schreurs, K. M. G., 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

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford University Press.

Herrmann, L., Reiss, F., Becker-Hebly, I., Baldus, C., Gilbert, M., Stadler, G., Kaman, A., Graumann, L., & Ravens-Sieberer, U. (2024). Systematic Review of Gender-Specific Child and Adolescent Mental Health Care. *Child Psychiatry & Human Development*, 55(6), 1487–1501. https://doi.org/10.1007/s10578-023-01506-z

* Hudley, C., Birtch, B., Wakefield, W., Smith, T., Demorat, M., & Cho, S.-J. (1998). An attribution retraining program to reduce aggression in elementary school students. *Psychology in the Schools*, 35, 271–282. https://doi.org/10.1002/(SICI)15206807(199807)35:3%253C271::AID-PITS7%253E3.0.CO;2-Q

* Hudley, C., & Graham, S. (1993). An Attributional Intervention to Reduce Peer-Directed Aggression among African-American Boys. *Child Development*, 64(1), 124–138. https://doi.org/10.2307/1131441

Im, S. (2023). *A Meta-Analysis of Culturally Adapted Social Emotional Learning Programs for Racially and Ethnically Minoritized Youth*. The University of Wisconsin-Madison. https://search.proquest.com/openview/97589cd4a9789860cb48d4c81406afcd/1?pq-origsite=gscholar&cbl=18750&diss=y

* Kazdin, A., Esveldt-Dawson, K., French, N., & Unis, A. (1987). Problem-solving skills training and relationship therapy in the treatment of antisocial child behavior. *J Consult Clin Psychol*, 55(1), 76–85. https://doi.org/10.1037//0022-006x.55.1.76

Keenan, C., Mallion, J. S., Hedges, S., Nugent, R., Idrees, I., Murphy, L., Allen, K., Torres, O., Hinks, J., Towers, J. (2026). *Social and Emotional Learning - Universal*. Youth Endowment Fund.

Li, A., Miller, F., & Williams, S. (2024). Cultural adaptations to social–emotional learning programs: A systematic review. *School Psychology*. https://psycnet.apa.org/record/2025-06564-001

Lipsey, M. W., & Cullen, F. T. (2007). The Effectiveness of Correctional Rehabilitation: A Review of Systematic Reviews. *Annual Review of Law and Social Science*, 3(1), 297–320. https://doi.org/10.1146/annurev.lawsocsci.3.081806.112833

* Lochmann, J. E., Coie D, J., Underwood, M. K., & Terry, R. (1993). Effectiveness of a social relations intervention program for aggressive and nonaggressive, rejected children. *Journal of Consulting and Clinical Psychology*, 61, 1053–1058. https://doi.org/10.1037/0022006X.61.6.1053

Maguire, L. K., Niens, U., McCann, M., & Connolly, P. (2016). Emotional development among early school-age children: Gender differences in the role of problem behaviours. *Educational Psychology*, 36(8), 1408–1428. https://doi.org/10.1080/01443410.2015.1034090

* Mannarino, A. P., Christy, M., Durlak, J. A., & Magnussen, M. G. (1982). Evaluation of social competence training in the schools. *Journal of School Psychology*, 20(1), 11–19. https://doi.org/10.1016/0022-4405(82)90036-x

* Murano, D., Sawyer, J. E., & Lipnevich, A. A. (2020). A Meta-Analytic Review of Preschool Social and Emotional Learning Interventions. *Review of Educational Research*, 90(2), 227–263. https://doi.org/10.3102/0034654320914743

Parker, K., Nunns, M., Xiao, Z., Ford, T., Stallard, P., Kuyken, W., Axford, N., & Ukoumunne, O. C. (2025). Patterns of intra-cluster correlation coefficients in school-based cluster randomised controlled trials of interventions for improving social-emotional functioning outcomes in pupils: A secondary data analysis of five UK-based studies. *BMC Medical Research Methodology*, 25(1), 120. https://doi.org/10.1186/s12874-025-02574-6

* Pedro-Carroll, J. (1983). The Children of Divorce Intervention Project: An Investigation of the Efficacy of a School-Based Prevention Program. *Unpublished Doctoral Dissertation*, Ann Arbor, MI: University Microfilms.

Pieper, D., Antoine, S.-L., Mathes, T., Neugebauer, E. A., & Eikermann, M. (2014). Systematic review finds overlapping reviews were not mentioned in every other overview. *Journal of Clinical Epidemiology*, 67(4), 368–375.

Piquero, A. R., Jennings, W. G., Farrington, D. P., Diamond, B., & Gonzalez, J. M. R. (2016a). A meta-analysis update on the effectiveness of early self-control improvement programs to improve self-control and reduce delinquency. *Journal of Experimental Criminology*, 12(2), 249–264. https://doi.org/10.1007/s11292-016-9257-z

Piquero, A. R., Jennings, W. G., Farrington, D. P., Diamond, B., & Gonzalez, J. M. R. (2016b). A meta-analysis update on the effectiveness of early self-control improvement programs to improve self-control and reduce delinquency. *Journal of Experimental Criminology*, 12(2), 249–264. https://doi.org/10.1007/s11292-016-9257-z

* Prinz, R., Blechman, E., & Dumas, J. (1994). An evaluation of peer coping-skills training for childhood aggression. *Journal of Clinical Child & Adolescent Psychology*, 23(2), 193–203. https://doi.org/10.1207/s15374424jccp2302_8

Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science: The Official Journal of the Society for Prevention Research*, 23(3), 425–438. https://doi.org/Meta-analysis%2520with%2520Robust%2520Variance%2520Estimation:%2520Expanding%2520the%2520Range%2520of%2520Working%2520Models

* Quinn, M. M. (2002). Changing antisocial behavior patterns in young boys: A structured cooperative learning approach. *Education & Treatment of Children,* 25, 380–395.

Ross, K. M., Kim, H., Tolan, P. H., & Jennings, P. A. (2019). An exploration of normative social and emotional skill growth trajectories during adolescence. *Journal of Applied Developmental Psychology*, 62, 102–115.

Rutterford, C., Copas, A., & Eldridge, S. (2015). Methods for sample size determination in cluster randomized trials. *International Journal of Epidemiology*, 44(3), 1051–1067. https://doi.org/10.1093/ije/dyv113

* Schneider, H. B., & Byrne, M. B. (1987). Individualizing social skills training for behavior-disordered children. *J Consult Clin Psychol*, 55(3), 444–445. https://doi.org/10.1037//0022-006x.55.3.444

* Shelton, T. L., Barkley, R., Crosswait, C., Moorehouse, M., Fletcher, K., Barrett, S., Jenkins, L., & Metevia, L. (2000). Multimethod psychoeducational intervention for preschool children with disruptive behavior: Two-year post-treatment follow-up. *J Abnorm Child Psychol*, 28(3), 253–266. https://doi.org/10.1023/a:1005144304071

* Smokowski, P. R., Fraser, M. W., Day, S., Galinsky, M., & Bacallao, M. L. (2004). School-Based Skills Training to Prevent Aggressive Behavior and Peer Rejection in Childhood: Evaluating the Making Choices Program. *The Journal of Primary Prevention*, 25(2), 233–251. https://doi.org/10.1023/b:Jopp.0000042392.57611.05

Spence, S. H., & Marzillier, J. S. (1981). Social skills training with adolescent male offenders: II. Short-term, long-term and generalized effects. *Behaviour Research and Therapy*, 19(4), 349–368. https://doi.org/10.1016/0005-7967(81)90056-5

* Splett, J. D., Maras, M. A., & Brooks, C. M. (2015). GIRLSS: A Randomized, Pilot Study of a Multisystemic, School-Based Intervention to Reduce Relational Aggression. *Journal of Child and Family Studies,* 24, 2250–2261. https://doi.org/10.1007/s10826-014-0027-0

* Stouwe, T., Asscher, J., Hoeve, M., Laan, P., & Stams, G. (2016). Social skills training for juvenile delinquents: Post-treatment changes. *Journal of Experimental Criminology*, 12, 515.

