Supplementary Material 1: Full table of included study details, effectiveness and quality ratings for peer learner outcomes

Author, Year	Study Design	Nation	Sample size	Health Area	Outcome & effectiveness	Effectiveness	Quality Score	Quality Rating
Aslan & Sahin, 2007	Non-randomised, pre-post design	Turkey	279	Alcohol, smoking, substance use	↑ Knowledge & attitudes (See paper for a breakdown of results as reported by the author at the level of multiple individual statements)	Effective	2	Medium
Bloor et al. 1999	Non-randomised, intervention vs. control	UK	1300	Alcohol, smoking, substance use	↔ Change in smoking behaviour immediately post intervention (p=0.22) ↑ Smoking abstinence in ex-smokers (p =0.004)	Mixed	1	Low
Botvin et al. 1990	Randomised, peer-led vs. teacher-led	USA	1185	Alcohol, smoking, substance use	Peer intervention, peer intervention + booster ↔,↑ Smoking use (NS), (<.01) ↔,↔ Alcohol use (NS), (NS) ↔,↑ Marijuana use (NS), (<.05) ↑,↑ Knowledge: Tobacco (p < .0001), (p < .0001) ↑,↑ Knowledge: Alcohol (p < .01), (p < .0001) ↔, ↔ Knowledge: Marijuana (NS), (NS) ↔,↑ Attitudes: Tobacco (NS), (p < .01) ↔, ↔ Attitudes: Alcohol (NS), (NS) ↔, ↔ Marijuana (NS), (NS) ↔,↔ Personality: Assertiveness (NS), (NS)	Ineffective	2	Medium

					↔, ↔ Personality: Locus of control (NS), (p < .01) ↔, ↑ Personality: Social anxiety (NS), (NS) ↔, ↔ Personality: Self-esteem (NS), (NS) ↔, ↔ Personality: Smoking influenceability (NS), (NS) ↔, ↔ Personality: General influenceability (NS), (NS)			
Campbell et al. 2008	Cluster randomised controlled trial, intervention vs. control	UK	10,730	Alcohol, smoking, substance use	↑ All students: Smoking prevalence immediately after intervention (*p* =0.031)	Effective	5	High
Demirezen et al. 2020	Non-randomised, pre-post survey	Turkey	663	Alcohol, smoking, substance use	↑ Knowledge and attitude scores (*p* = < .01) ↑ ↔ Incidence of smoking: just once (*p* =<.001), yes (NS) ↔↔ Incidence of alcohol use: just once (NS), yes (NS) ↔↔ Incidence of cannabis use: just once (NS), yes (NS) ↔↔ Incidence of using ecstasy: just once (NS), yes (NS) ↔↔ Incidence of using volatiles: just once (NS), yes (NS) ↑↔ Incidence of using non-prescription drugs: just once (*p* =<.001), yes (NS)	Ineffective	3	Medium
Erhard, 1999	Non-randomised, peer-led vs. adult-led	Israel	2447	Alcohol, smoking, substance use	↑ Evaluation of knowledge: (*p* =< .05) ↑ Evaluation of motivation: (*p* =< .05) ↑	Effective	2	Medium

					Increased curiosity: (p =< .05) ↑ Facilitators as an address: (p =< .05) ↑ Improved relations: (p =< .05) ↑ Satisfaction: (p = < .05)			
LaChause, 2008	Randomised, intervention vs. control	USA	114	Alcohol, smoking, substance use	↑ FAS knowledge (p =.01) ↔ Attitudes toward alcohol use during pregnancy (NS) ↔ Likelihood to use alcohol during pregnancy (NS) ↔ Intention to use alcohol during pregnancy (NS)	Ineffective	1	Low
Lotrean et al. 2010	Randomised, intervention vs. control	Romania	1071	Alcohol, smoking, substance use	↑ Pros of smoking (p = <0.01) ↔ Cons of smoking (NS) ↑ Self-efficacy social (p = <0.05) ↔ Self-efficacy emotional (NS) ↔ Self-efficacy situational (NS) ↑ Intention to smoke (p=<0.05)	Mixed	4	High
Mall & Bhagyalaxmi, 2017	Cluster randomised trial, intervention vs. control	India	802	Alcohol, smoking, substance use	↑ Tobacco consumption at follow-up (p =< 0.01)	Effective	1	Low
Mohammadi et al. 2019	Cluster randomised trial,	Iran	1807	Alcohol, smoking, substance use	↑ Smoking-related knowledge: (p= <0.001) ↔ Attitude toward smoking : (p=0.124)	Effective	4	High

