Appendix 6. Evidence summary RQ2

Question (PICO)	What is the effect of educational programs aimed at dealing with mental health (problems), social skills, recognizing signs of mental health problems or referring to professional help (I), on the mental health or mental health knowledge/skills/attitudes and/or behavior (O) in children and adolescents between 5 and 18 years old (P)?
Search Strategy	See Appendix 1
Search date	7/7/2021
In/Exclusion criteria	See Appendix 2

Characteristics of included studies

Author, year, Country	Study design	Population	Comparison	Remarks
Aguirre Velasco, 2020, UK	Systematic review of 36 experimental studies, but data of only 3 studies (2 cluster RCTs and 1 RCT) were eligible for this summary	Studies included 9154 adolescents between 10 and 19 years old. Studies were conducted in Australia, UK and USA. [data from the 55 studies about barriers and facilitators were not extracted; data from 33 interventions studies were not extracted (see "Comparison")].	**Interventions targeting help-seeking** for common mental health problems: Classroom-based interventions based on psychoeducation, with a focus on general mental health knowledge or specifically addressing stigma: • MAKINGtheLINK program (1 study): 5 interactive activities provided by trained teachers to year 9 students (14-15 years): ○ Recognizing when a friend needs help ○ Identifying types of helpers available as well as obligations related to professional confidentiality ○ Understanding myths and facts about substance use and mental health ○ Identifying and overcoming barriers to professional help-seeking ○ How to assist a friend to access professional help and how to access reliable sources of help. After one month, the interactive activities were followed by a booster session Control: waiting list • Dissemination of a student booklet about	Search date: April 2019 Databases searched: MEDLINE, Embase, PsycINFO, Web of Science, Google scholar, Open Grey This systematic review was used as a source of studies as not all studies fulfilled our selection criteria. We used the following studies: Lubman 2016, Sharpe 2016, Saporito 2013. As Lubman 2016 is a study protocol, we searched for the corresponding study, which was published as Lubman 2020. Outcome: help-seeking behaviour, measured via: • Simplified version of the Actual Help Seeking Questionnaire (AHSQ) adapted to include substance use and mental health, outcome measured at 12-month follow-up (Lubman 2016, 2020)

			help-seeking and self-management support (1 study) to year 7 students (11-12 years); control: waiting list • School-based interactive session and video with case example, with focus on stigma and myths regarding mental illness, to adolescents from public high school (1 study); control: educational presentation with content unrelated to mental health [Interventions targeting help-seeking for specific mental health problems such as suicide or depression were not included; outreach interventions (delivered by professional health care providers) were not included]	• A 4-point scale to assess help-seeking behaviour, outcome measured at 12-month follow-up (Sharpe, 2016) • A 7-point Likert Scale regarding 'Willingness to seek treatment', timing outcome measurement not specified (Saporito, 2013)
de Mooij, 2020, The Netherlands	Systematic review of 66 RCT studies and 32 non-RCTs/observational studies	Studies included 71,226 children and adolescents between 3 and 17 years old from a general population. Special populations (e.g. children with ASD or ADHD) are excluded. Studies were conducted in Asia, Australia, Eastern and Western Europe, India, Middle East and North America.	**Social skills training programs vs no social skills training programs** Intervention: 60 unique social skills training programs aimed at teaching or developing children's adaptive social behaviour to improve their success in social interactions**:** • 19 social-emotional learning programs (42 studies) • 6 programs targeting bullying behaviour (9 studies) • 10 programs targeting (social) anxiety (12 studies) • 11 programs targeting disruptive behaviour (12 studies) • 5 programs targeting resilience and self-esteem (9 studies) • 9 programs targeting prosocial interactions (14 studies) Controls: "care as usual" or "no treatment" or "attention control" **Inclusion of specific training components in**	Search date: October 2018 Databases searched: PsycINFO, MEDLINE, Scopus, ERIC and Google Scholar Outcomes: Interpersonal skills, emotional skills, peer relationship problems, internalizing problem behaviour and externalizing problem behaviour The systematic review included studies in which the post-intervention measurements ranged from <6 months to >2 years.

			the social skills training program vs without training components Training components: • Psychoeducation • Psychophysical components • Skill-building components **Inclusion of specific booster components in the social skills training program vs without booster components** Booster components: • Class management • Rewarding • Goal setting • Generalization • Coaching • (Self-)monitoring **Influence of specific program characteristics of the social skills training program** • Setting: indicated program vs universal program • Duration of program: 1-9 weeks vs 10-11 weeks vs 12-16 weeks vs 17-26 weeks vs >27 weeks • Type of trainer: school personnel vs mental health professional vs non-school personnel • Schooling required for trainer: schooling vs no schooling vs not specified • Mode of delivery: computer program vs face-to-face • Age of participants: primary school age vs secondary school age vs children and adolescents	
Mertens, 2020, The Netherlands	Systematic review of 99 experimental (RCT studies and non-RCTs) and observational studies	Studies included 97,884 adolescents between 11 and 18 years old (average age 13.70 years). Studies were conducted in USA, Canada,	**Universal secondary school-based programs aiming to stimulate students intra- and interpersonal domains vs control** Universal secondary school-based interventions were defined as interventions delivered	Search date: April 2019 Databases searched: PsycINFO, PubMed, ERIC and Cochrane CENTRAL Systematic review includes 104

		Europe, Australia, Asia and Africa.	to students during regular school hours, targeting all students. The intrapersonal domain was defined as managing one's own feelings, emotions, and attitudes pertained to the individual self in which one can experience competencies (e.g., resilience, self-esteem, self-regulation, general wellbeing) and problems (e.g., internalizing behavior). The interpersonal domain was defined as the ability of an individual to build and maintain positive relationships with others and understanding social situations, roles and norms, and respond appropriately in which one can experience competencies (e.g., sexual health, social competence, positive school climate) and problems (e.g., aggression, bullying). The following components of the programs were separately analysed: **Content components** (a definition of each component is presented in appendix 1 at the end of this evidence summary): • Emotion regulation • Assertiveness • Self-efficacy • Self-control • Insight building • Cognitive coping • Relaxation • Social skills • Problem solving • Peer resistance **Instructional components:** • Practice • Modelling • Discussion • Goal setting • (Self-)monitoring • Multimedia • Homework • Didactic instruction Controls:	publications reporting on 99 unique experimental studies. Outcomes: • Intrapersonal domain with the following subdomains: resilience, self-esteem, self-regulation, general wellbeing and internalizing problems • Interpersonal domain with the following subdomains: social competence, school climate, aggression and bullying [outcome sexual health was not extracted] Outcomes were measured within 6 months post-intervention.

			• In 47 studies, an active control group was used (i.e. Care As Usual or another intervention). • In 57 studies, a passive control group was used (i.e. waiting list or no intervention). [Interventions targeting the inclusion of structural components were not extracted]	
Moy, 2018, USA	Systematic review of 27 experimental (RCT studies and non-RCTs) and observational studies	Studies included 18,847 children between 4 and 14 years old. Studies were conducted in Germany, Guatemala, Norway and USA.	**The Second Step, a universal social-emotional learning program vs no Second Step program** Second Step is based on a blend of theoretical foundations, including the cognitive-behavioural model, social learning theory, social information processing and verbal self-regulation. Control: not specified **Additional comparison:** Second Step program in pre-kindergarten vs Second Step program in multiple grades	The systematic review included published and unpublished research reports from 1984-2016. Search date not reported. Databases searched: Academic Search Complete, Child Development and Adolescent Studies, Education Research Complete, Education Administration Abstracts, ERIC, OmniFile Full Test Select, Professional Development Collection, PsycINFO, PsycArticles, Social Work Abstracts, Teacher Reference Center, Social Work Reference Center, ProQuest Dissertations and Theses. Outcomes: • Antisocial behaviour (physical aggression, bullying, peer victimization, sexual violence and other antisocial behaviours) • Prosocial behaviour (coping, cooperative behaviour, conflict resolution,

				positive social behaviour, social problem solving, empathy and social competence) • Content knowledge of Second Step lessons (i.e. knowledge or attitudes about violence or violence prevention) Timing outcome measurement not specified.
Ng, 2020a, Singapore	Systematic review of 2 RCTs and 15 cluster RCTs	Studies included 35,694 children and adolescents between 10 and 18 years old. Studies were conducted in Australia, Austria, Belgium, Brazil, Finland, Germany, Italy, Romania, South Africa, Spain, USA	**Anti-(cyber)-bullying program vs no anti-(cyber)-bullying program** • Educational programs with a set curricula focused on bullying prevention. • Seven intervention programs were designed to reduce traditional bullying. • Five intervention programs were designed to reduce cyberbullying. • One antibullying program was used for both traditional and cyberbullying. Controls received usual lessons, treatment-as-usual bullying prevention programs, placebo interventions or waiting list control. **Influence of specific program characteristics:** • Personnel delivering program: ○ Teachers/school staff vs control ○ Content expert vs control ○ Teachers/school staff vs content expert • Location of program: ○ School vs control ○ Classroom vs control ○ School vs classroom • Program duration:	Search date: June 2019 Databases searched: PubMed, Embase, PsycINFO, Cumulative Index to Nursing and Allied Health Literature, Google Scholar and ProQuest Dissertations and Theses Cluster RCTs were randomized with schools or classes as common clusters. Outcomes: • Traditional bullying victimization • Traditional bullying perpetration • Cyberbullying victimization • Cyberbullying perpetration Immediate post-intervention values were of primary interest in the systematic review as not all studies conducted follow-up measurements.

			○ Up to 3 months vs control ○ 3<X>6 months vs control ○ More than 6 months vs control ○ Up to 3 months vs 3<X>6 months vs more than 6 months • Parental involvement: ○ Parental involvement vs control ○ No parental involvement vs control ○ Parental involvement vs no parental involvement	
Ng, 2020b, Singapore	Systematic review of 14 experimental studies, but data of only 2 cluster-randomised crossover trials were eligible for this summary	Studies included 1605 adolescents between 15 and 17 years old. Studies were conducted in Australia. [data from 12 experimental studies were not extracted (see "Comparison")].	**Teen Mental Health First Aid program (tMHFA) vs physical first aid training (2 studies):** • tMHFA is a program delivered to students from secondary schools, using age-appropriate learning materials: • Three 75-minute classroom sessions presented by trained external instructors according to a manualised curriculum to students of 15-17 years old. • Training involved a PowerPoint presentation, videos, role-plays, group discussion, small group and workbook activities [Interventions targeting the effect of YMHFA on adults were not included; uncontrolled studies were not included]	Search date: September 2020 Databases searched: PubMed, Embase, PsycINFO, ERIC and Cochrane CENTRAL. This systematic review was used as a source of studies as not all studies fulfilled our selection criteria. We used the following studies: Hart 2018 and Hart 2020. The papers of Hart 2018 and Hart 2020 refer to the same intervention study with the same population, but present different outcomes. Outcomes: • recognition of mental illness • mental health knowledge • stigma • helping intentions • confidence Outcomes were measured at 1 week post-intervention and at 12-month follow-up via a depression (post-intervention and follow-up) and

				anxiety vignette (only post-intervention).
Russell, 2021, USA	Systematic review of 10 RCTs	Studies included 15953 adolescents between 11 and 18 years old. Studies were conducted in the USA.	**Adolescent Dating Violence (ADV) prevention program vs no program or waiting list control** All included studies used an educational program as the intervention. Examples of included ADV prevention programs: • Teen Choices (1 study): a 3-session online program that delivers assessments and individualized guidance matched to dating history, dating violence experiences, and stage of readiness for using healthy relationship skills. • Building a Lasting Love (1 study): 4 sessions focussing on i.a. signs of healthy versus unhealthy romantic relationships, healthy couple communication, assertiveness, problem-solving techniques, and conflict management strategies. • Fourth R: Skills for Youth Relationships (1 study): a 21-lesson curriculum focussing on dating violence and relationship skills.	Search date: April 2019 Databases searched: Academic Search Complete, CINAHL, ERIC, Humanities International Complete, MEDLINE, PsycINFO, Psychology and Behavioural Sciences Collection, Social Work Abstracts, SocINDEX and Cochrane CENTRAL. Outcomes: • Perpetration: overall, emotional, physical, sexual and threatening • Victimization: overall, emotional, physical, sexual and threatening The systematic review included studies in which the follow-up measurements ranged from 6 weeks to 2.5 years.
Seedaket, 2020, Thailand	Systematic review of 7 experimental studies, but data of only 5 studies (2 RCTs, 2 cluster RCTs and 1 non-RCT) were eligible for this summary	Studies included 9432 adolescents between 10 and 19 years old. Studies were conducted in Australia, Canada, Norway UK and USA. [data from the two studies about community-based interventions were not	**Mental Health Literacy (MHL) programs:** School-based intervention with two strategies: • Education stand-alone intervention (4 studies): ○ "Mental Health for Everyone" program provided by trained teachers or researchers vs classes as usual. ○ "HeadStrong" program provided by schoolteacher	Search date: December 2019 Databases searched: ScienceDirect, Scopus, PubMed, Cochrane and CINHAL. This systematic review was used as a source of studies as not all studies fulfilled our selection criteria. We used the following studies:

