



## Public Health Research

Volume 12 • Issue 3 • February 2024

ISSN 2050-4381

# School-based interventions TO Prevent Dating and Relationship Violence and Gender-Based Violence: STOP-DRV-GBV systematic review

*GJ Melendez-Torres, Noreen Orr, Caroline Farmer, Naomi Shaw, Annah Chollet,  
Andrew J Rizzo, Fraizer Kiff, Emma Rigby, Ann Hagell, Sidnei R Priolo Filho, Bruce Taylor,  
Honor Young, Chris Bonell and Vashti Berry*





# School-based interventions TO Prevent Dating and Relationship Violence and Gender-Based Violence: STOP-DRV-GBV systematic review

GJ Melendez-Torres<sup>1\*</sup>, Noreen Orr<sup>1</sup>, Caroline Farmer<sup>1</sup>, Naomi Shaw<sup>1</sup>, Annah Chollet<sup>2</sup>, Andrew J Rizzo<sup>3</sup>, Fraizer Kiff<sup>1</sup>, Emma Rigby<sup>4</sup>, Ann Hagell<sup>4</sup>, Sidnei R Priolo Filho<sup>5</sup>, Bruce Taylor<sup>6</sup>, Honor Young<sup>7</sup>, Chris Bonell<sup>8</sup> and Vashti Berry<sup>1</sup>

<sup>1</sup>Faculty of Health and Life Sciences, University of Exeter, Exeter, UK

<sup>2</sup>Department of Social Policy and Intervention, University of Oxford, Oxford, UK

<sup>3</sup>College of Health and Human Performance, University of Florida, Gainesville, FL, USA

<sup>4</sup>Association for Young People's Health, London, UK

<sup>5</sup>Universidade Tuiuti do Paraná, Curitiba, Paraná, Brazil

<sup>6</sup>National Opinion Research Center, University of Chicago, Chicago, IL, USA

<sup>7</sup>Centre for Development, Evaluation, Complexity and Implementation in Public Health Improvement (DECIPHer), Cardiff University, Cardiff, UK

<sup>8</sup>Department of Public Health, Environments and Society, London School of Hygiene and Tropical Medicine, London, UK

\*Corresponding author

## Disclosure of interests

**Full disclosure of interests:** Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/KTWR6997>.

**Primary conflicts of interest:** GJ Melendez-Torres reports part-support from NIHR PenARC; grant and salary support from several NIHR programmes (PHR, HS&DR, HTA, PRP, ESP); and membership of NIHR PGfAR sub-committee, and was co-investigator of one of the studies in this review (NIHR PHR 15/03/09). Ann Hagell reports a role as Trustee and Chair of the charity Habitats and Heritage. Honor Young was co-investigator of one of the studies in this review (NIHR PHR 15/03/09). Sidnei R Priolo Filho (Filho 2017) and Bruce Taylor (Taylor 2010, 2011, 2015) were lead authors of several studies included in this review. Chris Bonell reports membership of the NIHR PHR Research Funding Board (2013–19), grants from the NIHR PHR programme, and a current NIHR Senior Investigator Award; and was chief investigator of one of the studies in this review (NIHR PHR 15/03/09).



Published February 2024  
DOI: 10.3310/KTWR6997

This report should be referenced as follows:

Melendez-Torres GJ, Orr N, Farmer C, Shaw N, Chollet A, Rizzo AJ, *et al.* School-based interventions TO Prevent Dating and Relationship Violence and Gender-Based Violence: STOP-DRV-GBV systematic review. *Public Health Res* 2024;**12**(3). <https://doi.org/10.3310/KTWR6997>



# Public Health Research

ISSN 2050-4381 (Print)

ISSN 2050-439X (Online)

A list of Journals Library editors can be found on the [NIHR Journals Library website](#)

*Public Health Research* (PHR) was launched in 2013 and is indexed by Europe PMC, NCBI Bookshelf, DOAJ, INAHTA, Ulrichsweb™ (ProQuest LLC, Ann Arbor, MI, USA) and MEDLINE.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)

The full PHR archive is freely available to view online at [www.journalslibrary.nihr.ac.uk/phr](http://www.journalslibrary.nihr.ac.uk/phr).

## Criteria for inclusion in the *Public Health Research* journal

Manuscripts are published in *Public Health Research* (PHR) if (1) they have resulted from work for the PHR programme and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Public Health Research* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## PHR programme

The Public Health Research (PHR) programme, part of the National Institute for Health and Care Research (NIHR), is the leading UK funder of public health research, evaluating public health interventions, providing new knowledge on the benefits, costs, acceptability and wider impacts of non-NHS interventions intended to improve the health of the public and reduce inequalities in health. The scope of the programme is multi-disciplinary and broad, covering a range of interventions that improve public health.

For more information about the PHR programme please visit the website: <https://www.nihr.ac.uk/explore-nihr/funding-programmes/public-health-research.htm>

## This manuscript

The research reported in this issue of the journal was funded by the PHR programme as project number NIHR130144. The contractual start date was in July 2020. The final report began editorial review in May 2022 and was accepted for publication in April 2023. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PHR editors and production house have tried to ensure the accuracy of the authors' manuscript and would like to thank the reviewers for their constructive comments on the final manuscript document. However, they do not accept liability for damages or losses arising from material published in this manuscript.

This manuscript presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the PHR programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the PHR programme or the Department of Health and Social Care.

Copyright © 2024 Melendez-Torres *et al.* This work was produced by Melendez-Torres *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library ([www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)), produced by Newgen Digitalworks Pvt Ltd, Chennai, India ([www.newgen.co](http://www.newgen.co)).





# Abstract

## School-based interventions TO Prevent Dating and Relationship Violence and Gender-Based Violence: STOP-DRV-GBV systematic review

GJ Melendez-Torres<sup>1\*</sup>, Noreen Orr<sup>1</sup>, Caroline Farmer<sup>1</sup>, Naomi Shaw<sup>1</sup>,  
Annah Chollet<sup>2</sup>, Andrew J Rizzo<sup>3</sup>, Fraizer Kiff<sup>1</sup>, Emma Rigby<sup>4</sup>,  
Ann Hagell<sup>4</sup>, Sidnei R Priolo Filho<sup>5</sup>, Bruce Taylor<sup>6</sup>,  
Honor Young<sup>7</sup>, Chris Bonell<sup>8</sup> and Vashti Berry<sup>1</sup>

<sup>1</sup>Faculty of Health and Life Sciences, University of Exeter, Exeter, UK

<sup>2</sup>Department of Social Policy and Intervention, University of Oxford, Oxford, UK

<sup>3</sup>College of Health and Human Performance, University of Florida, Gainesville, FL, USA

<sup>4</sup>Association for Young People's Health, London, UK

<sup>5</sup>Universidade Tuiuti do Paraná, Curitiba, Paraná, Brazil

<sup>6</sup>National Opinion Research Center, University of Chicago, Chicago, IL, USA

<sup>7</sup>Centre for Development, Evaluation, Complexity and Implementation in Public Health Improvement (DECIPHer), Cardiff University, Cardiff, UK

<sup>8</sup>Department of Public Health, Environments and Society, London School of Hygiene and Tropical Medicine, London, UK

\*Corresponding author [g.j.melendez-torres@exeter.ac.uk](mailto:g.j.melendez-torres@exeter.ac.uk)

**Background:** Schools have a duty of care to prevent violence between students but a significant amount of dating and relationship violence and gender-based violence occurs in schools. These are important public health issues with important longitudinal consequences for young people.

**Objectives:** To understand functioning and effectiveness of school-based interventions for the prevention of dating and relationship violence and gender-based violence.

**Review methods:** We undertook a mixed-methods systematic review to synthesise different types of evidence relating to school-based interventions for the prevention of dating and relationship violence and gender-based violence to understand if, how and in what ways these interventions are effective.

We searched 21 databases and 2 trial registers and undertook forwards and backwards citation chasing, author contact and other supplementary search methods. Searches identified all literature published to June 2021. All screening was undertaken in duplicate and independently, and we quality appraised all included studies.

**Results:** We included 247 reports (68 outcome evaluations, 137 process evaluations). Synthesis of intervention components produced an intervention typology: single-component, curricular, multicomponent, and multilevel programmes. Synthesis of intervention theories suggested that interventions aiming to increase students' sense of school belonging and sense of safety in the school building could encourage increased learning of prosocial skills and increased prosocial peer norms, and so potentially reducing dating and relationship violence and gender-based violence. Synthesis of factors affecting delivery highlighted school organisation and leaders who believed in the importance of addressing dating and relationship violence/gender-based violence, along with time and resources

to deliver the interventions. The ease with which the intervention could be delivered and modified was also important.

Meta-analysis found stronger evidence for intervention effectiveness in reducing dating and relationship violence than for gender-based violence, with significant long-term impacts on dating and relationship violence victimisation and perpetration, and some evidence that interventions in high-income countries could be effective for reducing victimisation and perpetration of gender-based violence in the long-term. Impacts on knowledge and attitudes were primarily short-term. Network meta-analysis did not suggest superiority of any intervention type. Moderation evidence suggested interventions reduced dating and relationship violence perpetration in boys more than girls, but reduced gender-based violence perpetration more in girls. Metaregression by intervention component did not explain heterogeneity in effectiveness, but qualitative comparative analysis suggested that reducing perpetration was important to reducing victimisation, and that perpetration could be reduced via focus on interpersonal skills, guided practice and (for gender-based violence) implementation of social structural components.

**Limitations:** Despite an exhaustive search, trials may have been missed and risk of publication bias was high for several analyses.

**Conclusions:** This is the most comprehensive systematic review of school-based interventions for dating and relationship violence and gender-based violence to date. It is clear that the prevention of dating and relationship violence and gender-based violence in schools will require longer-term investment to show benefit.

**Future work:** Future research is needed to understand why intervention effectiveness appears stronger for dating and relationship violence than gender-based violence.