	intervention vs. control				↑ Normative beliefs: (p=<0.001) ↑ Intention to tobacco use: (p = <0.013) ↔			
Murray et al. 1987	Non-randomised, pre-post survey	USA	Study 1: 3154 Study 2: 3820	Alcohol, smoking, substance use	↑ Reduced smoking incidence at 1 year ↑ Smoking intensity levels at 1 year ↑ Peer-led social influence condition	Effective	2	Medium
Perry et al. 1980	Non-randomised, intervention vs. control	USA	707	Alcohol, smoking, substance use	↑ Smoked in past week ↑ Smoked in past month	Effective	4	High
Perry et al. 1989	Randomised, peer-led, teacher-led, control	Multinational (Australia, Chile, Norway Switzerland)	2536	Alcohol, smoking, substance use	↑ Alcohol use (p =<.005) ↑ Knowledge (p =<.005) ↑ Attitudes (p =<.005) ↓ Skills (p =<.08) ↑ Friends' Drinking (p =<.005)	Effective	2	Medium
Severson et al. 1991	Randomised, peer-led vs. teacher-led	USA	1768	Alcohol, smoking, substance use	↑ Smokeless tobacco use in males (p =<0.05) ↑ Smokeless tobacco cessation rate ↔ Cigarette Smoking (NS) ↓ Carbon monoxide level (p =<0.05) ↔ Alcohol use (NS)	Ineffective	4	High

					↔ Marijuana use (NS)			
Weichold & Silbereisen, 2012	Randomised, peer-led, teacher-led, control	Germany	105	Alcohol, smoking, substance use	↓ Alcohol consumption – higher in peer-led condition compared to teacher-led (p = < .05 for T2) than in control (p = < .01 for T2). ↓ Frequency of drinking - higher in peer-led condition compared to teacher-led condition (p = < .001 for T2) ↓ Cigarette consumption – lower in students of the teacher-led condition compared to peer-led condition (p = < .05 for T2 and to those of the control group (p =< .01 for T2). ↓ Resistance skills - lower in the peer-led intervention (p = < .05 for T2) compared to control group (p =< .05 for T2)	Ineffective	2	Medium
Al-Sheyab et al. 2012	Cluster randomised controlled trial, intervention vs. control	USA	240	Asthma	↑ Health-related quality of life (p =<.05) ↑ Self-efficacy to resist smoking (p =<.05) ↑ Knowledge of asthma self-management (p =<.05)	Effective	2	Medium
Al-Sheyab et al. 2016	Cluster, randomised controlled trial, intervention vs. control	USA	486	Asthma	↑ Smoking related knowledge (p = < 0.001) ↑ Level of nicotine dependence (p = < 0.001) ↑ Asthmas control scores (p =< 0.001)	Effective	2	Medium
Gibson et al. 1998	Controlled trial,	Australia	935	Asthma	↑ Mean scores of asthma Knowledge (p =< 0.001)	Effective	3	Medium

	intervention vs. control							
McCallum et al. 2017	Non-randomised, pre – post survey	Australia	203	Asthma	↔ 'Asthma control test score: no change	Ineffective	2	Medium
Shah et al. 2001	Cluster randomised controlled trial, intervention vs. control	Australia	272	Asthma	↑ Quality of life (p =0.01) ↑ Activities domain (p =0.028) ↑ Emotions domain (males only), (p =0.02) ↑ School 6bsenteeism (p =<.05) ↔ Asthma attacks at school (NS)	Effective	3	Medium
Palladino et al. 2016	Non-randomised, intervention vs. control	Italy	Trial 1: 622 Trial 2: 461	Bullying	↑ Victimization: (p =.003) ↑ Bullying – (p =.010) ↑ Cybervictimization – (p =0.002) ↑ Cyberbullying – (p = 0.009)	Effective	4	High
Isik et al. 2013	Non-randomised, cross-sectional survey	Turkey	2829	Disease prevention	↑ Intention to participate in BB rituals (p =<0.001)	Effective	3	Medium
Sadoh et al. 2018	Non-randomised, pre-post survey	Nigeria	1443	Disease prevention	↑ Awareness of cervical cancer ↑ Knowledge of risk factors and cause of cervical cancer	Effective	2	Medium
Acemoglu, 2011	Non-randomised,	Turkey	2930	Disease prevention	↑ Knowledge scores (p =0.001)	Effective	1	Low