		extracted (see "Comparison")].	vs classes as usual. ○ "Mental Health and High School Curriculum Guide" delivered by trained teachers vs teaching as usual. ○ "Adolescent Depression Awareness Program (ADAP)" delivered by trained teachers vs routine health curriculum. • Educational program on mental health complemented with contact module vs educational program on mental health alone (1 study): The intervention included an interactive session with a young person with experience of mental illness. The main procedure of all studies was focused on imparting the information about mental illness and available mental health resources for adolescents. The MHL programs included interactive teaching methods, use of various media such as group discussion, videos and movies. [Community-based interventions were not included]	Skre 2013, Perry 2014, Milin 2016, Swartz 2017 and Chisholm 2016. Outcomes: mental health knowledge, attitudes or stigma, help-seeking efficacies, resilience and emotional well-being. Measured via: • A 66-item questionnaire, outcome measured at 2-month follow-up (Skre, 2013). • A Depression Literacy Scale, a Depression Stigma Scale and the Inventory of Attitudes towards Seeking Mental Health Services, outcome measured at post-intervention and 6-month follow-up (Perry, 2014) • A questionnaire consisting of 15 multiple choice questions (mental health knowledge) and 8 statements rated on a Likert scale (stigma), outcome measured at post-intervention (Milin, 2016). • An Adolescent Depression Knowledge Questionnaire and a modified Reported and Intended Behaviour Scale, outcome measured at 6-week postintervention and 4-month

				follow-up (Swartz, 2017). • The Mental Health Knowledge Schedule, the Reported and Intended Behaviour Scale (stigma), a questionnaire on attitudes to help-seeking, a 15-item version of the Resilience Scale and Strength and Difficulties questionnaire (emotional well-being), outcome measured at 2-week follow-up (Chisholm, 2016).
Tejada-Gallardo, 2020, Spain	Systematic review of 9 experimental studies (7 RCTs and 2 non-RCTs)	Studies included 4898 adolescents between 10 and 18 years old. Studies were conducted in Australia, Israel, Italy, Portugal, UK and USA.	**Multicomponent Positive Psychology Intervention program (MPPI) vs placebo program or waiting list control** The MPPI program in the included studies is based on a variety of individual exercises targeting at least two theoretically relevant well-being components (subjective and psychological well-being). Two of the included studies were combined with another technique, i.e. acceptance and commitment therapy and positive youth development. The number of sessions varied between 6 and 18 and the duration of the program ranged from 4 to 30 weeks. Controls: • Placebo (5 studies): anxiety-management school protocol • Waiting list (4 studies)	Search date: July 2019 Databases searched: PsycINFO, PubMed and Scopus databases. Outcomes: • Subjective well-being: satisfaction with life, positive affect, student's life satisfaction. • Psychological well-being: i.a. self-efficacy, autonomy, environmental well-being, personal growth, positive relationships, purpose in life and self-acceptance. Six of the included studies also performed follow-up measurements which ranged from 5 weeks to 12 months.

Synthesis of findings

Outcome	Comparison	Effect Size	#studies, # participants	Reference
Interventions targeting help-seeking for common mental health problems				
Help-seeking behaviour at 12 months	MAKINGtheLINK program vs waiting list	Not statistically significant: 243/1130 vs 271/1317 OR: 1.00, 95%CI [0.70;1.42] ¥ (p=0.99)	1, 1130 vs 1317	Aguirre Velasco, 2020 (Lubman, 2016, Lubman 2020)
Help-seeking from formal sources (vs informal sources)		Statistically significant: 109/243 vs 83/271 § OR: 1.81, 95%CI [1.19;2.75] (p=0.005) *In favour of help-seeking from formal sources*	1, 243 vs 271	
Help-seeking behaviour	Student booklet vs waiting list	Not statistically significant: £† OR: 1.01, 95%CI [0.95;1.07] (p>0.05)	1, 4865 vs 1686	Aguirre Velasco, 2020 (Sharpe, 2016)
Willingness to seek help	Stigma-based interactive session and video vs another presentation	Statistically significant: 24.6±5.6 vs 23.6±5.9 $F_{(1, 146)}$=6.64, η^2_p=0.04 ¤ (p=0.01) *In favour of stigma-based interactive session and video*	1, 80 vs 76 §	Aguirre Velasco, 2020 (Saporito, 2013)
Social skills training programs				
Interpersonal skills	Social skills training programs vs no social skills training programs	Statistically significant: £ SMD: 0.386, 95%CI [0.288;0.484] (p<0.001) *In favour of social skills training programs*	68, £££	de Mooij, 2020
Emotional skills		Statistically significant: £ SMD: 0.328, 95%CI [0.225;0.431] (p<0.001) *In favour of social skills training programs*	38, £££	
Peer relationships problems		Statistically significant: £ SMD: 0.255, 95%CI [0.095;0.415] (p=0.002) *In favour of social skills training programs*	27, £££	
Internalizing problem behaviour		Statistically significant: £ SMD: 0.233, 95%CI [0.159;0.306] (p<0.001) *In favour of social skills training programs*	52, £££	
Externalizing problem behaviour		Statistically significant: £ SMD: 0.172, 95%CI [0.078;0.266] (p<0.001) *In favour of social skills training programs*	60, £££	
Interpersonal skills	Social-emotional learning programs vs no social-emotional learning programs	Statistically significant: £ SMD: 0.290, 95%CI [0.187;0.393] (p<0.001)	42, 24278 vs 21276	

		In favour of social-emotional learning programs		
Emotional skills		Statistically significant: £ SMD: 0.249, 95%CI [0.142;0.355] (p<0.001) *In favour of social-emotional learning programs*		
Peer relationship problems		Not statistically significant: £† SMD: 0.249, 95%CI [0.142;0.355] (p<0.001)		
Internalizing problem behaviour		Statistically significant: £ SMD: 0.171, 95%CI [0.000;0.342] (p=0.05) *In favour of social-emotional learning programs*		
Externalizing problem behaviour		Not statistically significant: £† SMD: 0.095, 95%CI [-0.007;0.197] (p=0.067)		
Interpersonal skills	Programs targeting bullying behaviour vs no programs targeting bullying behaviour	Statistically significant: £ SMD: 0.709, 95%CI [0.367;1.050] (p<0.001) *In favour of programs targeting bullying behaviour*	9, 2975 vs 2524	
Emotional skills		Statistically significant: £ SMD: 0.864, 95%CI [0.742;0.987] (p<0.001) *In favour of programs targeting bullying behaviour*		
Peer relationship problems		Not statistically significant: £† SMD: 0.666, 95%CI [-0.010;1.342] (p=0.053)		
Internalizing problem behaviour		Statistically significant: £ SMD: 0.846, 95%CI [0.583;1.110] (p<0.001) *In favour of programs targeting bullying behaviour*		
Externalizing problem behaviour		Statistically significant: £ SMD: 0.774, 95%CI [0.099;1.448] (p<0.05) *In favour of programs targeting bullying behaviour*		
Interpersonal skills	Programs targeting (social) anxiety vs no programs targeting (social) anxiety	Not statistically significant: £† SMD: 0.259, 95%CI [-0.248;0.766] (p=0.203)	12, 652 vs 524	
Emotional skills		Statistically significant:		

		£ SMD: 0.264, 95%CI [0.082;0.447] (p<0.05) *In favour of programs targeting (social) anxiety*		
Internalizing problem behaviour		<u>Statistically significant:</u> £ SMD: 0.384, 95%CI [0.134;0.634] (p<0.01) *In favour of programs targeting (social) anxiety*		
Externalizing problem behaviour		Not statistically significant: £† SMD: 0.402, 95%CI [-1.852;2.692] (p=0.256)		
Interpersonal skills	Programs targeting disruptive behaviour vs no programs targeting disruptive behaviour	<u>Statistically significant:</u> £ SMD: 0.253, 95%CI [0.127;0.378] (p<0.001) *In favour of programs targeting disruptive behaviour*	12, 1016 vs 759	
Emotional skills		Not statistically significant: £† SMD: 0.494, 95%CI [-0.316;1.304] (p=0.166)		
Peer relationship problems		Not statistically significant: £† SMD: 0.219, 95%CI [-0.740;1.179] (p=0.429)		
Internalizing problem behaviour		<u>Statistically significant:</u> £ SMD: 0.348, 95%CI [0.092;0.603] (p<0.05) *In favour of programs targeting disruptive behaviour*		
Externalizing problem behaviour		<u>Statistically significant:</u> £ SMD: 0.245, 95%CI [0.086;0.405] (p<0.01) *In favour of programs targeting disruptive behaviour*		
Interpersonal skills	Programs targeting resilience and self-esteem vs no programs targeting resilience and self-esteem	Not statistically significant: £† SMD: -0.006, 95%CI [-0.094;0.83] (p=0.893)	9, 3507 vs 2234	
Emotional skills		<u>Statistically significant:</u> £ SMD: 0.287, 95%CI [0.015;0.559] (p<0.05) *In favour of programs targeting resilience and self-esteem*		
Internalizing problem behaviour		<u>Statistically significant:</u> £ SMD: 0.112, 95%CI [0.046;0.178]		

		(p<0.01) *In favour of programs targeting disruptive behaviour*		
Externalizing problem behaviour		Not statistically significant: £† SMD: 0.165, 95%CI [-0.795;1.125] (p=0.273)		
Interpersonal skills	Programs targeting prosocial interactions vs no programs targeting prosocial interactions	<u>Statistically significant:</u> £ SMD: 0.660, 95%CI [0.273;1.048] (p<0.01) *In favour of programs targeting prosocial interactions*	14, 5486 vs 4047	
Emotional skills		Not statistically significant: £† SMD: 0.098, 95%CI [-0.088;0.284] (p=0.281)		
Peer relationship problems		Not statistically significant: £† SMD: 0.149, 95%CI [-0.373;0.670] (p=0.344)		
Internalizing problem behaviour		<u>Statistically significant:</u> £ SMD: 0.198, 95%CI [0.028;0.369] (p<0.05) *In favour of programs targeting prosocial interactions*		
Externalizing problem behaviour		Not statistically significant: £† SMD: 0.045, 95%CI [-0.145;0.235] (p=0.638)		
Interpersonal and emotional skills (estimate of effect size, 95%CI)	Social skills training programs with psychoeducation components vs without psychoeducation components	<u>Statistically significant:</u> 0.415, 95%CI [0.331;0.499] vs 0.181, 95%CI [0.014;0.348] ££ (p=0.015) *In favour of social skills training programs with psychoeducation components*	77, £££	
Peer relationship problems (estimate of effect size, 95%CI)		Not statistically significant: 0.282, 95%CI [0.097;0.468] vs 0.170, 95%CI [-0.163;0.503] ££† (p=0.558)	27, £££	
Internalizing problem behaviour (estimate of effect size, 95%CI)		Not statistically significant: 0.239, 95%CI [0.159;0.320] vs 0.199, 95%CI [0.013;0.384] ££† (p=0.691)	52, £££	
Externalizing problem behaviour (estimate of effect size, 95%CI)		Not statistically significant: 0.189, 95%CI [0.085;0.294] vs 0.099, 95%CI [-0.116;0.315] ££† (p=0.460)	60, £££	
Interpersonal and emotional skills (estimate of effect size, 95%CI)	Social skills training programs with psychophysical components vs without	Not statistically significant: 0.353, 95%CI [0.246;0.461] vs 0.388, 95%CI [0.275;0.501] ££† (p=0.655)	77, £££	