**Study registration:** The study is registered as PROSPERO CRD42020190463.

**Funding:** This award was funded by the National Institute for Health and Care Research (NIHR) Public Health Research programme (NIHR award ref: NIHR130144) and is published in full in *Public Health Research*; Vol. 12, No. 3. See the NIHR Funding and Awards website for further award information.

# Contents

List of tables	xv
List of figures	xvii
List of supplementary materials	xix
List of abbreviations	xxi
Plain language summary	xxiii
Scientific summary	xxv
<b>Chapter 1 Background</b>	<b>1</b>
Description of the problem	1
<i>Dating and relationship violence and gender-based violence: joint constructs</i>	1
<i>Dating and relationship violence and gender-based violence: public health issues</i>	1
<i>School-based interventions for dating and relationship violence and gender-based violence in the UK</i>	2
Description of the intervention	2
Rationale for the current study	3
<i>UK policy context</i>	4
Review aim and objectives	5
<b>Chapter 2 Review methods</b>	<b>7</b>
Research design overview	7
Inclusion criteria for the review	7
<i>Types of population</i>	7
<i>Types of intervention</i>	7
<i>Types of control</i>	8
<i>Types of outcome</i>	8
<i>Types of study</i>	8
Search methods for the identification of studies	8
<i>Database search strategy</i>	8
<i>Search strategy for other literature sources</i>	9
Information management and study selection	10
Data extraction and appraisal	10
Assessment of quality and risk of bias	11
<i>Assessment of process evaluations</i>	11
<i>Assessment of outcome evaluations</i>	12
<i>Assessment of economic studies</i>	12
Data analysis	12
<i>Research question 1: theories and components</i>	13
<i>Research question 2: implementation and process evaluations</i>	13
<i>Research question 3: intervention effectiveness, moderation, mediation and economics</i>	14
<i>Research question 4: metaregression and qualitative comparative analysis</i>	16
<i>Research question 5: network meta-analysis</i>	17
<i>Research question 6: realist integration of findings</i>	17
Policy and practice consultation	18
Ethics approval	18

<b>Chapter 3</b> Overview of included studies	<b>19</b>
Results of the search	19
Included studies and reports	19
Study and intervention characteristics	20
<i>Study designs</i>	20
<i>Intervention characteristics</i>	24
<i>Participant characteristics</i>	26
<b>Chapter 4</b> Typology of intervention approaches and synthesis of theories of change (RQ1)	<b>31</b>
Included studies	31
<i>Intervention components analysis</i>	31
<i>Synthesis of theories of change</i>	32
Setting the scene: key theoretical frameworks	32
Intervention typology	33
<i>Intervention type</i>	33
<i>Intervention focus</i>	35
<i>Intervention domains of implementation</i>	36
Intervention theory of change	37
<i>Starting theoretical propositions</i>	37
<i>Intervention inputs</i>	37
<i>Intervention mechanisms of change</i>	43
<i>Hypothesised intervention outcomes</i>	49
Discussion and propositions	49
<b>Chapter 5</b> Synthesis of process evaluations (RQ2)	<b>53</b>
Included reports	53
Synthesis of process evaluation reports	53
Implementation factors at the school level	55
<i>School resources and infrastructure → ease of implementation and determined necessity of modification</i>	55
<i>Time constraints → dosage and fidelity</i>	56
<i>Dating and relationship violence/gender-based violence comprehension and perception → influences school and community support</i>	56
Implementation factors at the student level	57
<i>Intervention interactivity → improved student engagement and acceptability</i>	57
<i>Poor student attendance among certain populations → reduced intervention reach</i>	58
<i>Positive relationship formation → improves student comfort and engagement over time</i>	58
<i>Programme fit to student population → student and staff acceptability, student engagement and programme fidelity</i>	59
<i>Student immaturity and disruptive behaviours → inhibited student engagement and intervention delivery</i>	60
Implementation factors at the facilitator level	61
<i>Facilitator content knowledge → intervention fidelity and student acceptability</i>	61
<i>Support from the 'other side' → facilitate implementation and increase facilitator comfort</i>	61
<i>Embedding intervention into school curriculum → increase intervention sustainability</i>	62
Implementation factors at the intervention level	62
<i>Ease of programme delivery → increased programme fidelity</i>	62
<i>Ease of programme modification → increased programme fidelity and effectiveness</i>	62
<b>Chapter 6</b> Meta-analysis and narrative synthesis of intervention effectiveness (RQ3)	<b>65</b>
Reports included in outcome evaluation synthesis	65
Quality assessment	65
Outcomes presented	66

<i>Dating and relationship violence victimisation and perpetration outcomes</i>	66
<i>Gender-based violence victimisation and perpetration outcomes</i>	67
<i>Dating and relationship violence and gender-based violence knowledge, attitudes and behaviours</i>	67
Syntheses presented	69
<i>Dating and relationship violence</i>	69
<i>Gender-based violence</i>	73
<i>Role of knowledge and attitudes</i>	74
<i>Bystander outcomes</i>	77
<i>Comparability of dating and relationship violence and gender-based violence evidence</i>	77
Sensitivity and bias analyses	77
<i>Stratification by country context</i>	77
<i>Publication bias analyses</i>	78
GRADE analyses	79
Economic evaluations	79
<b>Chapter 7 Synthesis of mediation and moderation evidence (RQ3)</b>	<b>81</b>
Synthesis of mediation evidence	81
<i>Characteristics of included mediation evidence</i>	81
<i>Mediation of dating and relationship violence victimisation outcomes</i>	81
<i>Mediation of dating and relationship violence perpetration outcomes</i>	82
<i>Mediation of gender-based violence victimisation outcomes</i>	83
<i>Mediation of gender-based violence perpetration outcomes</i>	84
<i>Mediation of bystander outcomes</i>	84
Synthesis of moderation evidence: narrative synthesis and harvest plots	84
<i>Characteristics of included moderation evidence</i>	84
<i>Introducing harvest plots</i>	85
<i>Moderation of dating and relationship violence victimisation</i>	86
<i>Moderation of dating and relationship violence perpetration</i>	89
<i>Moderation of gender-based violence victimisation</i>	91
<i>Moderation of gender-based violence perpetration</i>	93
Metaregression analyses	95
<i>Dating and relationship violence victimisation</i>	96
<i>Dating and relationship violence perpetration</i>	96
<i>Gender-based violence victimisation</i>	96
<i>Gender-based violence perpetration</i>	97
<b>Chapter 8 Analysis of heterogeneity in intervention effectiveness (RQ4, RQ5)</b>	<b>99</b>
Metaregression of intervention components	99
<i>Discussion components</i>	99
<i>Visual components</i>	99
<i>Competition components</i>	99
<i>Guided practice components</i>	99
<i>Reflection components</i>	108
<i>Narrative-based components</i>	108
<i>School personnel components</i>	108
<i>Parent and family components</i>	108
<i>Structural environmental components</i>	108
<i>Structural social components</i>	108
<i>Guidance: self, peer, adult</i>	109
<i>Gender composition</i>	109
Qualitative comparative analysis	109
<i>Initial calibration and conditions</i>	109
<i>Dating and relationship violence victimisation</i>	110

<i>Dating and relationship violence perpetration</i>	112
<i>Gender-based violence victimisation: short-term</i>	114
<i>Gender-based violence perpetration: short-term</i>	115
Network meta-analysis	115
<i>Feasibility assessment</i>	115
<i>Dating and relationship violence victimisation: short-term</i>	116
<i>Dating and relationship violence victimisation: long-term</i>	116
<i>Dating and relationship violence perpetration: short-term</i>	119
<i>Dating and relationship violence perpetration: long-term</i>	119
<i>Gender-based violence victimisation: short-term</i>	119
<i>Gender-based violence victimisation: long-term</i>	122
<i>Gender-based violence perpetration: short-term</i>	122
<i>Gender-based violence perpetration: long-term</i>	122
<i>Component network meta-analyses</i>	122
<b>Chapter 9 Discussion</b>	<b>125</b>
Summary and realist reflection on empirical findings (RQ6)	125
<i>Programme components</i>	125
<i>Programme theory</i>	125
<i>Influences on successful implementation</i>	126
<i>Programme effects on outcomes</i>	127
<i>Evidence of programme effects on intermediate outcomes</i>	128
<i>Mediation</i>	128
<i>Moderation</i>	129
<i>Explaining heterogeneity in effectiveness</i>	130
<i>Realist integration</i>	130
Consultation with stakeholders	133
<i>What does gender-based violence/dating and relationship violence education look like in England right now?</i>	134
<i>What are the priorities and opportunities in gender-based violence/dating and relationship violence education for this research to focus on?</i>	134
<i>What are the gaps in the evidence base and what are the other ways in which we should consider/categorise the studies and the interventions in the review?</i>	135
<i>How can we make the most of the alignment of this review with existing agendas and approaches?</i>	135
<i>What do our findings mean for intervention functioning, and for current provision in the UK?</i>	136
Consultations with young people	137
<i>Immanence of gender-based violence and presence of dating and relationship violence</i>	137
<i>Saliency and appropriateness of dating and relationship violence and gender-based violence education</i>	137
<i>The value of proactive intervention</i>	137
<i>Reflections on consultation</i>	137
Implications of the findings	138
<i>In light of prior research</i>	138
<i>In light of policy and practice</i>	140
<i>Where further research is needed</i>	141
Strengths and limitations	142
Equity, diversity and inclusion	142
Conclusion	143

<b>Acknowledgements</b>	<b>145</b>
<b>References</b>	<b>147</b>
<b>Appendix 1</b> Process documentation	<b>169</b>
<b>Appendix 2</b> Extra information about review methods	<b>191</b>