* Tanner, V., & Holliman, W. B. (1988). Effectiveness of assertiveness training in modifying aggressive behaviors of young children. *Psychol Rep*, 62(1), 39–46. https://doi.org/10.2466/pr0.1988.62.1.39

* Webster-Stratton, C., Reid, J., & Hammond, M. (2001). Social skills and problem-solving training for children with early-onset conduct problems: Who benefits? *J Child Psychol Psychiatry*, 42(7), 943–952. https://doi.org/10.1111/1469-7610.00790

* Yang, W., Datu, J. A. D., Lin, X., Lau, M. M., & Li, H. (2019). Can Early Childhood Curriculum Enhance Social-Emotional Competence in Low-Income Children? A Meta-Analysis of the Educational Effects. *Early Education and Development,* 30(1), 36–59. https://doi.org/10.1080/10409289.2018.1539557

YEF. (2023). *Children, violence and vulnerability: The second annual Youth Endowment Fund report into young people's experiences of violence.* Youth Endowment Fund. https://youthendowmentfund.org.uk/wp-content/uploads/2023/11/YEF_Children_violence_and_vulnerability_2023_FINAL.pdf

Appendix 1. Methods of the systematic review

Protocol

Prior to initiating this systematic review, we developed a comprehensive protocol for an Evidence and Gap Map (EGM) outlining the research objectives, eligibility criteria, search strategy, data extraction, quality appraisal, and synthesis methods. This protocol was registered and is available on the Open Science Framework (OSF),[9] ensuring transparency and adherence to predefined methods.

The methods described below are aligned with the current Toolkit Strands on Problem-Oriented Policing, Precourt diversion, and Youth Clubs, ensuring a structured and rigorous approach to evidence synthesis.

Eligibility Criteria

To define the scope of relevant research, we applied the following criteria from the outset:

	Inclusion Criteria	Exclusion Criteria
Population	Children and young people aged 0–17 **with an identified need** (e.g., conduct problems, behavioural issues, aggression, or prior involvement in antisocial behaviour, violence or crime). Also aim to capture subgroup effects by race, gender, age, SES, or disability.	Adults aged 18 and older unless the target age group is also included and the outcomes can be disaggregated.

---

[9] Protocol is available to access here: https://osf.io/vamxy

Dose	Any, including one-off sessions to multi-year programmes. Programmes should devote ≥30% of contact time to SEL skill-building	Any programmes where the SEL component comprises <30% of contact time.
Setting	Schools, community settings, youth justice settings, or youth custody.	Purely therapeutic or clinical settings (e.g. individual psychotherapy or CBT delivered as treatment without a documented SEL curriculum).
Primary mechanism of change	Structured SEL skills development is **central and the most dominant or 'active' element of the intervention**, with a documented curriculum required (manuals, lesson plans, modules). The intervention **MUST** target/seek to reduce crime or offending; violence or violent crime; and externalising behaviours associated with violence (e.g., aggression, bullying, alcohol abuse, drug abuse). This criterion is used to determine study inclusion. Once the study is included, all *additional* outcomes that fall within the YEF Outcomes Framework will be extracted,	Programmes that do not include a structured SEL skills development curriculum, programme, manual or set of resources. Interventions that DO NOT target/seek to reduce crime or offending; violence or violent crime; and externalising behaviours associated with violence (e.g., aggression, bullying, alcohol abuse, drug abuse). Studies that only focus on outcomes unrelated to crime, violence or externalising behaviours (e.g., purely academic

	even if they are not crime or violence related.	outcomes, general wellbeing).
Participation	Voluntary or mandatory. May include children identified via referral (schools, children's services, youth justice), or in connection with out-of-court resolutions or youth custody.	N/A
Delivered by	Any providers, typically youth sector providers or teachers.	N/A
Timing	Any	N/A
Study design	Randomised controlled trials (RCTs) and robust quasi-experimental designs (QEDs) e.g. those using difference-in-difference, synthetic controls, regression discontinuity, propensity score matching, or instrumental variables.	Studies without a credible counterfactual e.g. pre/post designs, QEDs with non-robust matching techniques

Details of screening

Searches for both Targeted and Universal SEL were run simultaneously in OpenAlex and GoogleScholar and returned 30,246 records. Given the volume of research retrieved, we used a machine-learning–assisted screening approach to efficiently prioritise records for title and abstract review. OpenAlex bibliographic metadata were downloaded and uploaded into ASReview LAB v2, an open-source active-learning screening tool designed to support systematic title and abstract screening of large textual datasets. Within ASReview LAB, records were screened at title and abstract level using an active-learning workflow with a senior review

member (CK) making initial decisions. A total of 2,117 titles and abstracts were identified as potentially relevant following screening in ASReview LAB. These records were then exported and imported into EPPI-Reviewer for onward processing. The EPPI-Reviewer classifier conducted initial title and abstract screening. A total of 631 records were marked as included and 1,486 as excluded. A team of reviewers then screened the full text for all included studies. Senior reviewers conducted a check of any exclusions, and any discrepancies were discussed as a team and reconciled.

Of the 631 records screened at full-text level, 301 were excluded (for reasons, see Table 166). For inaccessible PDFs, the team attempted to contact lead authors to request access to the report or further data.

Table 16: *Full-text screening results*

Reason for exclusion	Number of Records Excluded at Full-Text Level
Did not target children and young people	10
PDF not accessible	26
Study Design not eligible	90
Outcomes or intervention not relevant	152
Excluded based on language	21
Duplicates	2

Following full-text screening, 330 studies were flagged as potentially relevant for inclusion. In addition, 116 studies were identified from the existing EGM; these had previously been coded under an earlier "social skills" category. Senior team members reviewed these at full text level to assess eligibility for the current SEL strand. All 446 studies had data extracted for the EGM, with reviewers checking the study design and relevant Toolkit strand. Of these 446 papers, six papers were excluded (two on study design and four on relevance to YEF), 186 were categorised into other Toolkit strands. Following this process, 71 were included in the Targeted SEL strand and 183 were included in the Universal SEL strand.