Peer relationship problems (estimate of effect size, 95%CI)	psychophysical components	Not statistically significant: 0.321, 95%CI [0.106;0.536] vs 0.172, 95%CI [-0.068;0.413] ££† (p=0.361)	27, £££	
Internalizing problem behaviour (estimate of effect size, 95%CI)		Not statistically significant: 0.223, 95%CI [0.119;0.326] vs 0.244, 95%CI [0.138;0.350] ££† (p=0.778)	52, £££	
Externalizing problem behaviour (estimate of effect size, 95%CI)		Not statistically significant: 0.128, 95%CI [-0.006;0.262] vs 0.214, 95%CI [0.083;0.346] ££ (p=0.365)	60, £££	
Interpersonal and emotional skills (estimate of effect size, 95%CI)	Social skills training programs with skill-building components vs without skill-building components	Not statistically significant: 0.372, 95%CI [0.292;0.451] vs 0.314, 95%CI [-0.110;0.739] ££† (p=0.794)	77, £££	
Peer relationship problems (estimate of effect size, 95%CI)		Not statistically significant: 0.254, 95%CI [0.086;0.422] vs 0.286, 95%CI [-0.364;0.936] ££ (p=0.924)	27, £££	
Internalizing problem behaviour (estimate of effect size, 95%CI)		Not statistically significant: 0.246, 95%CI [0.171;0.321] vs 0.017, 95%CI [-0.281;0.315] ££† (p=0.143)	52, £££	
Externalizing problem behaviour (estimate of effect size, 95%CI)		Not statistically significant: 0.169, 95%CI [0.073;0.265] vs 0.243, 95%CI [-0.215;0.700] ££† (p=0.756)	60, £££	
Interpersonal and emotional skills (estimate of effect size, 95%CI)	Social skills training programs with class management as a booster component vs without class management as a booster component	Not statistically significant: 0.375, 95%CI [0.293;0.457] vs 0.322, 95%CI [0.069;0.575] ££† (p=0.694)	77, £££	
	Social skills training programs with rewarding as a booster component vs without rewarding as a booster component	Not statistically significant: 0.295, 95%CI [0.159;0.431] vs 0.404, 95%CI [0.310;0.497] ££† (p=0.195)	77, £££	
	Social skills training programs with goal setting as a booster component vs without goal setting as a booster component	Not statistically significant: 0.335, 95%CI [0.170;0.501] vs 0.379, 95%CI [0.291;0.468] ££† (p=0.644)	77, £££	
	Social skills training programs with generalization as a booster component vs without generalization as a booster component	Not statistically significant: 0.368, 95%CI [0.274;0.462] vs 0.374, 95%CI [0.234;0.513] ££† (p=0.948)	77, £££	

	Social skills training programs with coaching as a booster component vs without coaching as a booster component	Not statistically significant: 0.368, 95%CI [0.279;0.457] vs 0.377, 95%CI [0.214;0.539] ££† (p=0.922)	77, £££	
	Social skills training programs with (self-)monitoring as a booster component vs without (self-)monitoring as a booster component	Not statistically significant: 0.316, 95%CI [0.141;0.490] vs 0.383, 95%CI [0.296;0.471] ££† (p=0.496)	77, £££	
	Influence of the setting of social skills training program: indicated program vs universal program	Not statistically significant: 0.356, 95%CI [0.155;0.556] vs 0.372, 95%CI [0.288;0.457] ££† (p=0.883)	77, £££	
	Influence of the duration of social skills training program: 1-9 weeks vs 10-11 weeks vs 12-16 weeks vs 17-26 weeks vs >27 weeks	Not statistically significant: 0.275, 95%CI [0.091;0.460] vs 0.452, 95%CI [0.313;0.592] vs 0.510, 95%CI [0.312;0.708] vs 0.376, 95%CI [0.199;0.552] vs 0.208, 95%CI [0.046;0.370] ££† (p=0.089)	74, £££	
	Influence of the type of trainer in the social skills training program: school personnel vs mental health professional vs non-school personnel	Not statistically significant: 0.343, 95%CI [0.255;0.431] vs 0.426, 95%CI [0.246;0.605] vs 0.403, 95%CI [-0.008;0.813] ££† (p=0.704)	74, £££	
	Influence of schooling required for trainer in the social skills training program: schooling vs no schooling vs not specified	Not statistically significant: 0.377, 95%CI [0.281;0.473] vs 0.390, 95%CI [0.217;0.564] vs 0.305, 95%CI [0.093;0.516] ££† (p=0.802)	77, £££	
	Influence of mode of delivery of the social skills training program: computer program vs face-to-face	Not statistically significant: 0.525, 95%CI [0.079;0.972] vs 0.365, 95%CI [0.286;0.444] ££† (p=0.487)	77, £££	
	Influence of the age of participants of the social skills training program: primary school age vs secondary school age vs children and adolescents	Not statistically significant: 0.391, 95%CI [0.295;0.486] vs 0.428, 95%CI [0.136;0.720] vs 0.298, 95%CI [0.142;0.455] ££† (p=0.565)	77, £££	
Programs aiming to stimulate students intra- and interpersonal domains				
Intrapersonal domain	Universal secondary school-based programs	Statistically significant: £ SMD: 0.19, 95%CI [0.13;0.25]	£££†	Mertens, 2020

	aiming to stimulate students intra- and interpersonal domains vs control	(p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
Resilience		Not statistically significant: £† SMD: 0.06, 95%CI [-0.01;0.14] (p>0.05)		
Self-esteem		<u>Statistically significant:</u> £ SMD: 0.25, 95%CI [0.11;0.39] (p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
Self-regulation		<u>Statistically significant:</u> £ SMD: 0.21, 95%CI [0.08;0.33] (p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
General wellbeing		<u>Statistically significant:</u> £ SMD: 0.13, 95%CI [0.08;0.19] (p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
Internalizing problems		<u>Statistically significant:</u> £ SMD: 0.19, 95%CI [0.10;0.29] (p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
Interpersonal domain		<u>Statistically significant:</u> £ SMD: 0.15, 95%CI [0.10;0.19] (p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
Social competence		<u>Statistically significant:</u> £ SMD: 0.16, 95%CI [0.10;0.23] (p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
School climate		Not statistically significant: £† SMD: 0.24, 95%CI [-0.11;0.58] (p>0.05)		
Aggression		<u>Statistically significant:</u> £ SMD: 0.10, 95%CI [0.03;0.17] (p<0.05)		

		In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains		
Bullying		<u>Statistically significant:</u> £ SMD: 0.13, 95%CI [0.03;0.24] (p<0.05) *In favour of universal secondary school-based programs aiming to stimulate intra- and interpersonal domains*		
Self-esteem (effect size)	Programs focussing on emotion regulation vs not	Not statistically significant: 0.13 vs 0.37 Φ *B*: -0.24 ££† (p<0.10)		
Self-regulation (effect size)		Not statistically significant: 0.20 vs 0.23 Φ *B*: -0.03 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.12 vs 0.16 Φ *B*: -0.05 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.17 vs 0.21 Φ *B*: -0.03 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.17 vs 0.20 Φ *B*: -0.03 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: -0.04 vs 0.35 Φ *B*: -0.39 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.10 vs 0.10 Φ *B*: -0.01 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: 0.03 vs 0.18 Φ *B*: -0.16 ££† (p<0.10)		
Resilience (effect size)	Programs focussing on assertiveness vs not	Not statistically significant: 0.12 vs 0.04 Φ *B*: 0.08 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.25 vs 0.26 Φ *B*: -0.00 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.07 vs 0.29 Φ *B*: -0.22 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.15 vs 0.13 Φ *B*: 0.02 ££† (p>0.05)		
Internalizing problems (effect size)		<u>Statistically significant:</u> 0.04 vs 0.26 Φ *B*: -0.21 ££ (p<0.05)		

		In favour of programs not focussing on assertiveness		
Social competence (effect size)		Not statistically significant: 0.09 vs 0.23 Φ *B*: -0.14 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.46 vs 0.21 Φ *B*: 0.26 ££† (p>0.05)		
Aggression (effect size)		Statistically significant: -0.05 vs 0.13 Φ *B*: -0.19 ££ (p<0.05) *In favour of programs not focussing on assertiveness*		
Bullying (effect size)		Not statistically significant: 0.23 vs 0.11 Φ *B*: 0.12 ££† (p>0.05)		
Resilience (effect size)	Programs focussing on self-efficacy vs not	Not statistically significant: 0.07 vs 0.09 Φ *B*: -0.02 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.23 vs 0.27 Φ *B*: -0.04 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.13 vs 0.27 Φ *B*: -0.14 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.14 vs 0.13 Φ *B*: 0.02 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.08 vs 0.23 Φ *B*: -0.15 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.16 vs 0.16 Φ *B*: -0.01 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: -0.04 vs 0.11 Φ *B*: -0.15 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: 0.34 vs 0.11 Φ *B*: 0.23 ££† (p>0.05)		
Self-regulation (effect size)	Programs focussing on self-control vs not	Not statistically significant: 0.15 vs 0.23 Φ *B*: -0.08 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.05 vs 0.14 Φ *B*: -0.09 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.18 vs 0.25 Φ *B*: -0.07 ££† (p>0.05)		

Aggression (effect size)		Not statistically significant: 0.24 vs 0.08 Φ *B*: 0.16 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: 0.38 vs 0.11 Φ *B*: 0.27 ££† (p>0.05)		
Resilience (effect size)	Programs focussing on insight-building vs not	Not statistically significant: 0.22 vs 0.04 Φ *B*: 0.18 ££† (p<0.10)		
Self-esteem (effect size)		Not statistically significant: 0.20 vs 0.31 Φ *B*: -0.11 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.29 vs 0.14 Φ *B*: 0.15 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.12 vs 0.15 Φ *B*: -0.03 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.28 vs 0.15 Φ *B*: 0.13 ££† (p>0.05)		
Social competence (effect size)		Statistically significant: 0.24 vs 0.08 Φ *B*: 0.16 ££ (p<0.05) *In favour of programs focussing on insight-building*		
School climate (effect size)		Not statistically significant: 0.09 vs 0.38 Φ *B*: -0.29 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.08 vs 0.11 Φ *B*: -0.04 ££† (p>0.05)		
Bullying (effect size)		Statistically significant: 0.56 vs 0.02 Φ *B*: 0.55 ££ (p<0.01) *In favour of programs focussing on insight-building*		
Resilience (effect size)	Programs focussing on cognitive-coping vs not	Not statistically significant: 0.04 vs 0.15 Φ *B*: -0.11 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.22 vs 0.26 Φ *B*: -0.05 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.13 vs 0.31 Φ *B*: -0.18 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.15 vs 0.12 Φ *B*: 0.03 ££† (p>0.05)		

Internalizing problems (effect size)		Not statistically significant: 0.15 vs 0.23 Φ *B*: -0.09 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.12 vs 0.21 Φ *B*: -0.10 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: -0.01 vs 0.34 Φ *B*: -0.35 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.09 vs 0.11 Φ *B*: -0.02 ££† (p>0.05)		
Bullying (effect size)		<u>Statistically significant:</u> 0.13 vs 0.19 Φ *B*: -0.32 ££ (p<0.05) *In favour of programs not focussing on cognitive-coping*		
Self-esteem (effect size)	Programs focussing on relaxation vs not	Not statistically significant: 0.06 vs 0.27 Φ *B*: -0.21 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.19 vs 0.22 Φ *B*: -0.03 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.10 vs 0.14 Φ *B*: -0.04 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.19 vs 0.20 Φ *B*: -0.01 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.06 vs 0.21 Φ *B*: -0.15 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.18 vs 0.25 Φ *B*: -0.07 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.29 vs 0.08 Φ *B*: 0.21 ££† (p>0.05)		
Resilience (effect size)	Programs focussing on social skills vs not	Not statistically significant: 0.07 vs 0.09 Φ *B*: -0.02 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.31 vs 0.18 Φ *B*: 0.13 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.23 vs 0.16 Φ *B*: 0.07 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.09 vs 0.17 Φ *B*: -0.07 ££†		