	pre-post survey				↑ Attitudes of students towards people infected with hepatitis virus (p =0.002) ↑ Change in behaviours (p =0.000)			
Liu et al. 2015	Randomised, peer-led vs. teacher-led	China	1265	Disease prevention	↑ Pre-post knowledge, attitude & practice ↓ Knowledge, attitude & practice (larger improvement in teacher-led, p =<.01)	Mixed	3	Medium
Ajuwon & Ajuwon, 2019	Non-randomised, pre-post survey	Nigeria	120	Healthy lifestyles	↑ Knowledge of CHI resources ↑ Understanding of roles as peer educators	Effective	2	Medium
Bell et al. 2017	Randomised, intervention vs. control	UK	928	Healthy lifestyles	↑ Consumption of breakfast (p =0.004) ↔ Consumption of at least three portions of fruit a day (p= 0.055) ↔ No statistically significant difference in the reported consumption of other target foods ↑ Counts per minute (CPM) ↑ Sedentary minutes per day ↔ MVPA minutes per day	Mixed	4	High
Bogart et al. 2014	Randomised trial, intervention vs. wait-list control	USA	2997	Healthy lifestyles	↑ Fruit servings (p =<.10) ↔ Vegetable servings: ns ↑ All lunches (p =<.01) ↑ Free/reduced lunch (p =<.001)	Effective	4	High

					↑ Full-price lunch (p =<.01) ↑ Cafeteria attitudes (p =< .05) ↑ Tap water attitudes (p =< .05) ↑ Knowledge about healthy eating/physical activity (p = < .05) ↑ Intentions to drink tap water (p = < .05) ↑ Intentions to drink from refillable bottled (p =< .05) ↑ Tap water consumption (p = < .05) ↔ Refillable bottle use (NS)			
Bogart et al. 2016	Randomised trial, intervention vs. wait-list control	USA	4576	Healthy lifestyles	↑ Students obese at seventh grade (p =<.01) ↔ Students overweight at seventh grade (no intervention effect) ↔ Students healthy weight/ underweight at seventh grade (no intervention effect)	Ineffective	2	Medium
Cohen et al. 1989	Non-randomised, peer-led vs. teacher-led	USA	886	Healthy lifestyles	↔ Blood pressure behavioural capabilities (NS) ↔ Smoking experimentation (NS) ↔ Nutrition behavioural capabilities (NS) ↔ Family discussion (NS)	Ineffective	3	Medium
Cui et al. 2012	Randomised, intervention vs. control	China	682	Healthy lifestyles	↔ Decrease in time spent on sedentary behaviour: 3 months (p = 0.21) ↔	Ineffective	3	Medium

					Decrease in time spent on sedentary behaviour on weekdays: 3 months (p =0.21) ↔ Decrease in time spent on sedentary behaviour on weekends: 3 months (p =0.25) ↔, Time spent on computer: 3 months: (p =0.13) ↔ Time spent on computer on weekdays: 3 months - (p =0.07) ↔ Time spent on computer on weekend: 3 months (p =0.27) ↔ Time spent on Television and DVD: 3 months (p = 0.13) ↔ Time spent on Video games: 3 months (p=0.26) ↔ Time spent on extracurricular reading, writing, drawing and listening to music: 3 months (p = 0.4) ↔ Time spent on passive commuting: 3 months (p =0.22) ↔ Time spent sitting and talking: 3 months (p =0.16)			
Forneris et al 2010	Randomised, intervention vs. wait-list control	USA	523	Healthy lifestyles	↑ Self-Efficacy to Eat Healthy (p =<.05) ↔ Perceived Taste of Low-Fat Foods “not significant” ↑ Fat and Fiber Knowledge (p =<.003) ↔ Fat Food Frequency Score (p = .98) ↔ Fiber Food Frequency Score (p = .89)	Ineffective	1	Low

					↔ Fruit and Vegetable Food Frequency Score (*p* =.76)			
Haleem et al. 2012	Cluster randomised controlled trial, per-led, dentist-led and teacher-led	Pakistan	1657	Healthy lifestyles	↔ Oral health knowledge (NS) ↔ Oral hygiene status (NS) ↑ Oral health behaviour (*p* =<0.05)	Ineffective	5	High
Ishak et al. 2019	Non-randomised, intervention vs. control	Malaysia	76	Healthy lifestyles	↑ Knowledge score at time interval one (*p* = 0.006) ↑ Knowledge score at time interval two *(p* = 0.005)	Effective	4	High
Lytle et al. 2004	Group-randomised trial, intervention vs. control	USA	2883	Healthy lifestyles	↔ Scores for total fruit and veg (*p* =.11) ↔ Scores for energy from total fat (*p* =.90) ↑ Food choice score (*p* =p=.01)	Ineffective	4	High
Ping et al. 2014	Cluster randomised control trial, intervention vs. control	China	645	Healthy lifestyles	↑ Sun safe knowledge (*p* =<.001)	Effective	4	High
Sebire et al. 2018	Cluster randomised controlled trial, intervention vs. control	UK	427	Healthy lifestyles	↔ Mean weekday moderate-to-vigorous physical activity ↔ Mean weekend moderate-to-vigorous physical activity ↑ Mean weekday sedentary (mins) ↔ Mean weekend sedentary (mins)	Ineffective	5	High
Shankar et al. 2020	Non-randomised,	India	755	Healthy lifestyles	↑ Health knowledge (*p* =< 0.001)	Effective	1	Low