Following data extraction for the EGM, these 71 papers proceeded to full Toolkit data extraction. At this stage reviewers checked the study design in detail and excluded 10 studies where the comparison group was not 'business as usual' and 26 studies that did not meet the strand's robust causal study design criteria (i.e., the minimum requirements for credible, causal and robust evidence). A further study was excluded because the sample included 12 to 21-year-olds and outcomes could not be disaggregated for 12 to 17-year-olds. All exclusions following full text screening were checked by a senior reviewer.

Table 17: *Reasons for exclusion after full text screening*

Reason for exclusion	Number of Records Excluded at EGM Data Extraction Level
Study design not meeting robust causal threshold	26
Comparison group was not business as usual	10
Outcomes or intervention not relevant	1

There was a total of 34 papers that were eligible for inclusion. Of these, four papers related to the same study (the 'Fast Track' prevention study) and so three papers were linked to the paper with the most relevant information (the 'main study').

Table 18: *List of linked studies*

Master study	Linked studies
Conduct Problems Prevention Research Group (2010)	Bierman (2007), Dodge (2013), Kalvin (2019)

Among the 31 studies, five had only implementation evidence, two had both implementation and effectiveness evidence, and 24 had effectiveness evidence only, giving a final total of 26 effectiveness studies and seven implementation studies. One effectiveness study did not contain the statistical information needed to include in the meta-analysis (Spence & Marzillier, 1981). The author was contacted and responsive but was unable to provide further information beyond

what was published (due to the passage of time and technical limitations), so this study was excluded from the meta-analysis; it is included in the description of interventions section above but not in Appendix 3. The final total for the meta-analysis was therefore 26 studies with one study (Schneider & Byrne, 1987) providing two independent clusters based on a multi-arm trial.
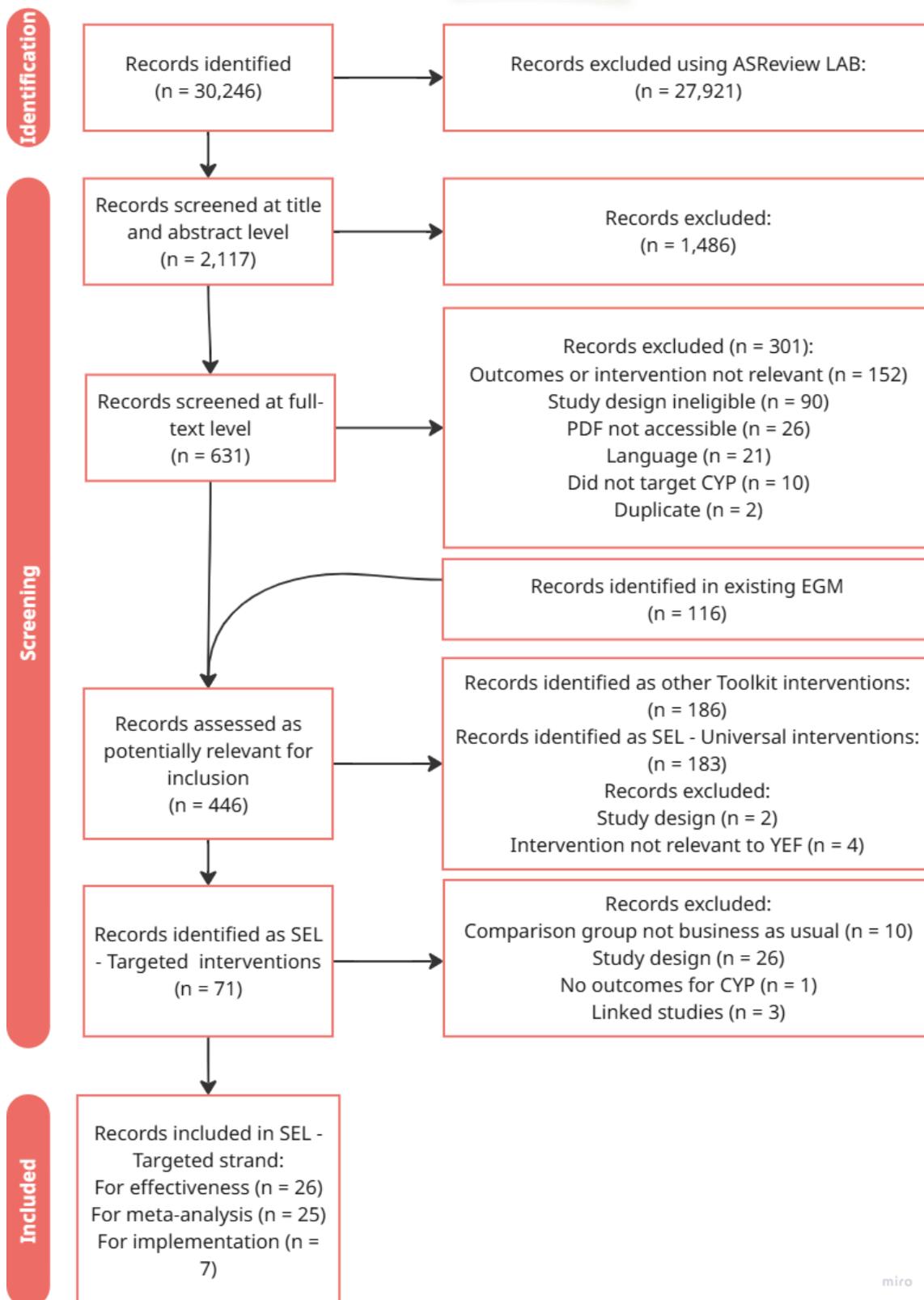
Figure 4: *PRISMA Flow Diagram*

Quality appraisal process for effectiveness studies

The YEF-EQA tool was used across all studies to systematically assess the quality, reliability, and relevance of the research.

Table 19: *Quality appraisal ratings for studies included in Targeted SEL*

Study ID	Overall quality of the study	Study Design
Block (1978)	Moderate	RCT
Burke and Loeber (2014)	Moderate	RCT
Camp et al. (1977)	High	RCT
Çelik et al. (2016)	Low	RCT
Collier (2002)	Moderate	RCT
Conduct Problems Prevention Research Group (2010)	High	RCT
Dong et al. (1979)	Low	RCT
Dupper and Krishef (1993)	Moderate	RCT
Edleson and Rose (1982)	Moderate	RCT
Feindler et al. (1984)	Low	RCT
Hudley et al. (1998)	Moderate	RCT
Hudley and Graham (1993)	Moderate	RCT
Kazdin et al. (1987)	Moderate	RCT
Lochmann et al. (1993)	Moderate	RCT
Mannarino et al. (1982)	Moderate	RCT

Pedro-Carroll (1983)	Low	RCT
Prinz et al. (1994)	Low	RCT
Quinn (2002)	Moderate	RCT
Schneider and Byrne (1987)	Moderate	RCT
Shelton et al. (2000)	Moderate	RCT
Smokowski et al. (2004)	Moderate	RCT
Spence and Marzillier (1981)	Low	RCT
Splett et al. (2015)	Moderate	RCT
Stouwe et al. (2016)	Moderate	QED
Tanner and Holliman (1988)	Low	RCT
Webster-Stratton et al. (2001)	Low	RCT

How the findings were analysed and combined

Preparing the data frame for analysis

Included studies comprised a mixture of individually randomised controlled trials, cluster randomised controlled trials, and robust quasi-experimental designs. For all studies, the primary effect size was the standardised mean difference (SMD) expressed as Hedges' g.