## List of tables

<b>TABLE 1</b> Criteria used to derive overall quality ratings of process evaluations	<b>11</b>
<b>TABLE 2</b> Continent/countries of included process evaluations	<b>21</b>
<b>TABLE 3</b> School settings of included outcome evaluations	<b>22</b>
<b>TABLE 4</b> Continent/country of included outcome evaluations	<b>23</b>
<b>TABLE 5</b> Head-to-head comparisons evaluated for outcome (efficacy)	<b>25</b>
<b>TABLE 6</b> Intervention facilitators (outcome evaluations)	<b>25</b>
<b>TABLE 7</b> Method for delivering the intervention (outcome evaluations)	<b>26</b>
<b>TABLE 8</b> Coding template	<b>32</b>
<b>TABLE 9</b> Mechanisms of change by ecological level	<b>50</b>
<b>TABLE 10</b> Summary of implementation syntheses	<b>53</b>
<b>TABLE 11</b> Meta-analysis results, DRV victimisation	<b>70</b>
<b>TABLE 12</b> Meta-analysis results, DRV perpetration	<b>70</b>
<b>TABLE 13</b> Meta-analysis results, DRV knowledge, attitudes and behaviours	<b>71</b>
<b>TABLE 14</b> Meta-analysis results, GBV victimisation	<b>74</b>
<b>TABLE 15</b> Meta-analysis results, GBV perpetration	<b>75</b>
<b>TABLE 16</b> Meta-analysis results, GBV knowledge, attitudes and behaviours	<b>76</b>
<b>TABLE 17</b> Metaregression results, country context	<b>78</b>
<b>TABLE 18</b> GRADE analyses for incidence of DRV and GBV	<b>80</b>
<b>TABLE 19</b> Metaregression, percentage of girl children	<b>96</b>
<b>TABLE 20</b> Component metaregression findings for DRV victimisation	<b>100</b>
<b>TABLE 21</b> Component metaregression findings for DRV perpetration	<b>102</b>
<b>TABLE 22</b> Component metaregression findings for GBV victimisation	<b>104</b>
<b>TABLE 23</b> Component metaregression findings for GBV perpetration	<b>106</b>
<b>TABLE 24</b> Qualitative comparative analysis for short-term DRV victimisation	<b>111</b>
<b>TABLE 25</b> Qualitative comparative analysis for long-term DRV victimisation	<b>112</b>

<b>TABLE 26</b>	Qualitative comparative analysis for long-term DRV perpetration	<b>113</b>
<b>TABLE 27</b>	Qualitative comparative analysis for short-term GBV victimisation	<b>114</b>
<b>TABLE 28</b>	Qualitative comparative analysis for short-term GBV perpetration	<b>115</b>
<b>TABLE 29</b>	Network meta-analysis results for DRV victimisation and perpetration outcomes	<b>117</b>
<b>TABLE 30</b>	Network meta-analysis results for GBV victimisation and perpetration outcomes	<b>120</b>
<b>TABLE 31</b>	Component network meta-analysis results for DRV outcomes	<b>123</b>
<b>TABLE 32</b>	Protocol amendments	<b>169</b>
<b>TABLE 33</b>	Protocol deviations and clarifications	<b>169</b>
<b>TABLE 34</b>	Overview of studies included in the review	<b>171</b>
<b>TABLE 35</b>	Full list of included publications	<b>186</b>
<b>TABLE 36</b>	Overview of data points extracted across included studies	<b>191</b>

# List of figures

<b>FIGURE 1</b> Initial logic model for understanding DRV and GBV outcomes	<b>3</b>
<b>FIGURE 2</b> Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart	<b>19</b>
<b>FIGURE 3</b> Location of process/implementation evaluations	<b>21</b>
<b>FIGURE 4</b> Location of outcome evaluations	<b>23</b>
<b>FIGURE 5</b> Duration of interventions according to intervention target (outcome evaluations)	<b>26</b>
<b>FIGURE 6</b> Majority ethnicity/race group of process evaluation samples (where reported)	<b>27</b>
<b>FIGURE 7</b> Majority ethnicity/race group of outcome evaluation samples (where reported)	<b>29</b>
<b>FIGURE 8</b> Integrated theory of change for school-based interventions	<b>38</b>
<b>FIGURE 9</b> Critical appraisal ratings for included cRCTs	<b>65</b>
<b>FIGURE 10</b> Critical appraisal ratings for included parallel RCTs	<b>66</b>
<b>FIGURE 11</b> Harvest plot, DRV victimisation, boys vs. girls	<b>87</b>
<b>FIGURE 12</b> Harvest plot, prior history of the outcome, all outcomes	<b>88</b>
<b>FIGURE 13</b> Harvest plot, DRV perpetration, boys vs. girls	<b>90</b>
<b>FIGURE 14</b> Harvest plot, GBV victimisation, boys vs. girls	<b>93</b>
<b>FIGURE 15</b> Harvest plot, GBV perpetration, boys vs. girls	<b>94</b>
<b>FIGURE 16</b> Network diagrams for DRV outcomes	<b>116</b>
<b>FIGURE 17</b> Network diagrams for GBV outcomes	<b>119</b>
<b>FIGURE 18</b> Hypotheses arising from syntheses	<b>132</b>
<b>FIGURE 19</b> Realist integration of findings	<b>133</b>



# List of supplementary materials

**Report Supplementary Material 1** Info about review methods and additional results

**Report Supplementary Material 2** Study-level effect estimates for RCTs included for research question 2

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/KTWR6997>).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.



## List of abbreviations

AAL	Athletes As Leaders	LGBTQ	lesbian, gay, bisexual, transgender, queer
CBIM	Coaching Boys into Men	LMIC	low- and middle-income countries
cRCT	cluster randomised controlled trial	NIHR	National Institute for Health and Care Research
DRV	dating and relationship violence	NMA	network meta-analysis
DV	dating violence	PSHE	personal, social, health and economic education
EPPI-Centre	Evidence for Policy and Practice Information and Co-ordinating Centre	QCA	qualitative comparative analysis
GBV	gender-based violence	RA	relationship abuse
GEMS	gender equity movement in schools	RCT	randomised controlled trial
GRADE	Grading of Recommendations, Assessment, Development and Evaluation	RSE	relationships and sex education
HIC	high-income country	SA	sexual abuse
ICC	intracluster correlation coefficient	SE	standard error
ITT	intention to treat	SES	socioeconomic status
KBEP	Katie Brown Educational Program	SUCRA	surface under the cumulative ranking curve
LGBT	lesbian, gay, bisexual, transgender	VAWG	violence against women and girls





## Plain language summary

Schools are places where dating and relationship violence and gender-based violence occur. Therefore, interventions conducted within schools are ideally placed to prevent and reduce dating and relationship violence and gender-based violence. We reviewed existing research on these interventions, exploring how they were expected to work, what factors affected their implementation in practice, how they had an impact on dating and relationship violence and gender-based violence, and what specific parts of the interventions were most effective and in what contexts. We defined what sort of evidence to include in the review, carried out a comprehensive search and found 247 reports on school-based interventions to prevent dating and relationship violence or gender-based violence, most of which were conducted in North America.

Most interventions aimed to provide students with the knowledge, attitudes and skills needed to prevent perpetration and victimisation. They varied in complexity; some had one activity, others had multiple activities, some were integrated into the existing school curricula and others were complex in that they sought to change how schools as a whole respond to dating and relationship violence or gender-based violence. We theorised that complex interventions would bring about greater and more sustainable change, but this was not supported by our findings. The implementation of interventions was affected by factors such as school organisation and leaders who believed in the importance of addressing dating and relationship violence/gender-based violence, along with time and resources to deliver the interventions. The ease with which the intervention could be delivered and modified was also important.

There was stronger evidence for intervention effectiveness in reducing dating and relationship violence than for gender-based violence, with significant long-term impacts on dating and relationship violence victimisation and perpetration. There was some evidence that interventions in high-income countries could be effective for reducing victimisation and perpetration of gender-based violence in the long-term. More complicated interventions were not more effective, and interventions tended to reduce dating and relationship violence perpetration in boys more than in girls. We have also been able to identify where there are gaps in available evidence, which may provide avenues for future research.



# Scientific summary

## Background and rationale

This systematic review focused on dating and relationship violence (DRV) and gender-based violence (GBV). The long-term impacts of both for young people are numerous and, consequently, addressing them is a public health priority. Adolescence is a crucial stage for focusing on the prevention of DRV and GBV and schools are an ideal location for this, as this is where young people are socialised into gender norms and where significant amounts of DRV and GBV occurs. Schools also offer an opportunity to reach many young people who may not otherwise be accessible for intervention. Previous reviews have evaluated the effectiveness of interventions for DRV or GBV for adolescents but do not specifically focus on interventions in schools or analyse outcomes jointly.

## Aim and review questions

We systematically searched for and synthesised the evidence for the following research questions (RQs):

1. What are the theories of change and components of evaluated interventions?
2. What factors affect the implementation of evaluated interventions?
3. Are interventions effective and cost-effective in preventing DRV and GBV and reducing social inequalities in these outcomes?
4. Based on the findings of RQs 1–3, what factors are important for joint effectiveness on DRV and GBV outcomes?
5. What is the comparative effectiveness of different approaches to DRV and GBV prevention?
6. What do the different sources of evidence suggest about intervention mechanisms and how these are contingent on context?

## Methods

### Inclusion criteria

We included randomised controlled trials and process evaluations of school-based interventions for DRV and/or GBV for children between 5 and 18 years. DRV was defined as physical, sexual and emotional violence (including coercive control) in relationships between young people. GBV was defined as violence rooted in gender equality and sexuality such as harassment or bullying on the basis of gender or sexuality, sexual violence, coercion and assault including rape, within or outside dating relationships.