	pre-post survey				↑ Health-related attitudes (p =< 0.001)			
Shrewsbury et al. 2020	Non-randomised, pre-post survey	Australia	2056	Healthy lifestyles	↔ ↑ Daily breakfast (behaviour, p =.675, intention, p =<.05) ↑ ↑ Fruit ⩾2 serves/day (behaviour, p=<.001), intention, p=<.001) ↔ ↑ Vegetables⩾5 serves/day (behaviour, p=.156, intention, p=<.01) ↑ <1 cup/day SSB (behaviour, p=<001.) ↔ ↑ Screen-time ⩽2 hours/day (behaviour, p=<.01, intention, p=<.001)	Effective	5	High
Tamiru et al. 2017	Non-randomised, pre-post survey	Ethiopia	992	Healthy lifestyles	↔ Dietary diversity score (p=0.093) ↔ Food variety score (p=0.075 ↑ Animal source food intake (p<0.001)	Ineffective	1	Low
Vangipuram et al. 2016	Randomised controlled trial, peer-led, dentist-led & control	India	450	Healthy lifestyles	↑ Oral health knowledge ↑ Attitude ↑ Oral hygiene practices ↑ Oral health status ↑ Oral health behaviour	Effective	1	Low
Ciao et al. 2015	Non-randomised, intervention	USA	50	Mental health	↑ Internalization (p =<0.001) ↑	Effective	3	Medium

	vs. wait-list control				Body dissatisfaction (p =<0.005) ↑ Eating pathology (p =<0.012)			
Eisenstein et al. 2019	Non-randomised, pre-post survey	UK	950	Mental health	↑ Understanding of key terms (p =<0.001) ↑ Key skills (p =<0.001) ↑ Confidence to talk about mental health (p =0.03) ↔ Knowledge of information and resources (p = 0.07) ↑ Readiness to support others (p <0.001) ↑ Emotional difficulties (p = 0.005) ↔ Behavioural difficulties (p = 0.10) ↑ School climate (p <0.001)	Effective	2	Medium
Ellis 2009	Non-randomised, pre-post survey, intervention vs. control	Australia	483	Mental health	↑ Pro bully attitude (p = < .05) ↔ Pro victim attitude (NS) ↑ Honesty/trustworthiness: (p =< .05) ↔ Self Confidence (NS) ↔ Global self esteem (NS) ↔ Self-efficacy (NS) ↔ Same sex relations (NS) ↑ Opposite-sex relations (p = < .01)	Ineffective	4	High

					↔ Cooperative teamwork (NS)			
Kaveh et al. 2014	Cluster randomised controlled trial, intervention vs. control	Iran	223	Mental health	↑ Self-esteem general scale (*p*=<0.001) ↑ Academic scale (*p*=<0.001) ↑ Social scale (*p*=<0.001) ↔ Family scale (*p*=0.529) ↑ Body scale (*p*=<0.001) ↑ Total scale (*p*=<0.001)	Effective	3	Medium
Parikh et al. 2018	Non-randomised, pre-post survey	USA	878	Mental health	↑ Confidence in identifying peers with depression (*p*<.001) ↑ Confidence in referring peers with depression (*p* =<.001) ↔ Knowing whether to tell someone if a friend informs them they are suicidal (NS) ↑ Feelings about visiting school social worker/psychologist: (*p* = <.001) ↑ Depression knowledge questions: (*p* = <.001) ↑ Knowledge of depression symptoms: (*p* = <.001) ↔ Attitudes to a new student with depression: ns ↑ Comfort speaking to other students about mental health (*p* =< 0.001)	Effective	1	Low

					↑ Willingness to confide in authority figures (p = <0.01)			
Aten et al. 2002	Non-randomised controlled trial, intervention vs. control	USA	1352	Sex education	Sexual activity at follow-up for students who were abstinent at pretest: ↔ Females: (p = 0.07) ↑ Males: (p = 0.04)	Mixed	3	Medium
Borgia et al. 2005	Randomised controlled trial, peer-led vs. teacher-led	Italy	1295	Sex education	↑↑↑↑ Pre-post: Skills, knowledge, attitudes, risk perception (p=<0.05) ↔↑↔↔ Between arms: Skills, knowledge, attitudes, risk perception (p=<0.05)	Effective	4	High
Caron et al. 2004	Non-randomised, intervention vs. control	Canada	1004	Sex education	↑ Postponing sexual intercourse: Intention (p=0.001) ↑ Postponing sexual intercourse: Attitude direct (p= < 0.001) ↑ Postponing sexual intercourse: Attitude indirect (p= < 0.001) ↑ Postponing sexual intercourse: Perceived self-efficacy (p= < 0.001) ↑ Postponing sexual intercourse: Role beliefs (p= < 0.001) ↑ Postponing sexual intercourse: Anticipated regret (p = < 0.001) ↑ Condom use: intention (p = < 0.001)	Effective	3	Medium