As is common in meta-analysis, the team encountered several challenges in harmonising effect sizes across studies. First, there was a need to standardise the direction of SMD values so that negative values consistently indicate a favourable intervention effect. For example, in Smokowski et al. (2004), a decrease in 'overt aggression' is positive, but a decrease in 'social competence' is negative. To achieve alignment across all outcomes, we reviewed all effect size direction labels and, where necessary, multiplied positive SMD values by −1 for outcomes where a lower score signified improvement This ensured that all SMDs reflected the same directional meaning, **that negative SMD values always indicate beneficial effects.** This approach supports clear interpretation and comparability across studies within the meta-analysis.

Four outcomes were odds ratios and so were approximately converted into an equivalent SMD by multiplying the natural log of the $OR \times \frac{\sqrt{3}}{\pi}$ (approximately 0.5513).[10] In practical terms:

$$SMD \ (Cohen's \ d) \ \approx \frac{\sqrt{3}}{\pi} \ln(OR) \ = \ \frac{\ln(OR)}{1.814}$$

Additionally, nine outcomes across three studies were from cluster randomised controlled trials. In a cluster trial, randomisation occurs at a group-level (e.g. school, classroom). People within the same cluster tend to be more similar to each other than to people in other clusters. There is therefore less information about the effectiveness of an intervention in a trial of 100 people in 10 schools than in a trial of 100 people randomised individually. This non-independence of

---

[10] Because this transformation is a linear scaling, the standard error of the SMD can be obtained by applying the same $\frac{\sqrt{3}}{\pi}$ factor to the standard error of the log OR

participants within clusters must be accounted for or the effect size will be over-estimated. For cluster randomised trials which reported the effect sizes from cluster-adjusted analyses (e.g. mixed-effects models, generalised estimating equations) along with their standard errors or confidence intervals, we converted those estimates directly to Hedges' g and derived the standard error of g from the reported uncertainty, with no further adjustment required. However, for cluster randomised trials analysed as if they were individually randomised, the effect size precision was adjusted using a design effect (Rutterford et al., 2015). The design effect (DE) was defined as:

$$DE = 1 + (M - 1)\rho$$

Where M is the average cluster size and ρ is the intracluster correlation coefficient (ICC). For these studies we retained the point estimate of g but inflated its standard error by DE (equivalently, we reduced the effective sample size in each arm by dividing the original sample size by DE before computing the sampling variance). When ICCs were not reported, we used values from similar studies in the same substantive area, matching outcome type and reporter (self-report, parent, or teacher) as closely as possible (Parker et al., 2025). The impact of these assumptions was examined in sensitivity analyses by varying the ICC over a plausible range.

Meta-analysis

A random-effects model was fitted to the data. The amount of heterogeneity (i.e., τ^2), was estimated using the restricted maximum-likelihood estimator (Viechtbauer, 2005). In addition to the estimate of τ^2, the Q-test for heterogeneity (Cochran, 1954) and the I^2 statistic (Higgins & Thompson, 2002) are reported. In case any amount of heterogeneity is detected, a prediction interval for the true outcomes is also provided (Riley et al., 2011).

Publication bias

To assess the presence of publication bias in our meta-analysis, we employed several statistical tests designed to detect funnel plot asymmetry, which can be indicative of such bias. A funnel plot was generated to visually inspect for publication bias. In the absence of bias, the plot should resemble a symmetrical

inverted funnel, with effect sizes from smaller studies scattering widely at the bottom and larger studies clustering near the top.

Sensitivity analyses

Studentized residuals and Cook's distances are used to examine whether studies may be outliers and/or influential in the context of the model (Viechtbauer & Cheung, 2010). Studies with a studentized residual larger than the $100 \times (1-0.05/(2 \times k))$th percentile of a standard normal distribution are considered potential outliers (i.e., using a Bonferroni correction with two-sided $\alpha = 0.05$ $\alpha = 0.05$ for k studies included in the meta-analysis). Studies with a Cook's distance larger than the median plus six times the interquartile range of the Cook's distances are considered to be influential. The rank correlation test (Begg & Mazumdar, 1994) and the regression test (Sterne & Egger, 2005), using the standard error of the observed outcomes as predictor, are used to check for funnel plot asymmetry. The analysis was carried out using R (version 4.4.2) (R Core Team, 2020) and the metafor package (version 4.8.0) (Viechtbauer, 2010).

Sub-group analysis

Subgroup-specific effects were estimated using robust variance estimation with CR2 adjustment, and statistical significance was assessed using t-tests with Satterthwaite small-sample degrees of freedom. We used coef_test() from the clubSandwich package to compute cluster-robust standard errors (CR2) with Satterthwaite degrees of freedom for hypothesis testing. Differences in effects across outcome domains were examined using a CR2-robust Wald chi-square test.

Results tables presented in the main report show the robust effect size (SMD) of Targeted SEL interventions on different subgroup domains, with robust standard errors (SE) and study counts (k = number of effect sizes; n = number of studies). Negative SMDs indicate a reduction/improvement for the Targeted SEL group compared to BAU control.

All estimates below are calculated using robust variance estimation (RVE) to account for dependence among effect sizes within studies. p-values correspond to two-tailed tests.

Meta-review

A meta-review was conducted to explore implementation factors related to Targeted SEL. A meta-review refers to an evidence synthesis approach which summarises findings from multiple existing systematic reviews and meta-analyses. Rather than analysing individual studies, a meta-review examines the consistency of the evidence across reviews, enabling a high-level summary of what is known. We undertook a meta-review for this strand due to the high volume of primary studies exploring the implementation of Targeted SEL interventions, which was not feasible to examine in the timescale and budget available.

The EGM had six systematic reviews, all of which were reviewed initially by a senior researcher. Of these, four were identified as having examined implementation factors and were included in the meta-review. The quality of these were assessed using the Chat-GPT AMSTAR-2 expert, which analyses systematic reviews using the AMSTAR 2 tool.

AMSTAR 2 (A Measurement Tool to Assess Systematic Reviews) is a critical appraisal tool used to evaluate the methodological quality of systematic reviews. It consists of 16 domains that assess how rigorously a systematic review was planned, conducted and reported, covering areas such as protocol registration, literature search quality, risk-of-bias assessment, data synthesis methods, and consideration of publication bias. Seven of these domains are considered critical, meaning weaknesses in these areas substantially reduce confidence in the review's findings. AMSTAR 2 rates confidence in the results of a review as high, moderate, low or critically low, based on the presence of critical and non-critical flaws.

Table 20: *Quality appraisal ratings for studies included in the meta-review*

Study ID	Overall quality of the study
Blewitt et al. (2019)	Low
Goldberg et al. (2019)	Critically Low
Murano et al. (2020)	Low
Yang et al. (2019)	Critically Low

A thematic analytic approach was taken to identify and structure recurring patterns in the reviews regarding implementation factors. An inductive approach was taken and began with a familiarisation with the reviews, involving reading these several times, noting initial observations and reflections on implementation factors included. Looking across reviews, themes were created on common implementation factors examined. Finally, themes were refined, named and defined and a narrative summary related to each theme was written for the report. If the systematic reviews report moderator analyses that contribute to our understanding of implementation, we describe this narratively.