### Searching information sources

The search strategy included both free-text terms and subject headings (e.g. MeSH in MEDLINE) for the school setting and DRV/GBV outcomes. In order to identify outcome, process and economic evaluations, we did not apply publication type or study design limitations. In July 2020, we searched 21 bibliographic databases from inception and without limitation on date or language: MEDLINE, EMBASE, PsycINFO, Social Policy and Practice (Ovid); Cumulative Index to Nursing and Allied Health Literature, Education Resources Information Center, British Education Index, Education Research Complete, EconLit, Criminal Justice Abstracts (EBSCOhost); Cochrane Database of Systematic Reviews and the Cochrane Central Register of Controlled Trials; NHS Economic Evaluation Database (via the Centre for Reviews and Dissemination); Social Science Citation Index and Conference Proceedings Citation Index (Web of Science, Clarivate Analytics); Australian Education Index, ProQuest Dissertations and Theses Global, Sociological Abstracts including Social Services Abstracts, Applied Social Sciences Index and Abstracts (ProQuest); Trials Register of Promoting Health Interventions and Bibliomap [Evidence for Policy and

Practice Information (EPPI)-Centre]; Campbell Systematic Reviews (Campbell Collaboration). The bibliographic database searches were updated in June 2021 with a revised strategy developed to improve precision, and added further search terms for named interventions.

We completed forwards and backwards citation chasing on included studies in Scopus (Elsevier), Web of Science (Clarivate Analytics) and Google Scholar, and reviewed the reference lists of relevant systematic reviews and reports. To identify linked studies and further grey literature, we conducted targeted searches in Web of Science and Scopus and searched Google Scholar for specific intervention names. We also searched publication lists on key websites.

### ***Information management and study selection***

Studies were screened by the title and abstract by six reviewers. The reviewers screened a random sample of 100 records and then discussed disagreements before proceeding to screening by title and abstract, and then full-text, independently and in duplicate.

### ***Data extraction and assessments of quality***

Data were extracted into a data extraction tool that was developed and piloted a priori. Data extracted included details about the study design, study sample, intervention characteristics, analysis methods and outcome data. Studies were appraised using the Cochrane risk of bias tool or the EPPI-Centre tool.

### ***Synthesis of theories and components***

We synthesised theories of change for the interventions via a staged approach: (1) using line-by-line coding and developing a coding template; (2) synthesising programme theories of change for the interventions and (3) using a meta-ethnographic approach to develop a line-of-argument for an overarching theory of change.

We undertook an intervention components analysis to analyse intervention descriptions. We used an inductive approach (open coding initially and then axial coding) to comprehensively describe and categorise intervention components.

### ***Synthesis of process data***

Process evaluations reported qualitative data and were synthesised qualitatively using thematic synthesis methods.

### ***Synthesis of effectiveness***

Pairwise meta-analysis tested effectiveness on victimisation and perpetration outcomes, knowledge and attitudes, using robust variance estimation meta-analyses with random effects.

### ***Synthesis of mediation and moderator data***

We used harvest plots to examine how interventions impact health inequalities, focusing on ethnicity, socioeconomic position, gender, sexuality and age. We narratively synthesised findings relating to mediation.

### ***Effectiveness of different approaches to dating and relationship violence and gender-based violence prevention***

We used metaregression to test if components explained heterogeneity, and qualitative comparative analysis (QCA) to consider pathways to effectiveness in victimisation and perpetration. We used network meta-analysis to understand the comparative effectiveness of intervention types on DRV and GBV victimisation and perpetration.

### ***Synthesis of evidence on intervention mechanisms contingent on context***

To understand the links between contexts and mechanisms in generating outcomes, we followed a realist synthesis approach and used findings from the syntheses of the theories of change and process

evaluations as a framework to infer and induce mechanisms from studies. This allowed previously 'untheorised' findings to emerge as relevant from our synthesis.

## Results

### *Included studies*

Searches identified 40,160 records after deduplication, of which 793 were screened in full text. Of these, 247 reports were identified as eligible for inclusion, and these were coded into 68 outcome evaluations and 137 process evaluations. No economic evaluations were identified, but we examined seven cost and resource use studies.

### *What are the theories of change and components of evaluated interventions?*

Most interventions focused on preventing victimisation and perpetration in DRV or GBV and fewer addressed bystander-intervention skills. Interventions had student components (e.g. group discussions, individual reflection), staff components (e.g. training, lesson plans), parent/family components (e.g. involvement in intervention) and school structures or physical environment (e.g. changes to school policies). Interventions were categorised as single-component interventions (usually short and focused on a novel technology or activity) and generally focused on a single change mechanism; curricular programmes integrated into the wider school curriculum, which generally focused on one or more change mechanisms at the student level; multicomponent programmes involving multiple modes of intervention and focusing on multiple change mechanisms operating at the student or staff level; and multilevel programmes, which were complex interventions involving multiple modes focused on change mechanisms within schools at multiple levels including the individual, classroom and school structural/environment context.

The interventions were theorised as aiming to trigger a complex set of mechanisms to promote students' school belonging, engagement with pro-social behaviours and avoidance of DRV/GBV behaviours. This was theorised to occur through interventions that 'weakened classification' and 'reframing'. 'Weakened classification' involved strengthening relationships between and among staff and students, between the classroom and the wider school, and between schools and their communities. 'Reframing' aimed to increase student involvement in decisions at the level of the classroom and the school. Both 'weakening classification' and 'reframing' were then theorised as increasing student belonging and a sense of safety in the school building which could encourage increased learning of prosocial skills and interactions. Not all interventions addressed all of these mechanisms; some multilevel interventions could trigger mechanisms at multiple levels of the school system; classroom-level interventions could trigger mechanisms at the level of staff-student relationships or could focus on mechanisms that sought, for example, to promote specific skills for preventing or reducing DRV/GBV. It was theorised that multilevel interventions triggering a complex set of mechanisms across multiple levels would achieve larger effects (e.g. at the school level) and be more sustainable than mechanisms at the individual and group level.

### *What factors affect the implementation of evaluated interventions?*

Key factors influencing the implementation of interventions were school resources and infrastructure, space and supplies, school organisation and leadership, perceived importance of addressing DRV/GBV, intervention interactivity and development of positive relations among students, facilitator content knowledge and availability of external support, ease of delivery and modification of interventions to suit the particularities of settings. Significant barriers to implementation included time constraints and competing priorities. However, strong staff commitment to prevention of DRV/GBV could offset time and resource limitations. Interventions may be best delivered when they align with school organisational readiness which includes a receptive school climate, staff buy-in and/or strong school leadership.

### ***Are interventions effective in preventing dating and relationship violence and gender-based violence and reducing social inequalities in these outcomes?***

Significant long-term, but not short-term, impacts on DRV victimisation [odds ratio (OR) = 0.82, 95% confidence interval (CI) (0.68 to 0.99)] and DRV perpetration [OR = 0.78, 95% CI (0.64 to 0.94)] were found. Although there was no overall effect for GBV victimisation [long-term OR = 0.93, 95% CI (0.80 to 1.08)] or perpetration [long-term OR = 0.90, 95% CI (0.73 to 1.12)], there was some evidence that interventions in high-income countries could be effective for reducing victimisation and perpetration of GBV in the long-term, and that the proportion of girls in the trial sample moderated the effect of the interventions for DRV and GBV victimisation, but not DRV or GBV perpetration. This may be where a critical mass of female students might have encouraged greater overall student engagement with the intervention. There was evidence for gender moderating programme effects on DRV perpetration with greater benefits for boys, particularly for emotional and physical DRV perpetration. Interventions improved DRV violence acceptance, knowledge and attitudes to personal help-seeking in the short term only. Interventions improved GBV violence acceptance, knowledge and individual self-efficacy in the short term, and violence acceptance in the long-term.

### ***What factors are important for joint effectiveness on dating and relationship violence and gender-based violence outcomes?***

Metaregression of intervention components did not explain heterogeneity in intervention effectiveness. We were able to estimate QCA models for short-term and long-term DRV victimisation, long-term DRV perpetration and short-term GBV victimisation and perpetration. A key finding from the QCA models for victimisation is that a central causal condition for reduction of victimisation is reduction of perpetration. There were a number of other pathways to the reduction of victimisation, namely the inclusion of single-gender components or a critical mass of girls. A critical mass of girls was especially important where interventions went beyond single components. There was also some evidence that components that were absent, for example, the absence of parental involvement was central to achieving effectiveness for long-term DRV victimisation. Perhaps the absence of such a component reduces opportunities to minimise GBV and DRV, or to receive conflicting messages about their importance. For long-term DRV perpetration and short-term GBV perpetration, interventions that were most effective incorporated a range of opportunities for guided practice of skills and attitudes, and interpersonal components focusing on student relationships. The implementation of social structural components was central to effectiveness for short-term GBV perpetration, but not for DRV perpetration.

### ***What is the comparative effectiveness of different approaches to dating and relationship violence and gender-based violence prevention?***

Network meta-analysis was consistent for DRV but not for GBV outcomes. No one intervention type was clearly more effective than any other, but single-component interventions may have been useful for reducing short-term and long-term DRV victimisation and perpetration, and short-term GBV victimisation. Multilevel interventions showed some effectiveness for DRV outcomes compared to other intervention formulations. For GBV outcomes, there was strong evidence for the role of curriculum interventions which were more successful than any other types at short-term follow-up victimisation outcomes, and short-term and long-term perpetration.

### ***What do the different sources of evidence suggest about intervention mechanisms and how these are contingent on context?***

There is evidence that the interventions in this review worked not by complex mechanisms (e.g. increasing school commitment), as initially hypothesised, but by the 'basic safety' mechanism which aimed to disrupt violent behaviours by communicating to students the unacceptability of violence. This simpler mechanism may have involved reductions in DRV perpetration among males. That simpler interventions if delivered well may be effective in reducing DRV would be a significant finding, especially for resource-poor settings. We argue that destabilising harmful practices is foundational whereas establishing prosocial behaviours will take more time and investment, so schools may focus on foundational mechanisms initially. Simpler interventions were easier to implement and receive so it may

be more important to deliver simpler programmes achieving basic levels of school safety with fidelity than attempt more complex, multicomponent or multilevel programmes. It may be particularly important to prioritise such programmes in resource-poor settings.

We also theorise that the individual-level basic safety mechanisms are more likely to effect change in DRV than in GBV perpetration because of the more dyadic, private nature of DRV. The more public nature of GBV means that it might be influenced by social norms that programmes do not seem to successfully address. We found that programmes could be effective in preventing GBV but that this was only likely in high-income settings. It may be that the transformation of school organisation and culture required to reduce GBV is beyond the reach of many schools in low- and middle-income settings because of low school organisational capacity and higher levels of inequality and GBV.