		(p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.16 vs 0.23 Φ *B*: -0.07 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.19 vs 0.18 Φ *B*: 0.01 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.13 vs 0.33 Φ *B*: -0.20 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.11 vs 0.09 Φ *B*: 0.02 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: 0.19 vs 0.08 Φ *B*: 0.12 ££† (p>0.05)		
Resilience (effect size)	Programs focussing on problem solving vs not	Not statistically significant: 0.19 vs 0.04 Φ *B*: 0.15 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.34 vs 0.21 Φ *B*: 0.12 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.30 vs 0.17 Φ *B*: 0.13 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.19 vs 0.13 Φ *B*: 0.06 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.20 vs 0.19 Φ *B*: 0.01 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.16 vs 0.16 Φ *B*: 0.00 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.67 vs 0.04 Φ *B*: 0.63 ££† (p<0.10)		
Aggression (effect size)		Not statistically significant: 0.13 vs 0.07 Φ *B*: 0.06 ££† (p>0.05)		
Bullying (effect size)		<u>Statistically significant:</u> 0.30 vs 0.03 Φ *B*: 0.27 ££ (p<0.01) *In favour of programs focussing on problem solving*		
Self-regulation (effect size)	Programs focussing on peer resistance vs not	Not statistically significant: 0.27 vs 0.20 Φ *B*: 0.08 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.13 vs 0.13 Φ		

		B: -0.01 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.21 vs 0.15 Φ *B*: 0.06 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: -0.02 vs 0.30 Φ *B*: -0.32 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: -0.02 vs 0.12 Φ *B*: -0.13 ££† (p>0.05)		
Self-esteem (effect size)	Programs including practice vs without	Not statistically significant: 0.34 vs 0.11 Φ *B*: 0.23 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.24 vs 0.12 Φ *B*: 0.11 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.14 vs 0.13 Φ *B*: 0.01 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.22 vs 0.13 Φ *B*: 0.09 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.19 vs 0.11 Φ *B*: 0.08 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.25 vs 0.22 Φ *B*: 0.03 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.11 vs 0.09 Φ *B*: 0.02 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: -0.01 vs -0.14 Φ *B*: 0.13 ££† (p>0.05)		
Self-esteem (effect size)	Programs including modelling vs without	Not statistically significant: 0.26 vs 0.25 Φ *B*: 0.00 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.08 vs 0.23 Φ *B*: -0.15 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.14 vs 0.13 Φ *B*: 0.01 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.08 vs 0.21 Φ *B*: -0.13 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.16 vs 0.20 Φ *B*: -0.03 ££†		

		(p>0.05)		
School climate (effect size)		Not statistically significant: -0.11 vs 0.31 Φ *B*: -0.42 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.15 vs 0.08 Φ *B*: 0.07 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: 0.10 vs 0.15 Φ *B*: -0.05 ££† (p>0.05)		
Resilience (effect size)	Programs including discussion vs without	Not statistically significant: £ *B*: 0.09 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.21 vs 0.41 £ *B*: -0.20 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.26 vs 0.14 £ *B*: 0.12 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.14 vs 0.13 £ *B*: 0.01 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.21 vs 0.16 £ *B*: 0.05 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.21 vs 0.14 £ *B*: 0.07 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.26 vs 0.02 £ *B*: 0.24 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.06 vs 0.20 £ *B*: -0.13 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: -0.04 vs -0.18 £ *B*: 0.14 ££† (p>0.05)		
Resilience (effect size)	Programs including goal setting vs without	Not statistically significant: 0.19 vs 0.04 £ *B*: 0.15 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.16 vs 0.28 £ *B*: -0.12 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.01 vs 0.25 £ *B*: -0.24 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.12 vs 0.14 £ *B*: -0.01 ££† (p>0.05)		

Social competence (effect size)		Not statistically significant: 0.16 vs 0.16 £ *B*: 0.00 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.02 vs 0.29 £ *B*: -0.27 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.04 vs 0.11 £ *B*: -0.07 ££† (p>0.05)		
Self-esteem (effect size)	Programs including (self-)monitoring vs without	Not statistically significant: 0.44 vs 0.23 £ *B*: 0.21 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.10 vs 0.14 £ *B*: -0.04 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.35 vs 0.16 £ *B*: 0.19 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.13 vs 0.17 £ *B*: -0.04 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: -0.02 vs 0.30 £ *B*: -0.32 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.16 vs 0.09 £ *B*: 0.07 ££† (p>0.05)		
Self-esteem (effect size)	Programs including multimedia vs without	Not statistically significant: 0.16 vs 0.31 £ *B*: -0.15 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.18 vs 0.21 £ *B*: -0.03 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.15 vs 0.12 £ *B*: 0.04 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.19 vs 0.19 £ *B*: -0.00 ££† (p>0.05)		
Social competence (effect size)		<u>Statistically significant:</u> 0.28 vs 0.13 £ *B*: 0.15 ££ (p<0.05) *In favour of programs including multimedia*		
School climate (effect size)		Not statistically significant: 0.05 vs 0.32 £ *B*: -0.27 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.08 vs 0.11 £ *B*: -0.03 ££†		

		(p>0.05)		
Bullying (effect size)		Not statistically significant: 0.16 vs 0.12 £ *B*: 0.05 ££† (p>0.05)		
Self-esteem (effect size)	Programs including homework vs without	Not statistically significant: 0.34 vs 0.24 £ *B*: 0.10 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.17 vs 0.21 £ *B*: -0.05 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.17 vs 0.12 £ *B*: 0.05 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.24 vs 0.18 £ *B*: 0.07 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.33 vs 0.18 £ *B*: 0.15 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.18 vs 0.09 £ *B*: 0.09 ££† (p>0.05)		
Bullying (effect size)		Not statistically significant: 0.37 vs 0.11 £ *B*: 0.26 ££† (p>0.05)		
Resilience (effect size)	Programs including didactic instruction vs without	Not statistically significant: 0.09 vs 0.07 £ *B*: 0.02 ££† (p>0.05)		
Self-esteem (effect size)		Not statistically significant: 0.28 vs 0.23 £ *B*: 0.05 ££† (p>0.05)		
Self-regulation (effect size)		Not statistically significant: 0.17 vs 0.25 £ *B*: -0.07 ££† (p>0.05)		
General wellbeing (effect size)		Not statistically significant: 0.14 vs 0.12 £ *B*: 0.02 ££† (p>0.05)		
Internalizing problems (effect size)		Not statistically significant: 0.23 vs 0.16 £ *B*: 0.07 ££† (p>0.05)		
Social competence (effect size)		Not statistically significant: 0.17 vs 0.15 £ *B*: 0.02 ££† (p>0.05)		
School climate (effect size)		Not statistically significant: 0.36 vs -0.08 £ *B*: 0.44 ££† (p>0.05)		
Aggression (effect size)		Not statistically significant: 0.16 vs 0.03 £ *B*: 0.13 ££† (p>0.05)		

Bullying (effect size)		Not statistically significant: 0.14 vs 0.13 £ *B*: 0.01 ££† (p>0.05)		
Second Step social-emotional learning program				
Antisocial behaviour	The Second Step social-emotional learning program vs no Second Step program	Not statistically significant: £† SMD: 0.22, 95%CI [-0.03;0.47] (p>0.05)	14, £££	Moy, 2018
Prosocial behaviour		Statistically significant: £ SMD: 0.19, 95%CI [0.08;0.31] Ω (p<0.05) *In favour of Second Step social-emotional learning program*	14, £££	
Content knowledge of Second Step lessons		Statistically significant: £ SMD: 1.08, 95%CI [0.55;1.60] Ω (p<0.05) *In favour of Second Step social-emotional learning program*	13, £££	
Antisocial behaviour	Second Step program in pre-kindergarten vs Second Step program in multiple grades	Statistically significant: £ Meta-regression coefficient (±SE): 0.95 ± 0.23 t-value (test for significance): 4.15 (p<0.05) *In favour of Second Step social-emotional learning program in kindergarten*	14, £££	
Anti-(cyber)-bullying programs				
Bullying victimization (dichotomous data)	Traditional anti-bullying program vs no traditional anti-bullying program	Not statistically significant: 1642/3956 vs 1580/3846 RR: 0.98, 95%CI [0.94;1.04] (p=0.54)	2, 3956 vs 3846	Ng, 2020a
Bullying victimization (continuous data)		Statistically significant: £ SMD: -0.18, 95%CI [-0.26;-0.10] (p<0.0001) *In favour of traditional anti-bullying program*	9, 2348 vs 1695	
Bullying perpetration (dichotomous data)		Not statistically significant: 1013/3950 vs 990/3847 RR: 0.98, 95%CI [0.89;1.08] (p=0.68)	2, 3950 vs 3847	
Bullying perpetration (continuous data)		Statistically significant: £ SMD: -0.30, 95%CI [-0.44;-0.15] (p<0.0001) *In favour of traditional anti-bullying program*	9, 2342 vs 1691	
Cyberbullying victimization (continuous data)	Anti-cyberbullying program vs no anti-cyberbullying program	Statistically significant: £ SMD: -0.13, 95%CI [-0.25;-0.02] (p=0.02) *In favour of anti-cyberbullying program*	5, 3632 vs 2787	
Cyberbullying perpetration (continuous data)		Statistically significant: £ SMD: -0.16, 95%CI [-0.29;-0.03] (p=0.01) *In favour of anti-cyberbullying program*	5, 3610 vs 2756	

Bullying victimization	Influence of personnel delivering program: teachers/school staff vs control	Statistically significant: £ SMD: -0.20, 95%CI [-0.32;-0.08] (p=0.002) *In favour of program delivered by teachers/school staff*	5, 1715 vs 1253	
Bullying perpetration		Statistically significant: £ SMD: -0.32, 95%CI [-0.53;-0.10] (p=0.004) *In favour of program delivered by teachers/school staff*	5, 1709 vs 1249	
Cyberbullying victimization		Statistically significant: £ SMD: -0.09, 95%CI [-0.19;-0.00] (p=0.05) *In favour of program delivered by teachers/school staff*	4, 3563 vs 2708	
Cyberbullying perpetration		Statistically significant: £ SMD: -0.11, 95%CI [-0.22;-0.00] (p=0.04) *In favour of program delivered by teachers/school staff*	4, 3541 vs 2677	
Bullying victimization	Influence of personnel delivering program: content expert vs control	Statistically significant: £ SMD: -0.13, 95%CI [-0.25;0.00] (p=0.05) *In favour of program delivered by content expert*	4, 633 vs 442	
Bullying perpetration		Statistically significant: £ SMD: -0.22, 95%CI [-0.35;-0.08] (p=0.002) *In favour of program delivered by content expert*	4, 633 vs 442	
Cyberbullying victimization		Statistically significant: £ SMD: -0.50, 95%CI [-0.83;-0.18] (p=0.003) *In favour of program delivered by content expert*	1, 69 vs 79 §	
Cyberbullying perpetration		Statistically significant: £ SMD: -0.58, 95%CI [-0.91;-0.25] (p=0.0005) *In favour of program delivered by content expert*	1, 69 vs 79 §	
Bullying victimization	Influence of personnel delivering program: teachers/school staff vs content expert	Not statistically significant: £† Chi^2: 0.63 (p=0.43)	9, 2348 vs 1695	
Bullying perpetration		Not statistically significant: £† Chi^2: 0.61 (p=0.43)	9, 2342 vs 1691	
Cyberbullying victimization		Statistically significant: £ Chi^2: 5.54 (p=0.02) *In favour of program delivered by content expert*	5, 3632 vs 2787	
Cyberbullying perpetration		Statistically significant: £ Chi^2: 7.09	5, 3610 vs 2756	

		(p=0.008) *In favour of program delivered by content expert*		
Bullying victimization	Influence of program location: school vs control	<u>Statistically significant:</u> £ SMD: -0.16, 95%CI [-0.31;-0.01] (p=0.03) *In favour of school-based program*	2, 1125 vs 678	
Bullying perpetration		<u>Statistically significant:</u> £ SMD: -0.15, 95%CI [-0.28;-0.03] (p=0.01) *In favour of school-based program*	2, 1124 vs 678	
Cyberbullying victimization		Not statistically significant: £† SMD: -0.10, 95%CI [-0.23;0.02] (p=0.09)	3, 3201 vs 2356	
Cyberbullying perpetration		Not statistically significant: £† SMD: -0.07, 95%CI [-0.17;0.03] (p=0.16)	3, 3179 vs 2330	
Bullying victimization	Influence of program location: classroom vs control	<u>Statistically significant:</u> £ SMD: -0.19, 95%CI [-0.31;-0.07] (p=0.001) *In favour of classroom-based program*	7, 1223 vs 1017	
Bullying perpetration		<u>Statistically significant:</u> £ SMD: -0.36, 95%CI [-0.55;-0.17] (p=0.0003) *In favour of classroom-based program*	7, 1218 vs 1013	
Cyberbullying victimization		Not statistically significant: £† SMD: -0.26, 95%CI [-0.68;0.15] (p=0.21)	2, 431 vs 431	
Cyberbullying perpetration		<u>Statistically significant:</u> £ SMD: -0.38, 95%CI [-0.71;-0.05] (p=0.03) *In favour of classroom-based program*	2, 431 vs 426	
Bullying victimization	Influence of program location: school vs classroom	Not statistically significant: £† Chi^2: 0.10 (p=0.75)	9, 2348 vs 1695	
Bullying perpetration		Not statistically significant: £† Chi^2: 3.14 (p=0.08)	9, 2342 vs 1691	
Cyberbullying victimization		Not statistically significant: £† Chi^2: 0.52 (p=0.47)	5, 3632 vs 2787	
Cyberbullying perpetration		Not statistically significant: £† Chi^2: 3.03 (p=0.08)	5, 3610 vs 2756	
Bullying victimization	Influence of program duration: up to 3 months vs control	<u>Statistically significant:</u> £ SMD: -0.16, 95%CI [-0.31;-0.01] (p=0.03)	1, 366 vs 352 §	