					↑ Condom use: roles beliefs (p = < 0.001)			
Fisher et al. 2002	Quasi-experimental controlled trial, classroom-based, peer-based and combined.	USA	1577	Sex education	**Information** ↔ sexually inexperienced (NS) ↑ sexually experienced (p= < .01) **Attitudes** ↔ sexually inexperienced (NS) ↑ sexually experienced: (p= <.01) **Norms** ↔ sexually inexperienced (NS) ↔ sexually experienced (NS) Intentions ↔ sexually inexperienced (NS) ↔ sexually experienced (NS) **Behavioural skills** ↔ sexually inexperienced (NS) ↔ sexually experienced (NS)	Ineffective	1	Low
Huang et al. 2008	Randomised, peer-led vs. teacher-led	China	3068	Sex education	↑ HIV-related knowledge (p =<0.001) ↑ Attitudes (p =<0.001) ↑ Behaviour intention (p =<0.05)	Effective	3	Medium

Jennings et al. 2014	Non-randomised intervention vs. control	USA	157	Sex education	↓ Knowledge (p=0.10) ↓ Parental communication (p=0.46) ↓ Self-efficacy (p=0.06) ↓ Sexual health information (p=0.06) ↑ Opportunity to practice skills (p=0.01) ↑ Intentions (p=0.04)	Ineffective	4	High
Mahat et al. 2010	Non-randomised, pre-post survey	USA	106	Sex education	↑ Mean score of total HIV/AIDS knowledge (p =< .001) ↑ Mean of the self-efficacy for limiting HIV risk behaviour score (p = < .001)	Effective	2	Medium
Mason-Jones, Mathews & Fisher 2011	Non-randomised, intervention vs. control	South Africa	2339	Sex education	↑ Not had sex (p=0.002) ↔ First age sex (p=0.84) ↔ Used a condom at last sex (p=0.32) ↔ Goal orientation (p=0.20) ↔ Decision-making (p=0.58) ↔ Future orientation score (p=0.32)	Ineffective	3	High
Mellanby et al. 2001	Non-randomised, peer-led vs. adult-led	UK	1320	Sex education	↓ Knowledge of STDs (greater increase in adult-led, p = < 0.001) ↔ Assertiveness score (p =0.32) ↑ Attitude scores (p =< 0.001)	Ineffective	3	Medium

					↓ Involvement in sessions (greater in adult-led, p = 0.016) ↓ Feeling embarrassed (greater in peer-led, p = <0.001) ↔ Talking to others about the sessions (p=0.12)			
Menna et al. 2015	Non-randomised, intervention vs. control	Ethiopia	560	Sex education	↑ Knowledge of HIV (p=0.004) ↔ Ever initiated sexual intercourse (p=0.72) ↔ Ever tested for HIV (0.53) ↑ Limiting sexual parter only to one in last 12 months (p=*0*.21) ↔ Consistent use of condom in last 12 months (p=0.24) ↑ Willingness to go for HIV counselling and testing (p=0.01)	Mixed	3	Medium
Merakou & Kourea-Kremastinou 2006	Non-randomised, intervention vs. control	Greece	702	Sex education	↑ Increased personal responsibility ↑ Safer behaviour in sexual practice ↔ Knowledge	Effective	1	Low
Michielsen et al. 2012	Non-randomised controlled trial, intervention vs. control	Rwanda	1400	Sex education	↔ Sexual risk behaviour (NS) ↔ Increased knowledge (NS) ↔ Perceived severity (NS) ↔ Perceived susceptibility (NS)	Ineffective	3	Medium