Based on the data available, findings are reported according to the following themes:

Implementation Processes and Delivery Capacity

- **Fidelity**

    Fidelity refers to how closely Targeted SEL interventions are delivered as originally designed, covering the quality and consistency of implementation.

- **Intervention Providers and Training**

    Intervention provider and training relate to who delivers the Targeted SEL intervention (e.g., classroom teachers, school staff, external specialists) and the nature of their training and ongoing support.

- **Family Involvement**

  Family involvement refers to the extent to which parents or carers are engaged in Targeted SEL interventions, including participation in workshops, home visits, and take-home activities, or acting as primary implementers through home-based parent training programmes.

Delivery Characteristics

- **Delivery Method**

  Delivery method refers to the way in which Targeted SEL content is delivered and experienced in practice, including the structure, format, and techniques used to engage learners. It shapes how participants interact with the material, how consistently skills are taught, and how feasible the intervention is to implement.

- **Duration / Dosage**

  Duration and dosage describe the length, frequency, and intensity of Targeted SEL interventions, including how long programmes run and the cumulative exposure children and young people receive.

- **Intervention Curriculum**

  Intervention curriculum refers to the specific content, concepts and skills covered in a Targeted SEL programme, along with how these elements are structured, sequenced and articulated for real-world delivery. It shapes what participants are exposed to, how complex delivery demands are, and how well core mechanisms of change are preserved.

The systematic reviews did not cover cost or children and young people's views. As such, we supplemented these sections with primary studies that were identified at the EGM or Toolkit extraction level as having cost-data or reported children and young people's views. Of these, two primary studies reported on cost-data and one on children and young people's views, and all three were also included in the meta-analysis.

Appendix 2. Location Details

	Number of UK Studies	Number (and Location) of International Studies
Contributing to Evidence Quality Rating	0	USA (n=21); Canada: (n = 3); Netherlands (n = 1); Turkey (n = 1)
Contributing to Estimated Impact on Violence	0	USA (n=9)
Contributing to Estimated Impact on Crime and Offending	0	USA (n=3)
Contributing to Cost Data[11]	1	USA (n=1)

---

[11] The meta-review included in this report examined existing systematic reviews and meta-analyses to examine implementation factors influencing the effectiveness of universal social and emotional learning interventions for children and young people. As a result, detailed study-level contextual information including the number of countries represented in individual primary studies contributing to the implementation evidence is provided in Appendix 5 for each study included in this meta-review.

Appendix 3. Characteristics of included studies for effectiveness

Authors (Year)	Country	Study Design	Intervention	Population/ Place	Comparison	Outcomes Measured	Quality Level	Findings
Block (1978)	US	RCT	A rational-emotive mental health program aimed at improving youth performance in schools.	High school students identified as being high-risk and failure-and misconduct prone.	Comparison groups included students who received human relation training and a group who received treatment as usual.	Behavioural difficulties School engagement	Moderate	Students who had received the rational-emotive mental health program had significantly higher GPA's, significantly lower incidents of disruptive behaviour and significantly lower levels of class cutting at both post-test, 1 week before the intervention ended and at the 4-months follow up

							compared to those students who received the alternative treatment and non-treatment control condition.
Burke and Loeber (2014)	US	RCT & PE	The Stop Now and Plan (SNAP) program had multiple components such as SNAP family counselling, booster sessions, academic tutoring, school advocacy, and mentoring.	This programme was designed for young people aged 6 to 11 who had come into contact with police but who are yet under the age of criminal responsibility.	Comparison group included received standard services.	Violence Crime and offending Behavioural difficulties General mental health	High	SNAP was effective at achieving an initial reduction in high-risk behaviours; however, long-term outcomes cannot be addressed using the data.

Camp et al. (1977)	US	RCT	Think Aloud was a 6-week training programme aimed at improving self-control.	6 to 8 year old boys who met the criteria on the aggressive scale were selected for participation.	Regular classroom instruction	Behavioural difficulties Ability to resolve conflicts	Low	There were no significant differences in teacher-rated classroom aggression between the Think Aloud group and the comparison group
Çelik et al. (2016)	Turkey	RCT & PE	First Step to Success Early Intervention Program (pre-school version: FSS-PSV) aimed to prevent problem/antisocial behaviours in children	Children aged 3 to 6 who were identified by teachers as having the highest risk of externalised problem behaviours	Business as usual	Behavioural difficulties Building and maintaining relationships (social-emotional skills)	Moderate	There were no significant differences in total problem behaviours or social skills between the FSS-PSV and the comparison group post-intervention

Collier (2002)	US	RCT	Sixteen pro-social skills training sessions to develop anger management, interpersonal, and problem-solving skills	Boys aged 7 to 14 who had received at least five referrals for behavioural problems during the past month	Business as usual	Behavioural difficulties	Moderate	After training, students who received pro-social skills training had fewer problem behaviours and fewer suspensions than students in the comparison group
Conduct Problems Prevention Research Group (2010)	US	RCT & PE	The Fast Track intervention (modelled on PATHS) lasted from 1st grade through to 10th grade and aimed to improve problem-solving, emotional resilience, peer relations, classroom environments, parenting practices, and home-school	Children who scored in the top 40% of their cohort for classroom conduct problems	Business as usual	Crime and offending	High	Fast Track students were less likely to be involved in a court-recorded arrest compared to the comparison group, but there was no difference between the groups in self-reported delinquency

			relations through a social skills curriculum, home visiting, parent training, academic tutoring, and mentoring					
Dong et al. (1979)	US	RCT	Group sessions to provide training in assertiveness	9th grade students with high overall aggression scores	Business as usual	Violence Behavioural difficulties Regulating and managing emotions	Low	Students in the intervention group scores higher on the assertiveness scale but did not differ from the comparison group on peer-rated aggression or self-rated aggression
Dupper and Krishef (1993)	US	RCT & PE	The social-cognitive skills group training programme aimed to help students acquire	6th and 7th grade students who had school	Business as usual	Regulating and managing emotions	Low	There were no significant differences in self-control between

			'school survival' skills to reduce the risk of school suspension	behaviour problems which placed them at-risk of school suspension				students who received the skills training and the comparison group
Edleson and Rose (1982)	US	RCT	Eight-week programme of group Social Skills Training sessions	4th, 5th and 6th grade students who were identified as having few friends or having difficulties in resolving peer conflicts	Wait-list control group	Violence Behavioural difficulties School engagement Ability to resolve conflicts Building and maintaining relationships (social-emotional skills)	Moderate	There were no significant post-training differences between students receiving the social skills training and the comparison group