		In favour of program duration of up to 3 months		
Bullying perpetration		<u>Statistically significant:</u> £ SMD: -0.31, 95%CI [-0.46;-0.16] (p<0.0001) *In favour of program duration of up to 3 months*	1, 361 vs 348 §	
Cyberbullying victimization		Not statistically significant: £† SMD: -0.26, 95%CI [-0.68;0.15] (p=0.21)	2, 431 vs 431	
Cyberbullying perpetration		<u>Statistically significant:</u> £ SMD: -0.38, 95%CI [-0.71;-0.05] (p=0.03) *In favour of program duration of up to 3 months*	2, 431 vs 426	
Bullying victimization	Influence of program duration: 3< X >6 months vs control	<u>Statistically significant:</u> £ SMD: -0.24, 95%CI [-0.44;-0.04] (p=0.02) *In favour of program duration of 3< X >6 months*	5, 463 vs 515	
Bullying perpetration		<u>Statistically significant:</u> £ SMD: -0.44, 95%CI [-0.75;-0.13] (p=0.006) *In favour of program duration of 3< X >6 months*	5, 463 vs 515	
Bullying victimization	Influence of program duration: more than 6 months vs control	<u>Statistically significant:</u> £ SMD: -0.15, 95%CI [-0.25;-0.05] (p=0.003) *In favour of program duration of more than 6 months*	3, 1519 vs 828	
Bullying perpetration		<u>Statistically significant:</u> £ SMD: -0.16, 95%CI [-0.25;-0.07] (p=0.0005) *In favour of program duration of more than 6 months*	3, 1518 vs 828	
Cyberbullying victimization		Not statistically significant: £† SMD: -0.10, 95%CI [-0.23;0.02] (p=0.09)	3, 3201 vs 2356	
Cyberbullying perpetration		Not statistically significant: £† SMD: -0.07, 95%CI [-0.17;0.03] (p=0.16)	3, 3179 vs 2330	
Bullying victimization	Influence of program duration: up to 3 months vs 3< X >6 months vs more than 6 months	Not statistically significant: £† Chi^2: 0.65 (p=0.72)	9, 2348 vs 1695	
Bullying perpetration		Not statistically significant: £† Chi^2: 5.10 (p=0.08)	9, 2342 vs 1691	
Cyberbullying victimization	Influence of program duration: up to 3 months vs more than 6 months	Not statistically significant: £† Chi^2: 0.52 (p=0.47)	5, 3632 vs 2787	
Cyberbullying perpetration		Not statistically significant: £† Chi^2: 3.03	5, 3610 vs 2756	

		(p=0.08)		
Bullying victimization	Influence of parental involvement: parental involvement vs control	Statistically significant: £ SMD: -0.22, 95%CI [-0.35;-0.08] (p=0.001) *In favour of parental involvement in the program*	4, 1651 vs 1193	
Bullying perpetration		Statistically significant: £ SMD: -0.34, 95%CI [-0.58;-0.09] (p=0.007) *In favour of parental involvement in the program*	4, 1645 vs 1189	
Cyberbullying victimization		Not statistically significant: £† SMD: -0.12, 95%CI [-0.26;0.03] (p=0.11)	3, 2186 vs 2043	
Cyberbullying perpetration		Not statistically significant: £† SMD: -0.14, 95%CI [-0.31;0.03] (p=0.12)	3, 2164 vs 2012	
Bullying victimization	Influence of parental involvement: no parental involvement vs control	Not statistically significant: £† SMD: -0.11, 95%CI [-0.23;0.00] (p=0.06)	5, 697 vs 502	
Bullying perpetration		Statistically significant: £ SMD: -0.21, 95%CI [-0.33;-0.09] (p=0.0005) *In favour of programs without parental involvement vs no anti-bullying program*	4, 697 vs 502	
Cyberbullying victimization		Not statistically significant: £† SMD: -0.25, 95%CI [-0.69;0.19] (p=0.26)	2, 1446 vs 744	
Cyberbullying perpetration		Not statistically significant: £† SMD: -0.30, 95%CI [-0.80;0.20] (p=0.24)	2, 1446 vs 744	
Bullying victimization	Influence of parental involvement: parental involvement vs no parental involvement	Not statistically significant: £† Chi^2: 1.23 (p=0.27)	9, 2348 vs 1695	
Bullying perpetration		Not statistically significant: £† Chi^2: 0.82 (p=0.36)	9, 2342 vs 1691	
Cyberbullying victimization		Not statistically significant: £† Chi^2: 0.31 (p=0.58)	5, 3632 vs 2787	
Cyberbullying perpetration		Not statistically significant: £† Chi^2: 0.37 (p=0.54)	5, 3610 vs 2756	
Mental Health First Aid programs				
Knowledge about depression at post-intervention	Teen Mental Health First Aid program vs physical first aid training	Statistically significant: 4.01±2.3 vs 3.23±2.2 MD: 0.78, 95%CI [0.59;0.96] (p<0.001) *In favour of Teen Mental Health First Aid program*	1, 542 vs 574	Ng, 2020b (Hart, 2018)

Knowledge about anxiety at post-intervention		<u>Statistically significant:</u> 3.67±2.3 vs 2.81±2.4 MD: 0.93, 95%CI [0.71;1.14] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Recognition of depression (suicidality) at post-intervention		<u>Statistically significant:</u> 90/542 vs 52/574 § OR: 1.97, 95%CI [1.14;3.39] (p=0.02) *In favour of Teen Mental Health First Aid program*		Ng, 2020b (Hart, 2020)
Recognition of depression (suicidality) at 12-month follow-up		Not statistically significant: 61/465 vs 37/429 § OR: 1.50, 95%CI [0.81;2.73] ¥ (p=0.195)	1, 465 vs 429	
Recognition of anxiety (social phobia/anxiety disorder) at post-intervention		<u>Statistically significant:</u> 315/542 vs 285/574 OR: 3.34, 95%CI [1.88;5.94] (p<0.001) *In favour of Teen Mental Health First Aid program*	1, 542 vs 574	Ng, 2020b (Hart, 2018)
Stigma regarding depression: "Weak-not-sick" at post-intervention		<u>Statistically significant:</u> 1.81±1.9 vs 2.11±1.9 MD: -0.20, 95%CI [-0.28;-0.13] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Stigma regarding depression: "Would not tell anyone" at post-intervention		<u>Statistically significant:</u> 2.17±1.2 vs 2.52±1.2 MD: -0.26, 95%CI [-0.38;-0.14] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Stigma regarding depression: "dangerous/unpre-dictable" at post-intervention		<u>Statistically significant:</u> 2.17±1.4 vs 2.46±1.0 MD: -0.19, 95%CI [-0.28;-0.11] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Stigma regarding anxiety: "Weak-not-sick" at post-intervention		<u>Statistically significant:</u> 1.90±2.30 vs 2.17±2.40 MD: -0.20, 95%CI [-0.28;-0.12] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Stigma regarding anxiety: "Would not tell anyone" at post-intervention		<u>Statistically significant:</u> 2.09±1.2 vs 2.46±1.2 MD: -0.30, 95%CI [-0.42;-0.17] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Helpful intentions regarding depression at post-intervention		<u>Statistically significant:</u> 4.61±1,4 vs 3.65±1,4 MD: 0.95, 95%CI [0.78;1.13] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Harmful intentions regarding depression at post-intervention		<u>Statistically significant:</u> 0.77±1.9 vs 1.14±1.9 MD: -0.33, 95%CI [-0.44;-0.21] (p<0.001) *In favour of Teen Mental Health First Aid program*		

Adequate suicide first aid intentions at post-intervention		Statistically significant: 334/542 vs 75/574 OR: 35.40, 95%CI [19.86;63.14] (p<0.001) *In favour of Teen Mental Health First Aid program*		Ng, 2020b (Hart, 2020)
Adequate suicide first aid intentions at 12-month follow-up		Statistically significant: 145/465 vs 34/429 § OR: 9.70, 95%CI [5.21;17.89] (p<0.001) *In favour of Teen Mental Health First Aid program*	1, 465 vs 429	
Avoid talking about suicide at post-intervention		Statistically significant: 113/542 vs 297/574 OR: 0.13, 95%CI [0.09;0.21] (p<0.001) *In favour of Teen Mental Health First Aid program*	1, 542 vs 574	
Avoid talking about suicide at 12-month follow-up		Statistically significant: 154/465 vs 191/429 OR: 0.50, 95%CI [0.30;0.72] (p=0.001) *In favour of Teen Mental Health First Aid program*	1, 465 vs 429	
Helpful intentions regarding anxiety at post-intervention		Statistically significant: 4.18±1.6 vs 3.32±1.7 MD: 0.75, 95%CI [0.57;0.93] (p<0.001) *In favour of Teen Mental Health First Aid program*	1, 542 vs 574	Ng, 2020b (Hart, 2018)
Harmful intentions regarding anxiety at post-intervention		Statistically significant: 0.86±2.8 vs 0.94±2.9 MD: -0.11, 95%CI [-0.23;0.1] (p<0.001) *In favour of Teen Mental Health First Aid program*		
Adolescent Dating Violence (ADV) prevention programs				
Overall ADV perpetration	Adolescent Dating Violence (ADV) prevention program vs no program or waiting list control	Not statistically significant: £† SMD: -0.04, 95%CI [-0.11;0.04] (p>0.05)	2, 2392 vs 2154	Russell, 2021
Emotional ADV perpetration (continuous data)		Statistically significant: £ SMD: -1.13, 95%CI [-2.09;-0.17] (p<0.05) *In favour of ADV prevention program*	3, 846 vs 756	
Emotional ADV perpetration (dichotomous data)		Statistically significant: £ RR: 0.75, 95%CI [0.70;0.80] (p<0.001) *In favour of ADV prevention program*	3, 2639 vs 2789	
Physical ADV perpetration (continuous data)		Not statistically significant: £† SMD: -0.13, 95%CI [-0.45;0.19] (p>0.05)	1, 121 vs 70 §	
Physical ADV perpetration (dichotomous data)		Statistically significant: £† RR: 0.77, 95%CI [0.63;0.94] (p<0.05) *In favour of ADV prevention program*	3, 3566 vs 3502	