					↑ Stigma (p=0.05)			
Ozcebe, Akin & Aslan, 2004	Non-randomised, pre-post survey	Turkey	369	Sex education	↑ HIV/AIDS knowledge scores (p =<.001)	Effective	4	High
Parwej, 2005	Randomised, peer-led vs. nurse-led	India	273	Sex education	↑ Reproductive health domains ↑ Reproductive health knowledge (*See full paper for full list of individual statements*)	Effective	3	Medium
Rotz et al. 2018	Matched comparison group design, intervention vs. comparison	USA	1522	Sex education	↔ Rates of sexual activity (p=.83) ↔ Rates of unprotected sex (p=.93) ↑ Exposure to information about sexual health topics (p=<.01) ↑ Knowledge of preventing pregnancy and transmission of STIs (p=<.01)	Mixed	3	Medium
Siegel et al. 1998	Non-randomised, intervention vs. control	USA	3635	Sex education	**Postintervention Knowledge Scores:** ↑ Middle School: Boys (p =<.001) ↑ Middle School: Girls (p =<.001) ↑ High School Boys (p =<.001) ↑ High School Girls (p =<.001) **Self-efficacy Scores:** ↔ Middle School: Boys (NS) ↔ Middle School: Girls (NS)	Effective	3	Medium

					↑ High School Boys (*p* =<.001) ↑ High School Girls (*p* =<.001) **Safe Behaviour Intention:** ↑ Middle School: Boys (*p* =<.001) ↑ Middle School: Girls (*p* =<.001) ↑ High School Boys (*p* =<.001) ↑ High School Girls (*p* =<.001)			
Siegel et al. 2001	Non-randomised, intervention vs. control	USA	4001	Sex education	**Postintervention Knowledge Scores:** ↑ Middle School: Boys (*p* =<.001) ↑ Middle School: Girls (*p* =<.001) ↑ High School Boys (*p* =<.001) ↑ High School Girls (*p* =<.001) **Self-efficacy Scores:** ↔ Middle School: Boys (NS) ↑ Middle School: Girls (*p* =<.001) ↔ High School Boys (NS) ↑ High School Girls (*p* =<.001) **Safe Behaviour Intention:** ↑ Middle School: Boys (*p* =<.001)	Mixed	3	Medium

					↔ Middle School: Girls (NS) ↔ High School Boys (NS) ↔ High School Girls (NS)			
Stephenson et al. 2004	Cluster-randomised trial, intervention vs. control (teacher-led)	UK	9508	Sex education	↑ Girls: Unprotected sex before age 16 (p=0.0008) ↔ Boys: Unprotected sex before age 16 (p=0·35) ↑ Girls: Knowledge of methods to prevent STIs (p=0.002) ↔ Boys: Knowledge of methods to prevent STIs (NS)	Mixed	3	Medium
Stephenson et al. 2008	Cluster-randomised trial, intervention vs. control (teacher-led)	UK	9000	Sex education	↔ Proportion of girls with one or more abortions by age 20 y (NS) ↔ Proportion of girls who had had sex by age 18.0 y (NS) ↔ Proportion of girls who had one or more live births by age 20.5 (NS)	Ineffective	3	Medium
Timol et al. 2016	Randomised, intervention vs. control	South Africa	2904	Sex education	↑ Sensation seeking (p =<.05) ↑ Self-efficacy in sexual relations (p =<.05) ↑ Future orientation (p =<.05) ↑ HIV transmission knowledge (p = <.01) ↑ HIV prevention knowledge (p =<.01) ↑ Health relationships (p = <.05)	Effective	3	Medium

Visser, 2007	Non-randomised, intervention vs. control	South Africa	4086	Sex education	↔ Psychological wellbeing (NS) ↔ Personal control (NS) ↔ School climate (NS)	Ineffective	1	Low

↑ = *improvement;* ↓ = *negative effect for peer education;* ↔ *no effect*

NS = no significant change where p-value is not reported

Effective = ≥60% outcomes with a positive finding; Mixed = 41-59% outcomes with a positive finding; Ineffective = ≥40% outcomes with a positive finding

High quality = 4-5; Medium quality = 2-3, Low quality = 0-1 (using the Mixed Methods Appraisal Tool, MMAT)

Supplementary Material 2: Full table of included study details, effectiveness and quality ratings for peer educator outcomes

Author, Year	Study Design	Nation	Sample Size	Health area	Outcome and effectiveness	Effectiveness	Quality Score	Quality Rating	Standalone paper?*
Audrey, Holliday & Campbell, 2006	Cluster randomised controlled trial, intervention vs. control	UK	10,730	Alcohol, smoking, substance use	Smoking prevalence at 1-year follow-up ↑ 'High risk group' (p = 0.045) ↔ Never smoked (p = 0.20) ↔ Regular (weekly) smokers (p= 0.68)	Ineffective	5	High	Yes
Zambuto et al. 2019	Non-randomised, pre-post survey	Italy	524	Bullying	↑ ↔ Victimization (p= .015 males, p= .093 females) ↔ ↔ Bullying (p= .068 males, p= .854 females) ↑↔ Prosocial behaviour (p= .002 males, p= .455 females) ↑ ↔ Social support (p= .036 males, p= .627 females) ↔↑ Defending Behaviour(p= .925 males, p= .015 females)	Ineffective	3	Medium	Yes
Acemoglu, 2011	Non-randomised, pre-post survey	Turkey	2930	Disease prevention	↑ Knowledge scores (p=0.000)	Effective	3	Medium	No
Foley et al. 2017	Non-randomised, pre-post survey	Australia	519	Healthy lifestyles	↑ Eating ≥2 serves fruit/day (p =<0.01) ↑ Eating ≥5 serves vegetables/day	Mixed	4	High	Yes