						General mental health		
Feindler et al. (1984)	US	RCT	10 bi-weekly anger control training programme	Junior high school students identified as having been suspended for offences at least twice during the previous school year	Business as usual	Violence Crime and offending Behavioural difficulties Building and maintaining relationships (socio-emotional skills) Regulating and managing emotions	Low	There were no significant differences between students who received the anger control training and the comparison group at post-testing, on the teacher-rated self-control scale nor for locus of control. However, students who received the training showed higher generation of problem-solving responses compared to those

							in the comparison group
Hudley et al. (1998)	US	RCT	The Brain Power Programme, an attributional retraining intervention	African American boys who had been classified as aggressive	Business as usual Attention group receiving training in non-social problem solving skills	Building and maintaining relationships (socio-emotional skills) Regulating and managing emotions	Moderate	The intervention showed moderate effects for some student but was not evident at all for some students
Hudley and Graham (1993)	US	RCT	12-session attributional intervention to reduce hostility towards peers	African-American elementary school boys	Business as usual Attention training	Violence Behavioural difficulties Helping others (prosocial behaviours)	Moderate	Results showed that aggressive students who received the attributional intervention were less likely to presume hostile intent and less likely to endorse hostile

						School engagement Ability to resolve conflicts		retaliation and engage in hostile behaviours and were rated as less aggressive by their teacher.
Kazdin et al. (1987)	US	RCT & PE	Cognitive-behavioural Problem-solving skills training (PSST)	Psychiatric inpatient children aged 7 to 13	Business as usual Nondirective relationship therapy (RT)	Violence Behavioural difficulties School engagement Building and maintaining relationships (social-emotional skills) General Mental Health	Moderate	Children who received the training resulted in a significant decrease in externalising and aggressive behaviours and overall behaviour problems at home and school and an increase in prosocial behaviours when compared to the

							two comparison groups.
Lochmann et al. (1993)	US	RCT	Social Relations Training Program made up of four components: a) social problem solving, b) positive play training, c) group-entry skills training, d) dealing effectively with strong emotions	Fourth grade students identified as either aggressive, rejected or non-aggressive rejected	Business as usual	Behavioural difficulties Helping others (prosocial behaviours) School engagement Self-esteem	Moderate	The social relations training program was found to be effective only with the aggressive, rejected children
Mannarino et al. (1982)	US	RCT	Social competence program	Children aged 6.5 to 8.8 who were identified as high-risk of classroom behaviour problems	Business as usual	Behavioural difficulties School engagement	Moderate	Children who received the social competence program showed significantly greater gains in classroom adjustment and peer acceptance

Pedro-Carroll (1983)	US	RCT	The Children of Divorce Intervention Project (CODIP)	Elementary school children whose parents had separated or divorced	Delayed treatment control group	Behavioural difficulties School engagement Building and maintaining relationships (social-emotional skills) General Mental Health	Low	Participating children, as well as teachers, parents, and group leaders, reported the program to be effective across multiple dimensions
Prinz et al. (1994)	US	RCT & PE	Peer coping-skills (PCS) training	School-aged children identified as showing high rates of aggressive behaviour and competent-	Business as usual	Behavioural difficulties School engagement Building and maintaining relationships (social-	Low	Children identified with high rates of aggressive behaviour showed an increase in prosocial coping via information, improve social skills

				non aggressive children		emotional skills)		and reduced aggression
Quinn (2002)	US	RCT	Six-week cooperative learning intervention	Young boys identified as being at risk for developing antisocial behaviour patterns	Business as usual	Behavioural difficulties School engagement	Moderate	Children who received the cooperative learning intervention showed an increase in academic engagement however, the intervention did not produce significantly lower rates of externalizing antisocial behaviours or negative

								playground interactions
Schneider and Byrne (1987)	Canada	RCT	Individualised social Skill training (SST) programme for behaviour-disordered children	Children aged 7 to 13 from a treatment centre for children with behaviour disorders	Alternative treatment – non-individualized training (NIT) Wait-list control (WLC) group	Behavioural difficulties Helping others (prosocial behaviours)	Moderate	Children who received the individualised training programme indicated an increase in cooperative play after treatment but there were no significant differences in observed aggression between the groups
Shelton et al. (2000)	US	RCT	Multimethod psychoeducational intervention	Kindergarten students identified with elevated ADHD	Business as usual	Behaviour difficulties School engagement	Moderate	The two-year post-treatment follow-up results showed no difference between

				scores or impulsive hyperactive factors and on the ODD or conduct problem items		Building and maintaining relationships (social-emotional skills) General Mental Health		the classrooms that received the intervention and classrooms who did not
Smokowski et al. (2004)	US	RCT	Making Choices program, a school-based prevention programme designed to decrease child aggression and peer rejection through teaching children problem solving and relationship engagement skills	Children with a mean age of 8.5 years	Business as usual	Violence Behavioural difficulties Building and maintaining relationships (social-emotional skills)	Moderate	Children who received the Making Choices intervention had significantly higher scores for social contact, cognitive concentration and displayed significantly lower aggression than students who did

							not receive the intervention
Splett et al. (2015)	US	RCT & PE	Growing Interpersonal Relationships through Learning and Systemic Supports (GIRLSS), a Multisystemic 10-week intervention	Middle school girls	Business as usual	Behavioural difficulties	Moderate	Intervention participants had significantly more changes in the desired direction when compared to control participants across school counsellors and an averaged score of school counsellor and teacher-reported relational aggression
Stouwe et al. (2016)	Netherlands	QED	Tools4U, an individual social skills training with a parental component	Juvenile delinquents who received a behavioural training	Business as usual	Behavioural difficulties	Moderate	The intervention showed small effects and no effects compared

				sentence or community service order by disposal by the district attorney or by the juvenile court		Ability to resolve conflicts Building and maintaining relationships (social-emotional skills) Parenting practices Regulating and managing emotions Self-esteem		to treatment as usual group
Tanner and Holliman (1988)	US	RCT	Assertiveness social-skills training	Elementary school children nominated by	Business as usual	Violence Behavioural difficulties	Low	Children who received the intervention showed increase in

				teachers who had observed aggressive behaviours towards other children		Building and maintaining relationships (social-emotional skills)		frequency of cooperative interaction and a decrease in physical aggression
Webster-Stratton et al. (2001)	US	RCT & PE	Incredible Years Dinosaur Social Skills and Problem-Solving Curriculum	Children aged 4 to 8 with early onset conduct problems	Waiting-list control group	Violence Behavioural difficulties Building and maintaining relationships (social-emotional skills) Parenting practices	Low	Children who received the intervention showed significantly fewer externalising problems at home, less aggression at school, more prosocial behaviour with peers, and a more positive conflict management strategies when compared to the waiting-list control

								group. Additionally, children who received the intervention showed clinically significant improvements of aggressive and noncompliant behaviour