## Conclusions

The evidence suggests that a well-delivered single-component intervention may be as effective as a complex, multicomponent or multilevel intervention. Such an intervention may be more acceptable within the school environment and receive greater staff buy-in. Preventing GBV may require normative and social structural change which may not be sufficiently triggered by current interventions. Overall, the evidence is more conclusive for DRV than GBV. It is possible that differential mechanisms will need to be activated for DRV than for GBV and that existing theories of change do not adequately account for differences between DRV and GBV in terms of intervention functioning.

## Study registration

This study is registered as PROSPERO CRD42020190463.

## Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Public Health Research programme (NIHR award ref: NIHR130144) and is published in full in *Public Health Research*; Vol. 12, No. 3. See the NIHR Funding and Awards website for further award information.





# Chapter 1 Background

## Description of the problem

### *Dating and relationship violence and gender-based violence: joint constructs*

Dating and relationship violence (DRV) refers to physical, sexual and emotional violence (including coercive control) in relationships between young people. Critically, it includes those aged under 16, who are not included in UK government definitions of domestic violence. Gender-based violence (GBV) refers to violence rooted in gender inequality and sexuality, such as harassment or bullying on the basis of gender or sexuality, sexual violence, coercion and assault including rape, within or outside dating relationships.<sup>1</sup> Although longitudinal evidence demonstrates that experience of DRV predicts young people's later GBV victimisation and that they share common risk factors,<sup>2,3</sup> they are rarely considered as joint constructs.<sup>4</sup>

Previous systematic reviews of interventions for young people have focused on DRV and have not meaningfully considered intervention impacts with GBV.<sup>5-8</sup> This is important because interventions nominally focusing on DRV may impact GBV and vice versa, underpinned by common mechanisms and structural features that lead to high rates of both in schools. Potentially shared aetiological mechanisms include gender norms at the societal level that are inequity-generating (i.e. creating unavoidable and unfair differences between groups); inconsistent development and enforcement of violence prevention policies at school and classroom levels, and the creation of spaces for violence to occur; and, at the individual level, exposure to and reinforcement of antisocial norms relating to gender, sexuality and violence and, conversely, insufficient exposure and reinforcement to prosocial norms relating to the same.<sup>4,9-11</sup> These multilevel influences are experienced across sexual and reproductive health, of which DRV and GBV are important determinants.<sup>12</sup> That is, DRV and GBV have a common basis in exploiting gender and sexual inequalities and in antisocial norms.<sup>13</sup>

Dating and relationship violence and GBV are amenable to change in school contexts via a range of approaches, from didactic (e.g. classroom-based) to structural (e.g. school policy changes).<sup>1</sup> Thus, considering DRV and GBV jointly is essential to develop, implement and adapt efficient interventions for schools.

### *Dating and relationship violence and gender-based violence: public health issues*

Dating and relationship violence and GBV are pressing public health problems with manifold and inequity-generating long-term impacts on health. While boys and girls both experience major burdens of emotional and physical DRV, impacts are disproportionately experienced by girls. A cross-sectional study based on a representative sample of 11- to 16-year-olds in Wales found that 28% of girls report emotional victimisation and 12% report physical victimisation by a partner over the course of adolescence, while 20% of boys report emotional victimisation and 17% physical victimisation.<sup>11</sup> However, age-related trajectories in victimisation are steeper in girls than in boys, suggesting that adolescence is a critical period to arrest inequalities arising from DRV and GBV. Girls are more likely than boys to report sexual dating violence victimisation,<sup>14</sup> and girls are more vulnerable to experiencing the detrimental impacts of victimisation such as feelings of fear, distress and post-traumatic stress compared to boys.<sup>15</sup> Nationally representative data from the US Youth Risk Behavioral Surveillance Survey from 2001 to 2019 revealed that rates of forced sex were maintained for girls and decreased for boys, and as girls and boys aged, the risk of forced sex increased.<sup>16</sup> In Britain, the median age for most recent occurrence of sex against one's will, a form of GBV, is 18 among men and 16 among women.<sup>17</sup> In addition, longitudinal evidence suggests that the onset of physical DRV and GBV peaks in mid-adolescence, while onset of sexual DRV and GBV is greatest in late adolescence.<sup>18</sup> This underscores the importance of interventions during school years.

Young people do not report perceived peer sanctions against GBV behaviours,<sup>1</sup> and at an individual level, norms accepting of GBV and harassment strongly correlate with DRV perpetration and victimisation, reinforcing the importance of considering these outcomes jointly.<sup>1,13,19,20</sup> In 2021 in the UK, 67% of girls aged 13–18 years reported sexual harassment at school or college; 18% experienced unwanted touching such as being pinned down, having their bra strap or skirt pulled. This increased from 9% for those aged 13–16 years, to 36% for those aged 17–18 years.<sup>21</sup> While it is well understood that sexual minority adolescents experience higher levels of GBV in terms of homophobic and transphobic bullying and sexual harassment,<sup>22,23</sup> these adolescents also experience higher rates of physical and sexual DRV.<sup>23–25</sup>

Longitudinal impacts of DRV and GBV are numerous. In adolescence, both perpetrators and victims report increased risky sexual behaviour, substance use and depressive symptoms;<sup>8,10,23,26</sup> in adulthood, survivors of DRV and GBV are more likely to be re-victimised<sup>27</sup> and more likely to report poorer mental and physical health.<sup>28</sup> In particular, a systematic review of longitudinal studies found that both DRV and GBV experiences as adolescents were predictive of adult experiences of domestic violence.<sup>29</sup> Another way in which DRV and GBV are inequity-generating is in their exacerbation of health inequalities between men and women;<sup>30</sup> in particular, earlier onset of intimate partner violence leads to greater impacts on mental and physical health in adulthood.<sup>28</sup> In addition, there are strong intersections with other inequalities, such as race/ethnicity and sexuality.<sup>22,23</sup> DRV and GBV generates inequalities between heterosexual and cisgender young people and their sexual minority peers, such as substance misuse and increased burden of suicidal ideation arising from experiences of DRV and GBV.<sup>22,23</sup> Importantly, a key source of these inequalities in mental health is the shared impact of school context, including prevalence and response to DRV and GBV, both of which point to the importance of school-based interventions.<sup>31</sup>

### **School-based interventions for dating and relationship violence and gender-based violence in the UK**

Adolescence is a crucial stage for focusing on the prevention of DRV and GBV. Schools, as important sites of gender socialisation (both in the UK and worldwide), have the potential to promote gender-equitable attitudes. In addition, schools are a critical site for the delivery of universal interventions. However, despite their role in establishing prosocial norms and behaviours and their duty of care to prevent violence between pupils, a significant amount of DRV and GBV occurs in schools (e.g. Ofsted 2021<sup>32</sup>). According to a recent report on sexual abuse in schools in England,<sup>32</sup> sexual harassment including online sexual abuse is normalised for children in schools and incidents of sexual harassment are so commonplace that they see little point in reporting them. Herbert *et al.* (2021)<sup>33</sup> observe that it is only relatively recently that there has been a sustained UK public health focus on DRV in younger people,<sup>10,11,34</sup> with the first trial of a school-based intervention to prevent DRV in young people, Project Respect,<sup>35</sup> trialled in the UK.

### **Description of the intervention**

This systematic review focuses on school-based interventions for the prevention of DRV or GBV, provided to students in compulsory education (aged 5–18). An indicative logic model is depicted in [Figure 1](#). These interventions draw on a range of approaches used in the school context, from 'traditional' classroom-based instruction as part of, or distinct from, relationships and sex education (RSE) through to school-level resourcing and restructuring. Didactically led programmes have been extensively evaluated. For example, Safe Dates included a RSE curriculum;<sup>36</sup> Second Step included classroom-based social and emotional learning;<sup>37</sup> and TakeCARE used a video-based programme to teach bystander behaviour, or increased self-efficacy to intervene, with the goal of reducing DRV and GBV.<sup>38</sup> However, explicit consideration of structural components is important given the presence of school 'hot spots' for violence, including DRV and GBV,<sup>39</sup> and the potential value of staff-led responses in terms of increased monitoring and other place-based approaches.<sup>34</sup> Previous reviews have not attended to these structural components. For example, Safe Dates increased services to adolescents in abusive relationships, and sought to upskill teachers and community service providers.<sup>36</sup> In addition,

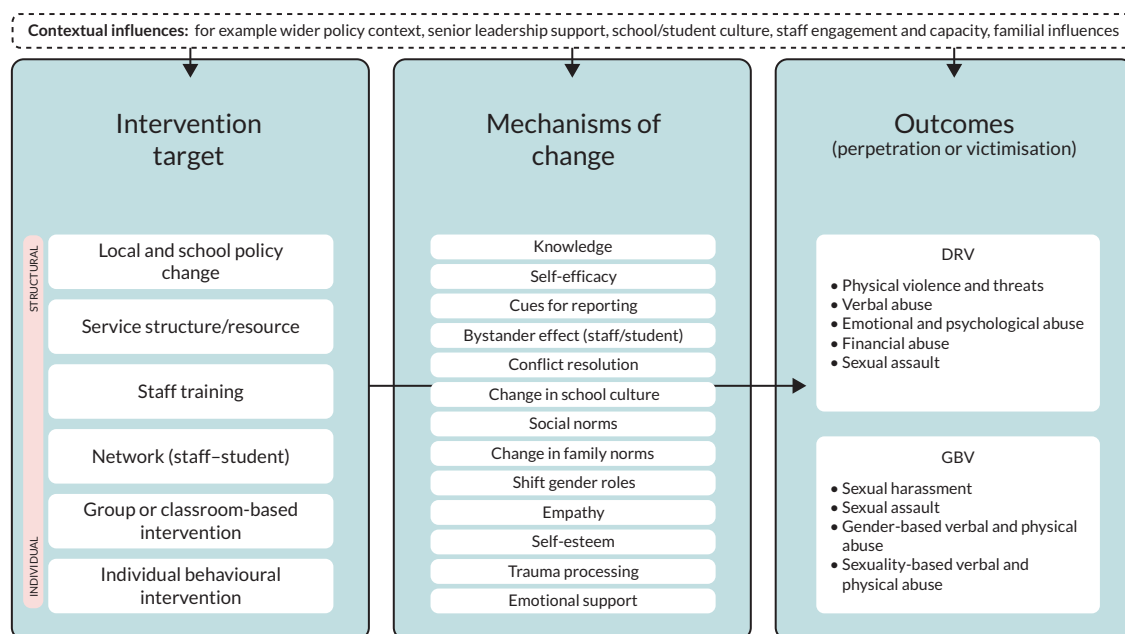


FIGURE 1 Initial logic model for understanding DRV and GBV outcomes.