Sexual ADV perpetration		<u>Statistically significant:</u> £ SMD: -0.14, 95%CI [-0.26;-0.03] (p<0.05) *In favour of ADV prevention program*	1, 686 vs 653	
Threatening perpetration		Not statistically significant: £† SMD: -0.08, 95%CI [-0.40;0.24] (p>0.05)	1, £££†	
Overall ADV victimization		Not statistically significant: £† SMD: -0.03, 95%CI [-0.17;0.11] (p>0.05)	1, 1384 vs 1156	
Emotional victimization (continuous data)		Not statistically significant: £† SMD: -0.07, 95%CI [-0.17;0.04] (p>0.05)	3, 846 vs 756	
Emotional victimization (dichotomous data)		<u>Statistically significant:</u> £† RR: 0.77, 95%CI [0.73;0.81] (p<0.001) *In favour of ADV prevention program*	3, 2639 vs 2789	
Physical ADV victimization (continuous data)		Not statistically significant: £† SMD: -0.02, 95%CI [-0.34;0.30] (p>0.05)	1, 121 vs 70 §	
Physical ADV victimization (dichotomous data)		<u>Statistically significant:</u> £† RR: 0.79, 95%CI [0.71;0.87] (p<0.001) *In favour of ADV prevention program*	2, 2598 vs 2748	
Sexual ADV victimization		Not statistically significant: £† SMD: -0.05, 95%CI [-0.16;0.06] (p>0.05)	1, 686 vs 653	
Threatening victimization		Not statistically significant: £† SMD: 0.08, 95%CI [-0.23;0.40] (p>0.05)	1, £££†	
Mental health literacy (MHL) programs				
Mental health knowledge: "symptom profile recognition"	"Mental Health for Everyone" program vs classes as usual	<u>Statistically significant:</u> 0.64±0.31 vs 0.31±0.38 MD: 0.33, 95%CI [0.28;0.38] (p<0.00001)* *In favour of "Mental Health for Everyone" program*	1, 399 vs 445	Seedaket, 2020 (Skre, 2013)
Attitudes or stigma: "prejudiced beliefs"		<u>Statistically significant:</u> 1.92±0.76 vs 2.28±0.97 MD: -0.36, 95%CI [-0.48;-0.24] (p<0.00001)* *In favour of "Mental Health for Everyone" program*		
Mental health knowledge ('literacy') at post-intervention	"HeadStrong" program vs classes as usual	<u>Statistically significant:</u> 14.76±3.84 vs 12.07±3.54 MD: 2.69, 95%CI [1.84;3.54] (p<0.00001)* *In favour of "HeadStrong" program*	1, 153 vs 134 §	Seedaket, 2020 (Perry, 2014)

Mental health knowledge ('literacy') at 6-month follow-up		Statistically significant: 14.27±4.65 vs 13.09±3.15 MD: 1.18, 95%CI [0.07;2.29] (p=0.04)* *In favour of "HeadStrong" program*	1, 128 vs 66 §	
Personal stigma towards depression at post-intervention		Statistically significant: 9.80±5.69 vs 11.79±5.68 MD: -1.99, 95%CI [-3.25;-0.73] (p=0.002)*δ *In favour of "HeadStrong" program*	1, 157 vs 155 §	
Personal stigma towards depression at 6-month follow-up		Not statistically significant: 8.61±5.29 vs 10.22±5.88 MD: -1.61, 95%CI [-3.27;0.05] (p=0.06)*δ	1, 137 vs 67 §	
Attitude towards help-seeking at post-intervention		Not statistically significant: 56.79±12.42 vs 56.17±12.50 MD: 0.62, 95%CI [-2.15;3.39] (p=0.66)*	1, 159 vs 153 §	
Attitude towards help-seeking at 6-month follow-up		Not statistically significant: 56.86±12.65 vs 57.13±13.41 MD: -0.27, 95%CI [-4.12;3.58] (p=0.89)*	1, 137 vs 67 §	
Mental health knowledge	"Mental Health and High School Curriculum Guide" vs teaching as usual	Not statistically significant: 8.82±2.41 vs 8.51±2.45 MD: 0.31 95%CI [-0.16;0.78] (p=0.19)*	1, 308 vs 157 §	Seedaket, 2020 (Milin, 2016)
Stigma: attitudes towards mental illness		Not statistically significant: 20.93±3.00 vs 20.70±2.96 MD: 0.23 95%CI [-0.34;0.80] (p=0.43)*		
Depression literacy at 6-week post-intervention	"Adolescent Depression Awareness Program (ADAP)" vs routine health curriculum	Statistically significant: 1538/2975 vs 818/2532 aOR: 3.10, 95%CI [2.0;5.0] (p<0.001) *In favour of "Adolescent Depression Awareness "*	1, 2975 vs 2532	Seedaket, 2020 (Swartz, 2017)
Depression literacy at 4-month follow-up		Statistically significant: 1220/2234 vs 482/1329 aOR: 3.30, 95%CI [2.2;5.0] (p<0.001) *In favour of "Adolescent Depression Awareness "*	1, 2234 vs 1329	
Mental health stigma at 6-week post-intervention		Not statistically significant: 98/2975 vs 101/2532 § aOR: 0.50, 95%CI [0.2;1.2] (p=0.1)	1, 2975 vs 2532	
Mental health stigma at 4-month follow-up		Not statistically significant: 78/2234 vs 53/1329 § aOR: 1.22, 95%CI [0.5;3.0] ¥ (p=0.7)	1, 2234 vs 1329	
Mental health knowledge	Educational program on mental health complemented with contact module vs educational program on mental health alone	Statistically significant: 42.98±5.77 vs 43.28±5.83 Unadjusted GEE: -0.65 ££ (p=0.008) *In favour of educational program on mental health alone*	1, 354 vs 303 §	Seedaket, 2020 (Chisholm, 2016)
Stigma of mental illness		Not statistically significant: 13.81±3.96 vs 13.85±3.83 Unadjusted GEE: -0.09 ££ † (p=0.5)		
Help-seeking		Statistically significant: 5.51±1.67 vs 5.48±1.62		

		Unadjusted GEE: -0.26 ££ (p=0.05) *In favour of educational program on mental health alone*		
Emotional well-being		Statistically significant: 9.15±5.90 8.87±5.87 Unadjusted GEE: 0.10 ££ (p=0.02) *In favour of educational program on mental health alone*		
Resilience		Not statistically significant: 82.50±15.75 vs 83.34±15.47 Unadjusted GEE: 0.19 ££ † (p=0.3)		
Multicomponent Positive Psychology Interventions				
Subjective well-being at post-intervention	Multicomponent Positive Psychology program vs placebo program or waiting list control	Statistically significant: £ SMD: 0.24, 95%CI [0.11;0.38] (p=0.000) *In favour of Multicomponent Positive Psychology program*	6, 2010 vs 1890	Tejada-Gallardo, 2020
Subjective well-being at follow-up		Statistically significant: £ SMD: 0.13, 95%CI [0.03;0.23] (p<0.05) *In favour of Multicomponent Positive Psychology program*	4, 1840 vs 1797	
Psychological well-being at post-intervention		Not statistically significant: £† SMD: 0.25, 95%CI [-0.01;0.51] (p=0.062) Ω	5, 936 vs 832	
Psychological well-being at follow-up		Not statistically significant: £† SMD: 0.44, 95%CI [-0.45;1.13] (p>0.05) Ω	3, £££†	

OR: Odds ratio, aOR: adjusted Odds ratio, CI: confidence interval, MD: mean difference, SMD: standardized mean difference, *B*: meta-regression coefficient, SE: standard error, GEE: generalised equation estimates.
Multivariate analysis of covariance (MANCOVA): $F_{(,)}$: F ratio with degrees of freedom; η^2_p: partial eta-squared
Chi^2: Test for subgroup differences
*Calculations of MD and p-value done by the reviewer using Review Manager software
£ Raw data intervention vs control not available
££ Effect size/CI cannot be calculated or is not available
£££ # studies and/or # participants not available
¤ In order to determine imprecision, calculation of MD and CI was done by the reviewer using Review Manager software
Φ CI of effect size not available (raw data)
Ω the value of the effect size and/or CI and/or p-value is different between table/forest plot and text in the systematic review
δ Our own calculations differ from the results in the individual study (i.e. statistically significant vs not statistically significant or vice versa)
¥ Imprecision (large variability of results)
† Imprecision (lack of data)
§ Imprecision (limited sample size)

Study limitations

Author, Year	Information about 'Study limitations' from the SRs
Aguirre Velasco, 2020	Tools used: Joanna Briggs Institute Critical Appraisal Checklist. Study limitations according to review authors: The majority of the studies were medium quality with moderate risk of bias. It was difficult to identify to what extent the groups were similar at baseline. Few studies included follow-up and the ones that did, had high attrition rates and short follow-up periods (up to 6 months). The randomized controlled

	trials presented difficulties in terms of the blinding of the research team and participants at different stages of the process. Some studies did not use valid and reliable (standardized) instruments for measuring help-seeking. Most of the studies only used self-report measures, increasing the risk of bias of the findings.
de Mooij, 2020	Tools used: Quality Assessment Tool for Quantitative Studies (QATQS). Study limitations according to review authors: According to the systematic review, the quality of most of the included studies (63 studies) was rated moderate to strong (as compared to 14 studies of weak quality). The quality of the study significantly influenced the estimated mean effect of social skills training programs on interpersonal and emotional skills ($p<0.001$): studies of moderate and strong quality yielded smaller effects compared to studies of weak quality.
Mertens, 2020	Tools used: Cochrane Risk of Bias 2.0 tool for Cluster Randomized Trials. Study limitations according to review authors: Most studies randomly assigned participants to the conditions (k = 70). The mean drop-out rate of participants was 12.33% (SD = 10.65). The analysed components were not implemented in isolation, but in the context of an intervention program consisting of multiple components. Interactions among components can affect their effectiveness. Moreover, it remains unclear how the components were implemented, how much time was allotted to certain components and what the quality of implementation of the component was. These aspects could influence the components' effectiveness. Whether or not participants were randomized, drop-out rate, and type of comparison group were not related to effect sizes concerning the intrapersonal domain or the subdomains. Whether or not participants were randomized was related to effect sizes concerning the interpersonal domain; randomized studies yielded stronger effects. Percentage of drop-out was related to effect sizes concerning social competence; studies with lower drop-out rates yielded stronger effects. Whether or not participants were randomized and drop-out rates were also related to effect sizes concerning bullying; randomized studies and higher drop-out rates yielded stronger effects.
Moy, 2018	The tool(s) used for the quality assessment of the included studies was/were not described in the systematic review. Study limitations according to review authors: 11 of the included studies were RCTs and 6 were quasi-experimental studies. The full impact of universal intervention may be difficult to capture with existing instruments that may be focused on identifying clinically significant levels of problem behaviour among a sample of general education students. Furthermore, primary studies on Second Step typically focused on immediate results following participation in the program, whereas program goals ultimately focus on the long-term development of student social competence.
Ng, 2020a	Tools used: Cochrane Risk of Bias 2.0 tool to assess the risk of bias of all included studies. Study limitations according to review authors: the included studies have a high risk of bias. Another limitation of the included studies was the paucity of follow-up assessments in adolescents.
Ng, 2020b	The tool(s) used for the quality assessment of the included studies are/were not described in the systematic review. Study limitations according to review authors: Quality of studies was mostly low with high risk of bias: blinding of participants in comparison group studies was not possible; regarding reporting outcomes, stigmatizing attitudes and confidence were always self-rated and at high risk of bias, especially social desirability for stigmatizing attitudes. Moreover, it is difficult to comment on the sustainability of the program effects since only few studies included a follow-up assessment.
Russell, 2021	Tools used: Cochrane Risk of Bias 2.0 tool to assess the risk of bias of the included studies. Study limitations according to review authors: In terms of risk of bias, all studies included were characterized as low risk to moderate risk. Of the included studies, lack of participant and personnel blinding and incomplete outcome data (due to high attrition rates) were the most common risk categories ranked as "high". Finally, it is possible that the differing follow-up periods of the included studies could have impacted the findings of the meta-analysis.
Seedaket, 2020	Tools used: Jadad scale to assess the quality of reports of randomized clinical trials (Subscales comprised in Jadad scale: randomisation, double-blinding, a description of withdrawals and dropouts). Study limitations according to review authors: Jadad scale scores (scores ranging from 1 (very poor) to 5 (rigorous)) of included studies: Skre, 2016: score 1; Perry, 2014: score 5; Milin, 2016: score 5; Swartz, 2017: score 3 and Chisholm, 2016: score 3

	All included studies used self-report assessments to measure "mental health literacy" outcomes (subjective).
Tejada-Gallardo, 2020	Tools used: Cochrane Risk of Bias 2.0 tool to assess the risk of bias of all included studies. Study limitations according to review authors: Three studies were rated as being at high risk of bias (i.e. low quality) and four studies were rated as having some concerns. The randomization process domain was the most poorly rated due to the non-randomized controlled trials included in the study. Moreover, many of the included studies lacked information regarding the allocation sequence of participants, session attendance and the blinding of the assessor to intervention status. The number of participants in some studies was weakly powered (i.e. less than 50 participants).