					(p = <0.01) ↔ Breakfast everyday (p = 0.053) ↑ Drinking <1 cup/day of Sugar sweetened beverages (p = <0.01) ↔ Screen Time ≤ 2 h/day (p = 0.592) ↔ Moderate-to-vigorous physical exercise ≥ 60 min/day (p =< 0.076)				
Eisenstein et al. 2019	Non-randomised, pre-post survey	UK	950	Mental health	↑ Understanding of key terms (p = 0.04) ↑ Key skills (p = <0.001) ↔ Confidence to talk about mental health (p =0.08) ↑ Knowledge of information and resources (p = 0.02) ↔ Readiness to support others (p =0.8) ↔ Emotional difficulties (p = 0.24) ↔ Behavioural difficulties (p =0.65) ↔ School climate (p =0.5)	Ineffective	2	Medium	No
Parikh et al. 2018	Non-randomised, pre-post survey	India	878	Mental health	↑ Confidence in identifying peers with depression (p = 0.036) ↑ Confidence in referring peers with depression (p= 0.023) ↓	Ineffective	1	Low	No

					Knowing whether to tell someone if a friend informs them they are suicidal (NS) ↓ Feelings about visiting school social worker/psychologist (NS) ↓ Depression knowledge questions (NS) ↓ Knowledge of depression symptoms (*p* = 0.07) ↑ Attitudes to a new student with depression (NS) ↓ Comfort speaking to other students about mental health (*p* = < 0.001) ↑ Willingness to confide in authority figures: peer educators: (NS)				
Wyman et al. 2010	Randomised, intervention vs. wait-list control	USA	4128	Mental health	↑ Help for suicidal peers (*p*=<.001) ↑ Rejects codes of silence (*p*=.002) ↑ Maladaptive coping (*p*=.013) ↑ Help-seeking from adults (*p*=<.001) ↑ Sources of Strength coping (*p*=.002) ↑ School engagement (*p*=.043) ↑ Trusted adults (*p*=<.001) ↑	Effective	4	High	Yes

					Referred distressed peers (p=.08) ↔ Support to peers (p=.015)				
Caron et al. 2004	Non-randomised, intervention vs. control	Canada	1004	Sex education	↑ Postponing sexual intercourse: Attitude direct (p =< 0.001) ↑ Postponing sexual intercourse: Attitude indirect (p =< 0.001) ↑ Postponing sexual intercourse: Perceived behavioural control direct (p = < 0.001) ↑ Postponing sexual intercourse: personal normative beliefs (p =0.01) ↑ Postponing sexual intercourse: Role beliefs (p = < 0.001) ↑ Postponing sexual intercourse: Perceived self-efficacy (p = < 0.001) ↑ Postponing sexual intercourse: Anticipated regret (p =< 0.001) ↑ Condom use: behavioural control indirect (p =< 0.001) ↑ Condom use: perceived self-efficacy (p =< 0.001) ↑ Condom use: attitude direct measures (p =0.05) ↑ Condom use: attitude indirect measure	Effective	3	Medium	No

					(p =0.05) ↑ Condom use: personal normative beliefs (p =0.01) ↑ Condom use: anticipated regret (p =0.03) ↑ Condom use: perceived behavioural control direct measure (p =0.01) ↑ Condom use: role beliefs (p =0.05) ↑ Condom use: intention (p =0.02)				
Ebreo et al. 2002	Randomised, pre-post survey	USA	845	Sex education	↑ Had talked to their parents about the use of alcohol (p = < .05) ↓ Report having unwanted sex because they were intoxicated after the intervention (p =< .05) intervention inferior ↔ Individual Difference Variables (NS) ↔ Knowledge (NS) ↔ Parental Communication (NS) ↔ Peer Norms (NS) ↔ Self-efficacy (NS) ↔ Intentions to Have Sex (NS) ↔ Self-reported behaviours (NS) ↔	Ineffective	1	Low	Yes