Shifting Boundaries, which compared a didactic and structural package against a structural-only package (building-based restraining orders, greater faculty and security staff in hot spots, school media campaign) and against no intervention, found reductions in sexual violence perpetration in the structural-only intervention alone.<sup>40</sup> To our knowledge, structural components have not yet been considered in a systematic review. The mechanisms through which interventions may impact DRV and GBV outcomes are accordingly broad, including improved knowledge and self-efficacy, improved reporting and bystander behaviours and better conflict resolution skills through to changes in school culture and responses,<sup>31</sup> and social norms at the group level.<sup>12</sup>

## Rationale for the current study

There is no recent systematic review examining the evidence on the effectiveness of school-based interventions for GBV. Systematic reviews published since 2013 have focused on interventions for the prevention of DRV rather than GBV and have not specifically synthesised GBV-related evidence.<sup>5-8,41,42</sup> The search strategies for these reviews did not use comprehensive coverage of search terminology for both DRV and GBV outcomes. Consequently, they provide only a partial picture of the evidence base by omitting relevant studies, either with respect to the range of DRV and GBV outcomes or by restricting to DRV alone. Three reviews<sup>5,7,41</sup> also appeared to miss at least one randomised controlled trial (RCT) that was within scope, and since this review began, a further review was published that excluded numerous relevant RCTs.<sup>42</sup> This may be due to variation in the aims of the review authors, though may also suggest that evaluations of school-based interventions may sometimes be harder to identify by common approaches to search strategies (e.g. white papers reporting on the findings of an evaluation may only be found through a comprehensive grey literature search strategy). Reviews also commonly exclude research published not in English, excluding a substantial amount of work on GBV worldwide.

Arguably, artificial distinctions between DRV and GBV have led to variable inclusion of studies, outcomes and effect estimates both across and within reviews.<sup>5-8,43</sup> Furthermore, artificial distinctions between outcomes and intervention strategies preclude a clear picture of the evidence. The shared mechanisms linking DRV and GBV constitute an important reason to consider these outcomes jointly. Some reviews<sup>5,6,41</sup> have excluded important forms of GBV that may or may not occur in the context of dating relationships, such as unwanted sexting, coercive control and sexual harassment.

Some of the reviews cited<sup>6,8,41-43</sup> have included interventions across age ranges and settings rather than focusing interventions on compulsory education settings specifically, which is most relevant to inform policy. In addition, existing reviews have not synthesised evidence on structural intervention components (e.g. school-level policy change) and have not examined comprehensively heterogeneity in effectiveness by intervention characteristics.<sup>5-8,41-43</sup> Only one review considered implementation aspects and cost effectiveness of interventions but was focused on domestic abuse.<sup>7</sup> Our scoping work suggested that there is a rich evidence base sufficient to go beyond questions of effectiveness and given the size of the evidence base, there is a good rationale for a targeted synthesis that focuses on school-based interventions in compulsory education (as opposed to adolescents generally) to understand how school contexts specifically shape intervention functioning and effectiveness.

A review that addresses these important knowledge gaps is very timely because of the pressing need to prevent DRV and GBV in UK schools.<sup>32</sup> First, as schools move away from single-issue interventions in the context of increasingly crowded timetables, understanding how interventions can target multiple related outcomes will meet a growing policy and practice need.<sup>44,45</sup> DRV and GBV, given potentially shared aetiological mechanisms and risks, form an ideal candidate for 'joint action'. Similarly, while it is important to consider these outcomes jointly, it is also important to identify which intervention strategies are most effective for one type of outcome or another, and which intervention strategies are best bets across the range of DRV and GBV outcomes. Second, attention to how interventions include multiple and multilevel strategies is important as single-issue, single-component interventions may be more likely to wash out of complex systems such as schools.<sup>46,47</sup> Third, a focus on didactic interventions alone without synthesis of structural and organisational intervention strategies does not account for the potential role of multilevel interventions in reducing health inequalities.<sup>48</sup> Finally, attention to internet-mediated DRV<sup>32,49</sup> meets a pressing policy and practice need identified with stakeholders to understand how digital culture in school and youth contexts shapes newer forms of violence.

### **UK policy context**

Importantly, all of this is directly relevant to UK policy and practice. The Department for Education released the final guidance for statutory RSE requiring all schools in England to deliver statutory Relationships Education (RE) in primary schools and RSE in secondary. It came into effect in September 2020, but schools could start using it from September 2019. Those schools not in a position to implement in September 2020 because of the COVID-19 pandemic and closures, had until summer term 2021 to implement. Implementation analysis in schools that adopted the RSE curriculum early shows that challenges were encountered when developing and delivering the RSE curriculum.<sup>50</sup> According to a recent review of sexual abuse in schools in England, children were seldom positive about their RSE lessons and most felt the curriculum did not give them the information and advice they needed to navigate the reality of their lives.<sup>32</sup> This review also highlighted the normalisation of sexual harassment including online sexual abuse in schools for children and recommended that school leaders should take a whole-school approach (WSA) to create a culture where sexual harassment and online sexual abuse are not tolerated, and where they identify issues and intervene early to better protect their students. In light of these findings and in order to fulfil their statutory duties in relation to RSE, schools are likely to need to develop and implement programmes that address DRV and GBV prevention and improve responses to these outcomes.

Thus, there is a good rationale for a new systematic review that provides usable information for practitioners and policy-makers regarding the design and implementation of school-based interventions for DRV and GBV prevention. By synthesising which intervention characteristics are most important for preventing DRV and GBV, what kinds of issues implementers are likely to face, and how interventions are most likely to function in local contexts can support local decision-making and commissioning. Policy and practice stakeholders and young people consulted in preparation for this review identified the need for evidence that could be used to select, develop and implement locally relevant interventions. For example, the statutory guidance on RSE noted that in secondary school, students should understand 'how stereotypes, in particular stereotypes based on sex, gender, race, religion, sexual orientation or

disability, can cause damage (e.g. how they might normalise non-consensual behaviour or encourage prejudice) and 'what constitutes sexual harassment and sexual violence and why these are always unacceptable'.<sup>51</sup> Despite this guidance, policy and practice stakeholders noted a lack of clarity or clear evidence as to the best ways to achieve these goals. These stakeholders also noted the importance of including sexuality-based bullying as part of GBV, especially in light of the Government Equalities Office campaign against homophobic, transphobic and biphobic bullying.

From a research perspective, syntheses that go beyond intervention effectiveness can shape forward development and implementation of best bet interventions against developing evidence of heterogeneity in effectiveness.<sup>43</sup> This review is especially timely given our experience with the recently concluded National Institute for Health and Care Research (NIHR)-funded pilot trial of Project Respect, a school-based DRV prevention intervention. The first UK trial of an intervention to prevent DRV in young people, Project Respect,<sup>35</sup> initially intended to adapt two related interventions found to be effective in the USA: Safe Dates<sup>36</sup> and Shifting Boundaries.<sup>40</sup> The adaptation and optimisation process suggested that it would be unlikely to yield an effective intervention in UK school settings. Thus, a systematic review that 'deconstructs' existing interventions to better understand intervention functioning would be an essential starting point to develop a relevant evidence base on effective strategies for DRV and GBV prevention specifically in UK school contexts. Indeed, a key finding of Project Respect was that reconsidering types and combinations of components would be critical before undertaking further intervention development.

Finally, the lack of consideration of health inequalities in the evidence base is a major gap as yet unaddressed in a systematic review. This information is vitally important to prevent implementation of interventions that may exacerbate inequalities in respect of a problem already marked by exceptionally large social gradients in long-term impact.

In sum, this mixed-method systematic review including both DRV and GBV as eligible outcomes aims to:

- address inconsistencies across prior systematic reviews;
- ameliorate gaps in the understanding of GBV alongside DRV;
- consider the evaluability of the evidence base across multiple types of evidence;
- generate timely, relevant and innovative evidence for policy, practice and research.

## Review aim and objectives

Our overarching aim is to understand, via systematic review, the functions and effectiveness of school-based interventions for the prevention of DRV and GBV. This aim is supported by the following research questions (RQs):

- RQ1. What are the theories of change and components of evaluated interventions?
- RQ2. What factors affect the implementation of evaluated interventions?
- RQ3. Are interventions effective and cost-effective in preventing DRV and GBV and reducing social inequalities in these outcomes?
- RQ4. Based on the findings of RQs 1–3, what factors are important for determining the joint effectiveness of interventions for DRV and GBV outcomes?
- RQ5. What is the comparative effectiveness of different approaches to DRV and GBV prevention?
- RQ6. What do the different sources of evidence suggest about intervention mechanisms and how these are contingent on context?

Research questions 1–5 are addressed in the main results chapters (see [Chapters 4–8](#)), while RQ6 is addressed at the start of the discussion as part of a realist integration of findings (see [Chapter 9](#)).



## Chapter 2 Review methods

This chapter contains extracts of text that have been reproduced from the review protocol under the Creative Commons licence.