Appendix 4. Measured outcomes across included studies for effectiveness

YEF outcome framework category *Category description (as described in the framework)*	Measured outcomes *(descriptions by study authors)*	Studies
Violence	"Aggression: Self-rated"; "Aggression: peer-rated"; "Aggression"; "Mean double fines"; "Verbal behaviours – insult"; "Overt aggression"; "Self-reported relational aggression"; "Teacher reported relational aggression"; "School councillor reported relational aggression"; "Averaged teacher and school councillor-reported relational aggression"; "Physical aggression"	**(n = 9, k = 15)** Dong (1979); Edleson (1982); Feindler (1984); Hudley (1993); Kazdin (1987); Smokowski (2004); Splett (2015); Tanner (1988); Webster-Stratton (2001)
Crime and offending	"Had any charges"; "Charges against them"; "Rule breaking"; "Juvenile arrest index"; "Self-reported offence index"; "Mean double fines"	**(n = 3, k = 7)** Burke (2014); Conduct Problems Prevention Research Group (2010); Feindler (1984)

Behavioural difficulties	"Disruptive behaviour"; "Aggressive behaviour"; "Conduct problems"; "Externalizing"; "ODD (oppositional defiant disorder)"; "Conduct disorder"; "Aggression"; "Total problem behaviour scores"; "Problem behaviours in children"; "Aggression: Self-rated"; "Aggression: peer-rated"; "Mean single fines"; "Verbal behaviours – insult"; "Percent of clinically significant change in teacher-ratings of behaviour"; "Externalizing"; "Total behaviour problems"; "Acting out"; "Teacher-rated aggression"; "Aggression score"; "Home mean severity"; "Home CBCL aggression"; "Home CBCL delinquent"; "School CBCL-TRF: aggression"; "School CBCL-TRF: delinquent"; "Examiner ratings of behaviour"; "Overt aggression"; "Impulsivity"; "Physical aggression"; "Verbal aggression"; "Child conduct problems construct"; "Parent externalizing score"; "Child deviance"; "Teacher report of aggression"; "Total negative behaviour"; "Externalizing antisocial behaviour (EBA)"	**(n = 23, k = 50)** Block (1978); Burke (2014); Camp (1977); Çelik et al. (2016); Collier (2002); Dong (1979); Edleson (1982); Feindler (1984); Hudley (1993); Hudley (1998); Kazdin (1987); Lochmann (1993); Mannarino (1982); Pedro-Carroll (1983); Prinz (1994); Schneider (1987); Shelton (2000); Smokowski (2004); Stouwe (2016); Tanner (1988); Webster-Stratton (2001); Quinn (2002)
Bullying	"Self-reported RA (relational aggression)"; Teacher-reported RA; School counselor-reported RA"; "Averaged teacher and school counselor-reported RA"	**(n = 1, k = 4)** Splett (2015)

Building and maintaining relationships (social-emotional skills)	"Total social skills scores"; "Assertion scale"; "Sociometric ratings"; "CBCL: Social"; "Planned cell comparison between ARI group and ARC group social acceptance"; "Peer ratings"; "Peer sociability"; "Adaptive assertiveness"; "Observed communication effectiveness"; "Teacher-rated communication effectiveness"; "Teacher-rated social skills"; "Peer acceptance rating"; "home CBCL social problems"; "School SSRS: social skills"; "School SSRS: social problems"; "School CBCL-TRF: social problems"; "Social contact"; "Social competence"; "Peer acceptance"; "Social perspective taking – cognitive empathy"; "Critical reasoning – Self centering"; "Critical reasoning – blaming others"; "Critical reasoning – minimizing/mislabeling"; "Critical reasoning – assuming the worst"; "Developmental task-related skills – social acceptance"; "WALLY test variety of positive strategies"; "WALLY test number of positive strategies"	**(n = 12, k = 33)** Çelik et al. (2016); Dong (1979); Edleson (1982); Kazdin (1987); Lochmann (1993); Mannarino (1982); Pedro-Carroll (1983); Prinz (1994); Shelton (2000); Smokowski (2004); Stouwe (2016); Webster-Stratton (2001)
Regulating and managing emotions	"Locus of control scale"; "Teacher's Self-control rating scale"; "Self-control rating scale"; "Locus of control"; "Frustration tolerance"	**(n = 3, k = 5)** Dupper (1993); Feindler (1984); Pedro-Carroll (1983)
Ability to resolve conflict	"Means-ends problem solving inventory"; "Social problem-solving – confrontation"; "Social problem-solving – seeking social support"	**(n = 2, k = 3)** Feindler (1984); Stouwe (2016)

General mental health	"CBCL Internalizing total"; "CBCL withdrawn-depressed"; "CBCL anxious-depressed"; "CBCL somatic complaints"; "CSI: depression"; "CSI: Separation anxiety"; "CSI: General anxiety"; "ANX (Anxiety)"; "CBCL: Internalizing (post-treatment)"; "CARS FACTOR Shy-anxious"; "Teacher rated internalising problems"; "Home CBCL anxious"; "School CBCL-TRF: anxious"	**(n = 6, k = 14)** Burke (2014); Edleson (1982); Kazdin (1987); Pedro-Carroll (1983); Prinz (1994); Shelton (2000)
Helping others (prosocial behaviours)	"Cooperation score"; "Cooperative interactions"	**(n = 3, k = 4)** Schneider (1987); Tanner (1988)
School engagement	"Grade point average"; "Class cuts"; "School suspensions, Post"; "AD (Academic disability)"; "CBCL: School (post-treatment)": "SBCL: Global school adjustment (post treatment)"; "AML: Learning"; "CARS: Learning"; "CARS FACTOR learning"; "HRI FACTOR good student"; "HRI FACTOR follows rules"; "Academic engaged time (AET)"	**(n = 7, k = 14)** Block (1978); Collier (2002); Edleson (1982); Kazdin (1987); Mannarino (1982); Pedro-Carroll (1983); Quinn (2002)
Self-esteem	"Developmental task-related skills - Self-worth"	**(n = 1, k = 1)** Stouwe (2016)

Appendix 5. Characteristics of studies included in meta-review of implementation findings

Authors (Year)	Study Design	Number of studies included	Number and ages of children and young people involved	Countries of studies included	Implementation Factors Covered	AMSTAR 2 Rating
Blewitt et al. (2019)	Systematic review of Tier 2 (targeted) SEL interventions in early childhood education care settings.	19 studies	1,944 children participated	North America (n = 15), One study each in: Australia, China, Canada, and Belgium.	**Implementation Processes and Delivery Capacity:** • Intervention Providers and Training • Fidelity • Family Involvement **Delivery Characteristics:** • Delivery Method • Intervention Curriculum	Low
Goldberg et al. (2019)	Meta-analysis aimed at	45 studies included in	496,299 children and young	Twenty-five studies (56%) were	**Delivery Characteristics:**	Critically Low

	evaluating the effectiveness of interventions (universal and targeted) that use a whole-school approach to promote the social and emotional development of children and young people.	the meta-analysis representing 30 interventions. Of these 45 studies, 38 studies representing 14 interventions (47%) included a targeted component.	people between the ages of 4-16	carried out outside of the US, with the remaining 20 being US-based.	• Delivery Method	
Murano et al. (2020)	Meta-analysis summarising the effects of universal and Targeted SEL	48 articles (57 independent studies) including 33	15,498 preschool students	Not reported	**Implementation Processes and Delivery Capacity:** • Fidelity **Delivery Characteristics:**	Low

	interventions in preschool students.	universal and 15 Targeted SEL interventions.			• Intervention Curriculum	
Yang et al. (2019)	Meta-analysis synthesising findings from studies involving interventions that developed young children's social and emotional skills.	29 studies drawn from 14 reports	5,827 children in intervention groups and 10,845 in control groups. All children were aged 3-5 and from low-income families who were considered at risk for negative social and academic outcomes	Most studies were based in the United States (n=27), with one conducted in Australia and another in Germany	**Implementation Processes and Delivery Capacity:** • Fidelity **Delivery Characteristics:** • Duration/dosage • Intervention Curriculum	Critically Low