Certainty of the body of evidence

1) **Interventions targeting help-seeking for common mental health problems (Aguirre Velasco, 2020)**

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	-1	See table 'Study limitations'
Imprecision	-1	Limited sample sizes/low number of events/lack of data/large variability of the results
Inconsistency	0	
Indirectness	0	
Publication bias	0	
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Low [C]**	

2) **Social skills training programs (de Mooij, 2020)**

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	0	See table 'Study limitations'
Imprecision	-1	Lack of data
Inconsistency	0	
Indirectness	0	
Publication bias	-1	Publication bias assessed using the PET-PEESE method
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Low [C]**	

3) **Programs aiming to stimulate students intra- and interpersonal domains (Mertens, 2020)**

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	-1	See table 'Study limitations'
Imprecision	-1	Lack of data
Inconsistency	0	
Indirectness	0	
Publication bias	0	
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Low [C]**	

4) **Second Step social-emotional learning program (Moy, 2018)**

	Initial grading e.g. High [A]	Downgrading due to

Limitations of study design	-1	See table 'Study limitations'
Imprecision	0	
Inconsistency	0	
Indirectness	0	
Publication bias	0	
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Moderate [B]**	

5) Anti-(cyber)-bullying programs (Ng, 2020a)

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	-1	See table 'Study limitations'
Imprecision	0	
Inconsistency	0	
Indirectness	0	
Publication bias	-1	
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Low [C]**	

6) Mental Health First Aid programs (Ng, 2020b)

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	-1	See table 'Study limitations'
Imprecision	0	
Inconsistency	0	
Indirectness	0	
Publication bias	0	
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Moderate [B]**	

7) Adolescent Dating Violence (ADV) prevention programs (Russell, 2021)

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	-1	See table 'Study limitations'
Imprecision	-1	Lack of data/limited sample sizes
Inconsistency	0	
Indirectness	0	
Publication bias	0	
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Low [C]**	

8) Mental health literacy (MHL) programs (Seedaket, 2020)

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	-1	See table 'Study limitations'
Imprecision	-1	Limited sample sizes/large variability of the results
Inconsistency	0	
Indirectness	0	
Publication bias	0	
		Upgrading due to

Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Low [C]**	

9) Multicomponent Positive Psychology Interventions (Tejada-Gallardo, 2020)

	Initial grading e.g. High [A]	Downgrading due to
Limitations of study design	-1	See table 'Study limitations'
Imprecision	0	
Inconsistency	0	
Indirectness	0	
Publication bias	-1	Publication bias was assessed through funnel plots, the Egger's test, Duval and Tweedie's trim-and-fill procedure, and the fail-safe N: it is likely that missing publications might have affected the results of the present meta-analysis
		Upgrading due to
Large magnitude of effect	0	
Dose-response gradient	0	
Plausible confounding	0	
QUALITY (GRADE)	**Final grading Low [C]**	

Conclusion	**Help-seeking promoting programs** There is limited evidence in favour of classroom-based interventions based on psychoeducation, with a focus on general mental health knowledge or specifically addressing stigma and with the aim of improving help-seeking for common mental health programs. It was shown that the MAKINGtheLINK program, consisting of interactive activities concerning help-seeking to adolescents of 14-15 years old, resulted in a statistically significant increase of help-seeking from formal sources (versus informal sources), compared to a waiting list control (Aguirre Velasco 2020). However, a statistically significant increase in help-seeking behavior at 12 months following the program, could not be demonstrated (Aguirre Velasco 2020). It was also shown that a stigma-based interactive session and video with a case example to high school adolescents resulted in a statistically significant increase of willingness to seek help, compared to a presentation unrelated to mental health (Aguirre Velasco 2020). However, a statistically significant increase of help-seeking behavior could not be demonstrated when only disseminating a student booklet about help-seeking and self-management support, compared to a waiting list control (Aguirre Velasco 2020). Evidence is of low certainty and results cannot be considered precise due to limited sample size, low number of events, lack of data and/or large variability of results. **Social skills training programs** There is limited evidence in favour of classroom-based social skills training programs aimed at teaching or developing children's adaptive social behaviour to improve their success in social interactions. It was shown that social skills training programs overall resulted in a statistically significant increase of interpersonal skills and emotional skills, and a statistically significant decrease of peer relationship problems, internalizing problem behavior and externalizing problem behavior, compared to no social skills training programs (de Mooij 2020). However, depending on the type of social skills program, results might differ:

	• Social-emotional learning programs vs no program: It was shown that such programs resulted in a statistically significant increase of interpersonal skills and emotional skills, and a statistically significant decrease of internalizing problem behaviour. A statistically significant decrease of peer relationship problems and externalizing problem behaviour could not be demonstrated (de Mooij 2020). • Programs targeting bullying behaviour vs no program: It was shown that such programs resulted in a statistically significant increase of interpersonal skills and emotional skills, and a statistically significant decrease of internalizing and externalizing problem behaviour. A statistically significant decrease of peer relationship problems could not be demonstrated (de Mooij 2020). • Programs targeting (social) anxiety vs no program: It was shown that such programs resulted in a statistically significant increase of emotional skills and a statistically significant decrease of internalizing problem behaviour. A statistically significant increase of interpersonal skills and decrease of externalizing problem behaviour could not be demonstrated (de Mooij 2020). • Programs targeting disruptive behaviour vs no program: It was shown that such programs resulted in a statistically significant increase of interpersonal skills and a statistically significant decrease of internalizing and externalizing problem behaviour. A statistically significant increase of emotional skills and decrease of peer relationship problems could not be demonstrated (de Mooij 2020). • Programs targeting resilience and self-esteem vs no program: It was shown that such programs resulted in a statistically significant increase of emotional skills and a statistically significant decrease of internalizing problem behaviour. A statistically significant increase of interpersonal skills and decrease of externalizing problem behaviour could not be demonstrated (de Mooij 2020). • Programs targeting prosocial interactions vs no program: It was shown that such programs resulted in a statistically significant increase of interpersonal skills and a statistically significant decrease of internalizing problem behaviour. A statistically significant increase of emotional skills and decrease of peer relationship problems and externalizing problem behaviour could not be demonstrated (de Mooij 2020). Also, depending on the inclusion of specific training components, results might differ: • Social skills training programs with psychoeducation components vs programs without these components: It was shown that such programs resulted in a statistically significant increase of interpersonal and emotional skills. A statistically significant decrease of peer relationship problems, internalizing and externalizing problem behaviour could not be demonstrated (de Mooij 2020). • Social skills training programs with psychophysical components vs programs without these components: A statistically significant increase of interpersonal and emotional skills and decrease of peer relationship problems, internalizing and externalizing problem behaviour could not be demonstrated (de Mooij 2020). • Social skills training programs with skill-building components vs programs without these components: A statistically significant increase of interpersonal and emotional skills and decrease of peer relationship problems, internalizing and externalizing problem behaviour could not be demonstrated (de Mooij 2020). A statistically significant increase of interpersonal and emotional skills could not be demonstrated in social skills training programs in which there was focus on a specific "booster component", such as class management, rewarding, goal setting, generalization, coaching, or (self-)monitoring, compared to programs without these booster components (de Mooij 2020). A statistically significant increase of interpersonal and emotional skills could not be demonstrated in social skills training programs when there were changes in setting

(indicated vs universal programs), duration (1-9 weeks vs 10-11 weeks vs 12-16 weeks vs 17-26 weeks vs >27 weeks), type of trainer (school personnel vs mental health professional vs non-school personnel), training of the trainer (training vs no training), mode of delivery (computer program vs face-to-face) and age of the participants (primary school age vs secondary school age vs both children and adolescents) (de Mooij 2020).

Evidence is of low certainty and results cannot be considered precise due to lack of data.

Programs aiming to stimulate intra- and interpersonal domains

There is limited evidence in favour of universal secondary school-based programs aiming to stimulate students' intrapersonal (i.e. the ability to manage one's own feelings and emotions) and interpersonal (i.e. the ability of an individual to build and maintain positive relationships with others) domains.

It was shown that secondary school-based programs aiming to stimulate students' intra-and interpersonal domains resulted in a statistically significant improvement of the intrapersonal domain, self-esteem, self-regulation, general wellbeing, internalizing problems, the interpersonal domain, social competence, aggression and bullying, compared to control (usual care, another intervention, no intervention) (Mertens 2020). A statistically significant improvement of resilience and the school climate could not be demonstrated (Mertens 2020).

However, depending on a focus on a specific content component of the program, results might differ:

- Programs with focus on emotion regulation vs programs without: A statistically significant improvement for any of the above listed outcomes could not be demonstrated (Mertens 2020).
- Programs with focus on assertiveness vs programs without: It was shown that such programs resulted in a statistically significant decrease of internalizing problems and aggression. A statistically significant increase of resilience, self-esteem, self-regulation, general wellbeing, social competence, and school climate, and decrease of bullying could not be demonstrated (Mertens 2020).
- Programs with focus on self-efficacy vs programs without: A statistically significant improvement for any of the above listed outcomes could not be demonstrated (Mertens 2020).
- Programs with focus on self-control vs programs without: A statistically significant improvement for any of the above listed outcomes could not be demonstrated (Mertens 2020).
- Programs with focus on insight building vs programs without: It was shown that such programs resulted in a statistically significant increase of social competence and decrease of bullying. A statistically significant increase of resilience, self-esteem, self-regulation, general wellbeing, and school climate, and decrease of internalizing problems and aggression could not be demonstrated (Mertens 2020).
- Programs with focus on cognitive coping vs programs without: It was shown that such programs resulted in a statistically significant decrease of bullying. A statistically significant increase of resilience, self-esteem, self-regulation, general wellbeing, social competence and school climate, and decrease of internalizing problems and aggression could not be demonstrated (Mertens 2020).
- Programs with focus on relaxation vs programs without: A statistically significant improvement for any of the above listed outcomes could not be demonstrated (Mertens 2020).
- Programs with focus on social skills vs programs without: A statistically significant improvement for any of the above listed outcomes could not be demonstrated (Mertens 2020).
- Programs with focus on problem solving vs programs without: It was shown that such programs resulted in a statistically significant decrease of bullying. A statistically significant increase of resilience, self-esteem, self-regulation, general wellbeing, social competence and school climate, and decrease of internalizing problems and aggression could not be demonstrated (Mertens 2020).

- Programs with focus on peer resistance vs programs without: A statistically significant improvement for any of the above listed outcomes could not be demonstrated (Mertens 2020).

In addition, depending on the use of specific instructional components, results might differ:

- Programs including multimedia vs programs without: It was shown that such programs resulted in a statistically significant increase of social competence. A statistically significant increase of self-esteem, self-regulation, general wellbeing, and school climate, and decrease of internalizing problems, bullying and aggression could not be demonstrated (Mertens 2020).
- Programs including practice, modelling, discussion, goal setting, (self-)monitoring, multimedia, homework or didactic instruction as instructional component vs programs without: A statistically significant improvement for any of the above listed outcomes could not be demonstrated (Mertens 2020).

Evidence is of low certainty and results cannot be considered precise due to lack of data.

Second Step social-emotional learning programs

There is in favour of the Second Step social-emotional learning programs, based on a blend of theoretical foundations, including the cognitive-behavioural model, social learning theory, social information processing and verbal self-regulation.

It was shown that social-emotional learning programs resulted in a statistically significant increase of content knowledge of the lessons and prosocial behavior, compared to no such program (Moy 2018). A statistically significant decrease of antisocial behavior could not be demonstrated, although a statistically significant decrease could be shown when comparing programs in kindergartens versus multiple grades (Moy 2018).

Evidence is of moderate certainty.

Anti-(cyber)bullying programs

There is limited evidence in favour of anti-(cyber)bullying programs.

It was shown that traditional anti-bullying programs resulted in a statistically significant decrease of bullying victimization and perpetration (based on 9 studies with continuous data) (Ng 2020a). However, this could not be demonstrated in 2 studies with dichotomous data (Ng 2020a).

It was shown that anti-cyberbullying programs resulted in a statistically significant decrease of bullying victimization and perpetration, compared to no such program (Ng 2020a).