### Research design overview

We undertook a mixed-method systematic review in which different types of evidence relating to school-based interventions for the prevention of DRV and GBV were synthesised to understand if, how and in what ways these interventions are effective. Our systematic review followed best-practice conduct (York Centre for Reviews and Dissemination;<sup>52</sup> Cochrane Handbook)<sup>53</sup> and reporting guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analysis,<sup>54</sup> including extensions relating to equity; Enhancing Transparency in Reporting the Synthesis of Qualitative Research).<sup>55</sup> This protocol is registered in PROSPERO. Clarification and amendments to the review protocol are provided in [Tables 32](#) and [33](#) in [Appendix 1](#).

### Inclusion criteria for the review

#### Types of population

We included studies with children in compulsory education (e.g. aged 5–18 years) who are attending school.

#### Types of intervention

Guided by our logic model, we included evidence relating to interventions implemented in school contexts with students separate from, or as part of, RSE. These interventions could include one or more of:

- individual behavioural intervention (e.g. individual learning modules or apps);
- group or classroom-based intervention or practices (e.g. as part of RSE; delivering DRV and GBV prevention content in other academic sessions;<sup>44</sup> delivery of content in groups during school hours);
- network-based approaches, such as public opinion leader interventions;
- staff training and other service provision in schools (e.g. to recognise and respond better to sexual violence<sup>34</sup>);
- local and school policy change<sup>56</sup> to address structural factors relating to DRV or GBV, or to change school responses to DRV or GBV.

Interventions could be single-component or multicomponent, or implemented the same type of approach (e.g. group or classroom-based intervention) in a range of ways, for example by differentiating instruction over a range of school years. Included interventions focused in whole or in part on DRV and GBV, and could be universal, selective or indicated; could be primary prevention (reducing incidence of DRV and GBV) or secondary prevention (improving responses to DRV and GBV); and could focus on gender-specific groups (e.g. boys or girls only).

We excluded interventions that:

- did not seek to address DRV and GBV outcomes, for example interventions focusing on another health promotion topic, such as healthy eating, with an ‘opportunistic’ effect on DRV or GBV outcomes, but that did not describe prevention of DRV or GBV in intervention descriptions;
- were not delivered in compulsory education (e.g. university-based sexual violence prevention, or youth services);
- were not delivered at least in part in school contexts.

### *Types of control*

Comparators included business as usual, waitlist control or another active intervention.

### *Types of outcome*

We included outcomes relating to the full scope of DRV and GBV behaviours. These included:

- DRV perpetration or victimisation, including physical violence; emotional violence, including isolation; coercive control, including internet-mediated DRV; sexual assault in the context of relationships;
- GBV perpetration or victimisation, including harassment and bullying on the basis of gender or sexuality, including homophobic and transphobic bullying; internet-mediated GBV, such as unwanted sexting or forwarding of sexts; unwanted sexual contact, such as groping; sexual assault; sexual harassment and rape;
- DRV and GBV-related behaviours, such as harm reduction behaviours, help-seeking behaviours and bystander behaviours;
- knowledge and attitudes related to DRV and GBV, such as rape myth acceptance, bystander attitudes and GBV-condoning norms.

Outcomes included self-reported behaviours or experiences (e.g. were you groped in the last year; did you call someone names because of their sex or because you thought they were gay), teacher-reported behaviours (e.g. how many times did you see students engaging in sexual harassment) or official reports (e.g. how many sexual harassment incidents were reported in the last year). Included outcome measures were quantitative, expressed as categorical, continuous or count measures. Measures could be composite items (a range of DRV behaviours collected as a count of behaviours) or may be behaviour-specific. Behavioural outcomes could focus on behaviours over a specific period; frequency (monthly, weekly or daily); the number of episodes of a behaviour; or an index constructed from multiple measures. Economic analyses could also include health-related quality of life. We excluded knowledge and attitude outcomes relating to gender, or violence norms generally.

We did not include evaluations where outcomes related only to honour-based violence, forced marriage or female genital mutilation as these outcomes are potentially less amenable to curriculum-based school-based interventions.

### *Types of study*

Types of study included are categorised by research question.

- For RQ1, we used intervention descriptions and descriptions of theories of change across all included evidence.
- For RQ2, we drew on process and implementation evidence from eligible interventions that examined intervention delivery or receipt and how delivery was influenced by provider, user or context characteristics. This evidence could be reported in randomised or single-arm trials, retrospective or prospective evaluation studies, and evidence could be qualitative (e.g. description of acceptability of interventions) or quantitative (e.g. measurement of intervention fidelity).
- For RQ3 and RQ5, we drew on randomised trials, including cluster trials. We also sought any economic evaluations or modelling studies linked to these trials; that is, evidence that seeks to relate intervention costs and savings to health and well-being outcomes or benefits. Finally, we included moderation or subgroup analyses linked to these trials that explore equity-relevant characteristics.
- RQ4 and RQ6 drew on all included evidence.

## **Search methods for the identification of studies**

### *Database search strategy*

In July 2020, the following bibliographic databases were searched from inception and without limitation on date, language or publication type.



- MEDLINE, EMBASE, PsycINFO, Social Policy and Practice (Ovid).
- Cumulative Index to Nursing and Allied Health Literature, Education Resources Information Center, British Education Index, Education Research Complete, EconLit, Criminal Justice Abstracts (EBSCO).
- Cochrane Database of Systematic Reviews and the Cochrane Central Register of Controlled Trials.
- NHS Economic Evaluation Database (via the Centre for Reviews and Dissemination).
- Social Science Citation Index and Conference Proceedings Citation Index (Web of Science, Clarivate Analytics).
- Australian Education Index, ProQuest Dissertations and Theses Global, Sociological Abstracts including Social Services Abstracts, Applied Social Sciences Index and Abstracts (ProQuest).
- Trials Register of Promoting Health Interventions and Bibliomap [Evidence for Policy and Practice Information (EPPI)-Centre].
- Campbell Systematic Reviews (Campbell Collaboration).

The search strategies were developed by an experienced information specialist following extensive scoping searches and consideration of existing systematic review search strategies.<sup>6,8</sup> The Ovid MEDLINE search strategy was peer-reviewed by another experienced information specialist. The search strategy includes both free-text terms and subject headings for the school setting and DRV/GBV outcomes. In order to identify studies to answer review questions regarding outcome and economic evaluations, intervention theory, process and implementation evidence, and mediation and moderation evidence, no filters for specific study designs were applied. Searches were not restricted by date or language of publication. Full search strategies for each bibliographic database are available in [Report Supplementary Material 1](#).

We searched trial registers to identify ongoing or unpublished research [[www.clinicaltrials.gov](http://www.clinicaltrials.gov), WHO International Clinical Trials Registry Platform (ICTRP)] and we conducted searches for grey literature including conference abstracts, reports and theses from web searches, as well as searches of websites identified in initial scoping searches (including VAWnet; [www.vawnet.org](http://www.vawnet.org)).

The bibliographic database searches were updated in June 2021, with a revised strategy developed to improve precision. The revised strategy was based on analysis of titles, abstracts and index terms of included studies. The updated search strategy also incorporated programme names not identified in the initial scoping searches. We created a search summary table for all included studies to identify database sources that retrieved unique records, and this informed bibliographic database selection for update searches. Results from the update searches were not date limited, but instead de-duplicated against previous result sets to ensure records were not missed. Full search strategies for the update searches are available in [Report Supplementary Material 1](#).

### **Search strategy for other literature sources**

Supplementary search methods were used to identify studies not captured by our sensitive database strategies. The reference lists of existing systematic reviews or relevant reports were reviewed for relevant literature. This was the core of our cluster-based approach<sup>57</sup> to capture theory underpinning evaluated interventions and identify any missed 'sibling' studies. Forward and backward citation chasing was conducted on included studies identified from the June 2020 bibliographic database searches. Scopus (Elsevier), Web of Science (Clarivate) and Google Scholar were used for citation chasing, and bibliographies of included studies were manually checked where this information was incomplete on Web of Science and Scopus.

Targeted searches were conducted in Web of Science and Scopus using first and last author names for studies identified in bibliographic database searches in June 2020. Specific project names (e.g. Project Respect, Shifting Boundaries or Safe Dates) were included in the update search strategies in bibliographic databases, and Google Scholar. Results were screened in Google Scholar, with the first 200

records scanned for each search string. Websites identified in initial scoping searches were browsed or searched for additional reports (including VAWnet: [www.vawnet.org](http://www.vawnet.org); USAID: [www.usaid.gov](http://www.usaid.gov); AVA – Against Violence and Abuse; UNGEI, National Criminal Justice Reference Service: [www.ncjrs.gov](http://www.ncjrs.gov)). Full details of Google Scholar and website searches are available in [Report Supplementary Material 1](#).

Finally, we planned to hand-search journals that published included studies which we found only via reference checking and which are not indexed on databases we have searched. However, due to the exhaustive cluster-based search methods, this was unnecessary.

### Information management and study selection

Search results were downloaded into EndNote for deduplication. This was done using EndNote deduplication functionality, plus manual checking. Covidence provided further duplicate matching on import.

Following deduplication, a single search file was uploaded to Covidence software. Two reviewers piloted the screening of successive batches of 100 titles/abstracts, meeting to discuss disagreements, calling on a third reviewer where necessary. Once 90% agreement was reached; each title and abstract were reviewed independently and in duplicate. Records retained after this stage were accessed in full text and assessed against the inclusion criteria in duplicate, and assigned to one or more evidence types (implementation/process, outcome, economic evaluation, mediation and moderation).

### Data extraction and appraisal

One reviewer undertook data extraction on study characteristics independently using standardised, piloted forms. All data extraction on study characteristics was comprehensively quality assured by two reviewers for completeness and accuracy.