Appendix 6. Characteristics of studies on cost and children and young people's views

Authors (Year)	Country	Study Design	Intervention (Name and brief description)	Quality Level	Summary of Findings
Conduct Problems Prevention Research Group (2010)	USA	RCT and Process Evaluation	Fast Track preventative intervention; a multi-site, multi-component intervention program targeting children and young people at the highest risk for life-course persistent conduct problems.	Cost information provided online and not in the primary studies, meaning YEF-EQA could not be utilised. However, included studies linked to the cost data were rated as high quality	**Cost:** Costs associated with implementing Fast Track include start-up costs, training, implementation support, fidelity monitoring and resources for intervention providers and recipients.

Department for Education, (2016)	UK	Process Evaluation	Better by Design (BBD) aimed to build the social and emotional skills of young people in, or on the edge of care, who present with challenging behaviours	Very Low	**Cost:** When comparing the costs over the period April 2016 to February 2016 'with BBD' and 'without BBD', Frontier (2016) found that BBD has a net cost of £368,819 (i.e., additional costs exceed cost savings).
Splett et al. (2015)	USA	RCT and Process Evaluation	GIRLSS (Growing Interpersonal Relationships through Learning and Systemic Supports) - a 10-week intervention including school-based group counselling (*Learning*), and parent training and consultation (*Systematic Supports*) designed to target identified risk factors related to relational aggression	Moderate	**Children and young people's views:** Splett et al. (2015) reported on anonymous feedback from 22 child participants regarding the usefulness and quality of the group counselling sessions. Overall, the participants thought that the session topics were fairly easy to figure out and important and that group leaders did a good job during most sessions. Summative evaluation feedback from participants on the acceptability of all 10 weeks of the intervention indicated good to very good perceptions of the overall group, content, activities and group leaders.

Appendix 7. Moderator categories

Quality moderators

Study design

Category	Description
Randomised Controlled Trial (RCT)	A study in which participants are prospectively and randomly allocated between intervention and comparison groups
Quasi-experimental design (QED)	A study with a well-designed, closely matched but non-randomised comparison group, for example studies using propensity score matching, difference-in-differences, regression discontinuity, or synthetic control groups
Cluster Randomised Controlled Trial (cluster-RCT)	A study where groups, such as schools or classrooms, are randomly assigned to intervention or control conditions, rather than assigning individuals.

Study quality

Category	Description
Low	Study quality assessed using the YEF-EQA tool

Moderate	
High	

Setting-level moderators

Setting

Category	Description
Clinic or inpatient setting	Intervention took place in a clinic or inpatient psychiatric facility
Community	Intervention took place in a community setting
School	Intervention took place in school
School and home	Intervention included sessions at school and at home
Not specified	Study did not specify where the intervention took place

Intervention-level moderators

Duration

Category	Description
< 1 month	
1 – 3 months	
3 – 6 months	Length of time that the intervention lasts, as reported by the study
6 – 12 months	
> 5 years	
Unclear	

Intensity

Category	Description
Daily	
2 – 4 per week	Frequency of intervention sessions

Weekly	
Not stated / unclear	

Provider

Category	Description
Counsellor, therapist, psychologist or other health professional	Professional who delivered the intervention sessions
Researcher	
Service provider or unspecified leader	
Social worker	
Teacher or other education professional	
Not stated	

Provider training

Category	Description
Yes – manual or pre-existing qualification	No formal or bespoke training but providers received a manual or were required to have a specific qualification
Yes – pre-intervention	Providers received training before implementing the intervention
Yes – with ongoing supervision	Providers received training and ongoing supervision visits, calls or workshops to monitor fidelity and provide support
Not stated	Study does not specify what training providers received

Delivery method

Category	Description
Face to face	Intervention was delivered only through face to face sessions

Face to face and phone	Intervention was delivered through face to face sessions and through phone calls
Not specified	Study does not specify what training providers received

Individual or group

Category	Description
Group	Intervention was delivered only through group sessions
Individual	Intervention was delivered only through one-to-one sessions
Group and individual	Intervention was delivered through a mix of group and one-to-one sessions
Not specified	Study does not specify how the intervention was delivered

Population-level moderators

Age band

Category	Description
0 – 9	Pre-school and early childhood / early primary school

10 – 14	Middle childhood / early adolescence
15 – 17	Adolescence / middle to high school
Mixed	Intervention targets a broad range of ages

Ethnicity

Category	Description
Majority White Sample	>85% of sample identified as white
Some Diversity	Between 15–49% of participants identified as Black and Global Majority (i.e., Black, Hispanic/Latinx, Asian, Indigenous, etc.)
Diverse/Balanced	50% or more Black and Global Majority youth in the sample
Not stated/unclear	Study does not specify the ethnicity of participants

Sex/Gender

Category	Description
Girls only	Only girls were included
Boys only	Only boys were included
Majority boys	Sample included >70% boys
Mixed sex	40–70% male and 30–60% female
Not stated / unclear	Study did not report the gender breakdown of participants

Outcome moderators

Outcome type

Category	Description
Building and maintaining relationships	Social-emotional skills for listening, cooperating, and understanding emotions.

Regulating and managing emotions	Also called 'emotion regulation'. Having the skills and techniques to manage feelings and reactions to situations and events, reducing the intensity, duration, and impact of such feelings.
Ability to resolve conflicts	Skills to de-escalate conflicts and arguments.
Behavioural difficulties	Disruptive or aggressive behaviours due to distress or needs, excluding hyperactivity and neuro-diverse conditions.
Crime and offending	Criminal behaviours including non-violent acts (e.g., shoplifting), sexually violent crimes, and violent acts (e.g., assault, robbery).
General mental health	Encompasses psychological mental health, which specifically relates to an individual's cognitive and emotional functioning, including their ability to manage thoughts, behaviours, and cope with life's challenges. Includes mental health conditions like anxiety, depression, PTSD
Helping others (prosocial behaviour)	Engaging in positive, prosocial behaviours such as helping, comforting, or sharing.
School engagement	Indicators of active participation in education, including attendance, positive connections with peers and staff, avoiding exclusions, and achieving good grades
Self-esteem	Viewing yourself positively, including confidence in your abilities, appearance, and self-worth.

Data source

Category	Description
Administrative	From school or criminal justice system records e.g. exam grades, school attendance
Observation	Measured by the researcher directly observing the behaviour
Parent report	Reported by a parent through a survey or interview
Peer report	Reported by a peer through a survey or interview
Self-report	Reported by the young person through a survey or interview
Teacher report	Reported by a teacher through a survey or interview
Not stated	Study does not report how the outcome was measured