However, depending on the personnel delivering the intervention, the location of intervention, the duration of the intervention or the presence of parental involvement, results might differ:

- Anti-(cyber)bullying programs delivered by teachers/school staff (versus no program): It was shown that such programs resulted in a statistically significant decrease of (cyber)bullying victimization and perpetration (Ng 2020a).
- Anti-(cyber)bullying programs delivered by content experts (versus no program): It was shown that such programs resulted in a statistically significant decrease of (cyber)bullying victimization and perpetration (Ng 2020a).
- Anti-(cyber)bullying programs delivered by teachers/school staff versus content experts: It was shown that programs delivered by teachers/school staff resulted in a statistically significant less decrease of cyberbullying victimization and perpetration compared to programs delivered by content experts. A difference bullying victimization and perpetration could not be demonstrated (Ng 2020a).
- Anti-(cyber)bullying programs delivered in schools (versus no program): It was shown that such programs resulted in a statistically significant decrease

	of bullying victimization and perpetration. A decrease of cyberbullying victimization and perpetration could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs delivered in classrooms (versus no program): It was shown that such programs resulted in a statistically significant decrease of bullying victimization and perpetration. A decrease of cyberbullying victimization and perpetration could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs delivered in schools versus classrooms: A statistically significant difference in (cyber)bullying victimization and perpetration could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs up to 3 months (versus no program): It was shown that such programs resulted in a statistically significant decrease of bullying victimization and (cyber)bullying perpetration. A decrease of cyberbullying victimization could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs between 3 and 6 months (versus no program): It was shown that such programs resulted in a statistically significant decrease of bullying victimization and bullying perpetration (Ng 2020a). • Anti-(cyber)bullying programs of more than 6 months (versus no program): It was shown that such programs resulted in a statistically significant decrease of bullying victimization and bullying perpetration. A statistically significant difference in cyberbullying victimization and perpetration could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs up to 3 months versus between 3 and 6 months versus more than 6 months: A statistically significant difference in (cyber)bullying victimization and perpetration could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs with parental involvement (versus no program): It was shown that such programs resulted in a statistically significant decrease of bullying victimization and bullying perpetration. A statistically significant difference in cyberbullying victimization and perpetration could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs without parental involvement (versus no program): It was shown that such programs resulted in a statistically significant decrease of bullying perpetration. A statistically significant difference in bullying victimization and cyberbullying victimization and perpetration could not be demonstrated (Ng 2020a). • Anti-(cyber)bullying programs with parental involvement versus no parental involvement: A statistically significant difference in (cyber)bullying victimization and perpetration could not be demonstrated (Ng 2020a). Evidence is of low certainty. **Mental health first aid programs** There is evidence in favour of teen mental health first aid programs. It was shown that teen mental health first aid programs resulted in a statistically significant increase of knowledge about depression, knowledge about anxiety, recognition of depression, recognition of anxiety, stigma regarding depression, stigma regarding anxiety, helpful intentions regarding depression, adequate suicide first aid intentions, and helpful intentions regarding anxiety, and a statistically significant decrease of harmful intentions about depression, avoiding talking about suicide and harmful intentions about anxiety, when measured immediately after program implementation, compared to a physical first aid training (Ng 2020b). At 12 month follow-up it was shown that the program resulted in a statistically significant increase of adequate suicide first aid intentions and a decrease of avoiding talking about suicide, compared to a physical first aid training (Ng 2020b). However, at 12 months follow-up, a statistically significant increase of recognition of depression could not be demonstrated (Ng 2020b). Evidence is of moderate certainty. **Adolescent Dating Violence prevention programs**

There is limited evidence in favour of adolescent dating violence prevention programs.

It was shown that adolescent dating violence prevention programs resulted in a statistically significant decrease of emotional dating violence perpetration and sexual dating violence perpetration, compared to no such program or a waiting list control (Russell 2021). However, for the outcomes physical dating violence perpetration, emotional victimization and physical dating violence victimization, the effect depended on the type of data: a statistically significant decrease was found when measured by dichotomous data, but a significant decrease could not be demonstrated with continuous data (Russell 2021). In addition, a statistically significant decrease of overall data violence perpetration, threatening perpetration, overall dating violence victimization, sexual dating violence victimization and threatening victimization could not be demonstrated (Russell 2021).

Evidence is of low certainty and results cannot be considered precise due to limited sample size or lack of data.

Mental health literacy programs

There is limited evidence in favour of mental health literacy programs.

It was shown that the "Mental Health for Everyone program", the "HeadStrong program" and the "Adolescent Depression Awareness Program" resulted in a statistically significant increase of mental health knowledge, compared to classes as usual (Seedaket 2020). For the "Adolescent Depression Awareness Program" this effect was also shown at 4 months follow-up, and for the "HeadStrong program" at 6 months follow-up (Seedaket 2020). In addition, it was shown that an educational program with a contact module resulted in a statistically significant increase of mental health knowledge, compared to the same program without contact module (Seedaket 2020). However, for the "Mental Health and High School Curriculum Guide" a statistically significant increase of mental health knowledge could not be demonstrated (Seedaket 2020).

It was shown that the "Mental Health for Everyone program" and the "HeadStrong program" resulted in a statistically significant decrease of mental health stigma. However, for the "Mental Health and High School Curriculum Guide" and the "Adolescent Depression Awareness Program" a statistically significant decrease of stigma could not be demonstrated. For the "HeadStrong program" the decrease of stigma at 6 months-follow up could not be demonstrated. For an educational program with a contact module a statistically significant decrease in stigma, compared to the same program with contact module, could not be demonstrated (Seedaket 2020).

It was shown that an educational program with a contact module resulted in a statistically significant increase of help-seeking, compared to the same program without contact module (Seedaket 2020). However, for the "Headstrong program" a statistically significant increase of a help-seeking attitude could not be demonstrated (Seedaket 2020).

Finally, it was shown that an educational program with a contact module resulted in a statistically significant increase of emotional well-being and resilience, compared to the same program without contact module (Seedaket 2020).

Evidence is of low certainty and results cannot be considered precise due to limited sample size or large variability of results.

Multicomponent positive psychology interventions

There is limited evidence in favour of multicomponent positive psychology interventions.

It was shown that multicomponent positive psychology interventions resulted in a statistically significant increase of subjective wellbeing immediately following implementation of the program and at follow-up, compared to a placebo program or a waiting list control (Tejada-Gallardo 2020). However, a statistically significant

	increase of psychological wellbeing could not be demonstrated (Tejada-Gallardo 2020). Evidence is of low certainty.
Updatestatus	Update needed
Reference(s)	**Articles** Chisholm K, Patterson P, Torgerson C, Turner E, Jenkinson D, Birchwood M. *Impact of contact on adolescents' mental health literacy and stigma: the SchoolSpace cluster randomised controlled trial.* BMJ Open 2016, 6(2):e009435 Hart LM, Morgan AJ, Rossetto A, Kelly CM, Mackinnon A, Jorm AF. *Helping adolescents to better support their peers with a mental health problem: A cluster-randomised crossover trial of teen Mental Health First Aid.* Australian & New Zealand Journal of Psychiatry 2018, 52(7):638-651 Hart LM, Cropper P, Morgan AJ, Kelly CM, Jorm AF. *teen Mental Health First Aid as a school-based intervention for improving peer support of adolescents at risk of suicide: Outcomes from a cluster randomised crossover trial.* Australian & New Zealand Journal of Psychiatry 2020, 54(4):382-392 Lubman DI, Cheetham A, Jorm AF, Berridge BJ, Wilson C, Blee F, Mckay-Brown L, Allen N, Proimos J. *Australian adolescents' beliefs and help-seeking intentions towards peers experiencing symptoms of depression and alcohol misuse.* BMC Public Health 2017, 17(1):658 Lubman DI, Cheetham A, Sandral E, Wolfe R, Martin C, Blee F, Berridge BJ, Jorm AF, Wilson C, Allen NB, Mckay-Brown L, Proimos J. *Twelve-month outcomes of MAKINGtheLINK: A cluster randomized controlled trial of a school-based program to facilitate help-seeking for substance use and mental health problems.* EClinicalMedicine 2020, 18:100225 Milin R, Kutcher S, Lewis SP, Walker S, Wei Y, Ferrill N, Armstrong MA*. Impact of a Mental Health Curriculum on Knowledge and Stigma Among High School Students: A Randomized Controlled Trial.* J Am Acad Child Adolesc Psychiatry 2016, 55(5):383-391 Perry Y, Petrie K, Buckley H, Cavanagh L, Clarke D, Winslade M, Hadzi-Pavlovic D, Manicavasagar V, Christensen H. *Effects of a classroom-based educational resource on adolescent mental health literacy: a cluster randomized controlled trial.* J Adolesc 2014, 37(7):1143-1151 Saporito JM, Ryan C, Teachman BA. *Reducing stigma toward seeking mental health treatment among adolescents.* Stigma Res Action 2011, 1(2):9-21 Sharpe H, Patalay P, Vostanis P, Belsky J, Humphrey N, Wolpert M. *Use, acceptability and impact of booklets designed to support mental health self-management and help seeking in schools: results of a large randomised controlled trial in England.* Eur Child Adolesc Psychiatry 2017, 26(3):315-324 Skre I, Friborg O, Breivik C, Johnsen LI, Arnesen Y, Wang CE. *A school intervention for mental health literacy in adolescents: effects of a non-randomized cluster controlled trial.* BMC Public Health 2013, 13:873 Swartz K, Musci RJ, Beaudry MB, Heley K, Miller L, Alfes C, Townsend L, Thornicroft G, Wilcox HC. *School-Based Curriculum to Improve Depression Literacy Among US Secondary School Students: A Randomized Effectiveness Trial.* Am J Public Health 2017, 107(12):1970-1976 **Systematic reviews** Aguirre Velasco A, Cruz ISS, Billings J, Jimenez M, Rowe S. *What are the barriers, facilitators and interventions targeting help-seeking behaviours for common mental health problems in adolescents? A systematic review.* BMC Psychiatry 2020, 20(1):293 de Mooij B, Fekkes M, Scholte RHJ, Overbeek G. *Effective Components of Social Skills Training Programs for Children and Adolescents in Nonclinical Samples: A Multilevel Meta-analysis.* Clin Child Fam Psychol Rev. 2020, 23(2):250-264 Mertens E, Deković M, Leijten P, Van Londen M, Reitz E. *Components of School-Based Interventions Stimulating Students' Intrapersonal and Interpersonal Domains: A Meta-analysis.* Clin Child Fam Psychol Rev. 2020, 23(4):605-631 Moy G, Polanin JR, McPherson C, Phan T-V. *International adoption of the Second Step program: Moderating variables in treatment effects.* School Psychology International 2018, 39(4):333-359 Ng ED, Chua JYX, Shorey S. *The Effectiveness of Educational Interventions on Traditional Bullying and Cyberbullying Among Adolescents: A Systematic Review and Meta-Analysis.* Trauma Violence Abuse 2022, 23(1):132-151

	Ng SH, Tan NJH, Luo Y, Goh WS, Ho R, Ho CSH. *A Systematic Review of Youth and Teen Mental Health First Aid: Improving Adolescent Mental Health.* J Adolesc Health 2021, 69(2):199-210 Russell KN, Voith LA, Lee H. *Randomized controlled trials evaluating adolescent dating violence prevention programs with an outcome of reduced perpetration and/or victimization: A meta-analysis.* J Adolesc 2021, 87:6-14 Seedaket S, Turnbull N, Phajan T, Wanchai A. *Improving mental health literacy in adolescents: systematic review of supporting intervention studies.* Trop Med Int Health 2020, 25(9):1055-1064 Tejada-Gallardo C, Blasco-Belled A, Torrelles-Nadal C, Alsinet C. *Effects of School-based Multicomponent Positive Psychology Interventions on Well-being and Distress in Adolescents: A Systematic Review and Meta-analysis.* J Youth Adolesc 2020, 49(10):1943-1960

Appendix 1: definition of the program content components that were analysed (Mertens, 2020)

Content component	Definition
Emotion regulation	Strategies to help youth identify and appropriately express emotions (including aggression).
Assertiveness	Exercises designed to promote the youth's ability to assert his or her needs appropriately with others.
Self-efficacy	Techniques and training to enhance self-confidence and improve self-efficacy.
Self-control	Strategies to help youth interrupt undesired behavioural tendencies (e.g. impulses) and refrain from acting on them.
Insight building	Activities specifically designed to help a youth achieve greater self-understanding and adjust attitudes.
Cognitive coping	Any techniques designed to alter interpretation of events or deal with stressful situations through examination of the youth's reported thoughts (e.g. cognitive restructuring).
Relaxation	Techniques or exercises designed to induce physiological calming.
Social skills	Training youth how to communicate more effectively with others and providing constructive information, training and feedback to improve interpersonal verbal or non-verbal functioning.
Problem solving	Training in the use of techniques, discussions or activities designed to bring about solutions to social, emotional or behavioural problems.
Peer resistance	Techniques or training to learn youth how to resist pressure from peers.