					Student Course Evaluations (NS)				
Jennings et al. 2014	Non-randomised intervention vs. control	USA	160	Sex education	↔ Knowledge (p =0.10) ↔ Parental communication (p=0.46) ↔ Self-efficacy (p=0.06) ↔ Sexual health information (p=0.06) ↑ Opportunity to practice skills (p=0.01) ↑ Intentions (p=0.04)	Ineffective	4	High	No
Strange, Forrest & Oakley, 2002	Randomised controlled trial, intervention vs. control	UK	331	Sex education	↔ Knowledge; condoms (p = 0.109) ↔ Knowledge; the Pill (p =0.089) ↑ Knowledge; femidom (p =0.000) ↑ Knowledge; emergency contraception – (p =0.026) ↑ Knowledge; cap (p =0.000) ↑ Knowledge; coil (p =0.000) ↔ Attitude; sex before marriage (p = 0.336) ↑ Attitude; people having sex with someone the first time they meet them: (p =0.002) ↑ Attitude; two men having sex with each other (p =0.01) ↑	Mixed	2	Medium	Yes

	Attitude; two women having sex with each other (p =0.003) ↔ Attitude; abortion (p =0.7815) ↔ Attitude; using contraception (p =0.2414)				

↑ = improvement; ↓ = negative effect for peer education; ↔ no effect

NS = no significant change where p-value is not reported

Effective = ≥60% outcomes with a positive finding; Mixed = 41-59% outcomes with a positive finding; Ineffective = ≥40% outcomes with a positive finding

High quality = 4-5; Medium quality = 2-3, Low quality = 0-1 (using the Mixed Methods Appraisal Tool, MMAT)

**Standalone paper: a paper that reported peer educator outcomes only rather than peer learner and peer educator outcomes*

Supplementary Material 3: Example Search Strategy (Medline)

Search terms pertaining to schools and school students	Search terms pertaining to peer education	Search terms pertaining to health-related interventions, health-related outcome domains
(MH "Schools") OR TI ("Elementary School*" OR "Primary School*" OR "Junior School*" OR "Junior High School" OR "Secondary School*" OR "Middle School*" OR "Sixth Form*" OR "High School*" OR "Primary education" OR "Secondary education" OR "Special Education*" OR "School-based" OR "school based") OR AB ("Elementary School*" OR "Primary School*" OR "Junior School*" OR "Junior High School" OR "Secondary School*" OR "Middle School*" OR "Sixth Form*" OR "High School*" OR "Primary education" OR "Secondary education" OR "Special Education*" OR "School-based" OR "school based")	TI ("peer educat*" OR "peer-led" OR "Peer led" OR "Student-led" OR "Student led" OR "Peer teach*" OR "Peer Intervention* OR "Cross-age" or "Cross age" or "Peer to Peer" or "Peer-to-Peer" OR "Peer-delivered" or "Peer Delivered") OR AB ("peer educat*" OR "peer-led" OR "PEER LED" OR "STUDENT-LED" OR "STUDENT LED" OR "PEER TEACH*" OR "Peer Intervention* OR "Cross-age" or "Cross age" or "Peer to Peer" or "Peer-to-Peer" OR "Peer-delivered" or "Peer Delivered")	MH ("Health Education+" OR "Health Literacy+" OR “Health promotion” OR "School Health Services" OR "Preventive Health Services" OR “School Mental Health Services" OR "Anxiety" OR "Self Concept" OR "Depression” OR “Emotional Intelligence" OR “Psychological Distress” OR “Adaptation, Psychological" OR "Sex Education" OR "Sexual Behavior" OR "Safe Sex" OR "Tobacco Use" OR "Underage Drinking" OR "Alcoholism" OR "Substance-Related Disorders" OR "Substance Abuse, Intravenous” OR "Risk Reduction Behavior" OR "Impulsive Behavior" OR "Risk-Taking” OR “Health Risk Behaviors” OR "Health Behavior") OR TI (“Health Promotion" OR "Health Literacy" OR "health education" OR "School Based Intervention*" OR "Group Intervention*" OR “Smoking” OR “Drug abuse” OR “Substance abuse” OR “Drug education” OR "alcohol*" OR “Mental health” OR “Esteem” OR “Confidence” OR “Emotional” OR “behaviour*” OR “Depression” or “anxiety” OR “Sex education" OR "Safe sex" OR “RISK TAKING” OR “HEALTH RISK*”) OR AB ("Health Promotion" OR "Health Literacy" OR "health education" OR "School Based Intervention*" OR "Group Intervention*" OR “Smoking” OR “Drug abuse” OR “Substance abuse” OR “Drug education” OR "alcohol*" OR “Mental health” OR “Esteem” OR “Confidence” OR “Emotional” OR “behaviour*” OR “Depression” or “anxiety” OR “Sex education" OR "Safe sex" OR “RISK TAKING” OR “HEALTH RISK*”)