Further detail about the data extraction tool (DET) is provided in [Appendix 2](#). For all studies where relevant, we extracted information on basic study details (study location, timing and duration; individual and organisational participant characteristics); study design and methods (design, sampling and sample size, allocation, blinding, control of confounding, accounting for data clustering, data collection, attrition, analysis); process evaluation findings and interpretation; outcome measures [timing, reliability of measures, intraclass correlation coefficients (ICCs), effect sizes]; relevant mediation and moderation analyses; and economic data (inputs and outputs relating to costs, consequences/benefits, disaggregated by time period where appropriate). Intervention descriptions and theories of change were extracted as free text across included evidence. When extracting theories of change, we focused on constructs, mechanisms and any contextual contingencies affecting these, as well as other theories cited. Reviewers entered data into a Microsoft Excel spreadsheet. If included studies were reported in languages that cannot be translated by the review team, a review author completed data extraction in conjunction with a translation of the article.

Published reports may be incomplete in a wide range of ways. For example, they may not: present information on all the outcomes that were measured (possibly resulting in outcome reporting bias); provide sufficient information about the intervention for accurate characterisation; or report statistical information necessary for the calculation of effect sizes. In all cases where there was a danger of missing data affecting our analysis, we contacted authors of papers wherever possible to request additional information. If authors were not traceable, or sought information was unavailable from the authors within 2 months of contacting them, we recorded that the study information was missing on the data extraction form, and this was reflected in our risk of bias assessment for the study.<sup>58</sup>

## Assessment of quality and risk of bias

Two reviewers assessed the quality of each empirical report. The two reviewers then met to compare their assessments, resolving any differences through discussion and, where necessary, by calling on a third reviewer.

### Assessment of process evaluations

Process evaluations were appraised using the EPPI-Centre tool.<sup>59</sup> This addressed the rigour of sampling; data collection; data analysis; the extent to which the study findings are grounded in the data; whether the study privileges the perspectives of participants; the breadth of findings; and depth of findings. These assessments were then used to assign studies to two categories of 'weight of evidence' (low, medium or high) to rate the reliability or trustworthiness of the findings (the extent to which the methods employed were rigorous/could minimise bias and error in the findings), and to rate the usefulness of the findings for shedding light on factors relating to the research questions. Guidance was provided to reviewers to help them reach an assessment on each criterion and the final weight of evidence. Findings from critical appraisal were used to inform synthesis, including by describing the qualitative strength of findings in evidence syntheses. The tool was used in this review to assess process evaluations regardless of whether these drew on qualitative or quantitative data, because the criteria were judged applicable to both.

We chose the EPPI-Centre tool given our prior experience with this tool in other NIHR-funded systematic reviews of qualitative research.<sup>39,44,59</sup> A key strength of this tool is that it generates appraisal in terms of both study relevance and study trustworthiness, both of which are important in qualitative evidence synthesis. In addition, the tool reflects the degree to which qualitative findings privilege participant voices, which is important given our focus on generating evidence that speaks to local implementers' needs.

A final step in quality assessment was to assign studies two types of 'weight of evidence'. First, reviewers assigned a weight (i.e. low, medium or high) to rate the reliability or trustworthiness of the findings (i.e. the extent to which the methods employed were rigorous/could minimise bias and error in the findings). Second, reviewers assigned an additional weight (i.e. low, medium or high) to rate the usefulness of the findings for shedding light on factors relating to the review questions. The criteria guiding these decisions are shown in [Table 1](#). Quality was used to determine the qualitative weight given to findings in our synthesis, with none of the themes represented solely by studies judged as low on both dimensions. This is relevant because our experience with systematic reviews of process evaluations is that studies relating to exceptionally relevant interventions may not provide rich data; conversely, process evaluations of interventions that only address knowledge and attitudes, rather than behaviours, may provide especially meaningful data that illuminate both how interventions function and how interventions were implemented. Thus, distinguishing between these two weights of evidence sensitised our analysis and helped in prioritising most relevant and most trustworthy findings in synthesis.

**TABLE 1** Criteria used to derive overall quality ratings of process evaluations

	Reliability/trustworthiness of the data	Usefulness of the data
High	Rigour in at least three of the first four criteria	Studies privileged the perspectives of participants and to present findings that achieved both breadth and depth
Medium	Rigour in two out of the first four criteria	Studies partially met criteria for privileged perspectives of participants and presenting findings that achieved both breadth and depth
Low	Rigour in one or none of the first four criteria	Sufficient but limited findings

### **Assessment of outcome evaluations**

Critical appraisal of the cluster randomised controlled trials (cRCTs) and RCTs was conducted using the Cochrane Risk of Bias Tool version II.<sup>60</sup> For each study, reviewers judged the likelihood of bias for sequence generation and allocation concealment; blinding (of participants, personnel, or outcome assessors); incomplete outcome data; selective outcome reporting; other sources of bias (e.g. recruitment bias in cluster-randomised studies); and intensity/type of comparator. Each study was subsequently identified as 'high risk', 'low risk' or 'unclear risk' within each domain. Decisions were guided by the guidance associated with the tool,<sup>53</sup> with the following review-specific adaptations:

- Where trials did not report baseline characteristics of the included clusters or arms, it was assumed that there were no confounding factors unless the reviewer was concerned that the sample size would undermine effective randomisation.
- It was assumed that students would likely guess the aims of the intervention, unless evidence to the contrary.
- There appeared to be a low rate of switching in the trials, but there was limited information about the risk for other protocol deviations that may lead to performance bias. The following were not downgraded for performance bias if there was no switching and an intention to treat (ITT) analysis was reported, as the risk of contamination was considered low:
  - cluster trials where schools were the unit of randomisation
  - interventions lasting one class only.
- An absolute participant attrition rate > 30% was considered problematic.
- A scale outcomes measure was judged to be acceptable where items were considered by the reviewer to have face validity, and if internal reliability (generally the only psychometric outcome reported) was  $\geq 0.7$ .
- Trials were not downgraded for unblinded outcome assessors within the outcome measurement domain. Due to the nature of the interventions evaluated, it was not possible for the vast majority of trials to fully blind or obscure students from the intervention they were receiving. This was a limitation noted by many trial authors, who considered there to be a risk that outcomes completed by students could be subject to desirability bias, where students complete outcomes in a way they think is desired by researchers. Trials were not downgraded for this to avoid a floor effect in critical appraisal ratings across the evidence base. However, all trials included in the review are nevertheless considered to be at risk of this bias, in addition to the overall risk of bias across the other tool domains.

Finally, we assessed reporting bias in trials according to Cochrane Handbook guidance.<sup>53</sup> We reduced the effect of reporting bias by focusing synthesis on studies rather than publications and avoiding duplicated data. We attempted to detect duplicate studies and, if multiple articles report on the same study, we extracted data only once. We prevented location bias by searching across multiple databases. We minimised language bias by not excluding any article based on language.

### **Assessment of economic studies**

No formal health economic evaluations were identified by the review. Were we to have identified any eligible economic evaluations and modelling studies linked to randomised trials, we would have appraised them using the Drummond<sup>61</sup> or Philips<sup>62</sup> checklists, respectively. These checklists require the analyst to answer 24 questions regarding each study, ranging from the type of economic evaluation (e.g. cost-utility analysis) to the time horizon and rationale for the choice of modelling approach.

## **Data analysis**

Our analysis proceeded in a convergent design with RQs 1–5 informing RQ6.

### **Research question 1: theories and components**

We used intervention components analysis<sup>63</sup> to analyse intervention descriptions across all included evidence. Intervention components analysis is an inductive approach to comprehensively describing and categorising intervention components in a target body of evidence. This is an appropriate method to describe intervention components when these components do not fit into pre-existing taxonomies of behaviour change, which is especially the case in this review given the diverse, contextually situated and frequently multilevel nature of included interventions. Two reviewers used open coding to generate a comprehensive list of possible intervention descriptors from five different intervention descriptions. The two lists were compared and combined. Using principles of axial coding, the two reviewers proceeded through the remaining intervention descriptions, collapsing codes and adding new ones as required and meeting periodically to compare codes, determine if new axial codes were required and organise axial codes into categories. The final result was a comprehensive list of descriptors to characterise included interventions, organised by relevant categories. Our intervention components analysis ultimately included a set of descriptors for included interventions (type, level, focus), which was analysed across all included evidence, and a set of descriptors for intervention components specifically, which was applied only to outcome evaluations given the size of the evidence based and differences in depth of intervention description.

We also synthesised theories of change based on intervention descriptions and accounts embedded in individual studies' theories of change. Drawing on methods used in previous theory syntheses,<sup>64</sup> this inductive analysis included a lines-of-argument synthesis.<sup>65</sup> Lines-of-argument synthesis is an appropriate method based on its understanding that each included study examines a 'part of the whole'; that is, each study's account of theory of change represents one possible part of how school-based interventions can work to reduce DRV and GBV broadly. Analysis was undertaken using thematic synthesis and line-by-line coding.<sup>66</sup> We intended to analyse interventions within type first; however, this was not done as it would have limited the breadth of interventions analysed in initial coding, and ultimately did not generate any analytical advantages. Two reviewers began the theory synthesis with a subset of studies first to understand and agree a common approach, prioritising a selection of interventions that (a) were multiply evaluated and (b) represented the breadth of interventions. Subsequently, both reviewers proceeded through each set of studies, meeting to agree findings. The findings from each intervention (second-order constructs) were then compared against each other using a 'higher-level' iteration of these thematic grids to produce an overarching theory of change (third-order constructs). Coding unfolded over several cycles, proceeding from initial coding to saturation, followed by confirmatory coding on remaining papers in the analysis. The outcome of this synthesis was a hierarchically organised account of the mechanisms by which interventions are theorised to function in reducing and preventing DRV and GBV and the contexts in which these mechanisms are most likely to operate, broadly across interventions and specifically within intervention types. Specific methods for how the analysis 'unfolded' are presented in [Chapter 6](#).

For both the intervention components analysis and the synthesis of theories of change, two reviewers undertook analysis in parallel, with periodic meetings among the analysis team to validate the developing framework. We additionally checked findings with our advisory and stakeholder groups for face validity. These checks focused on identifying which intervention components are most promising in the UK context on the basis of relevance and fit with existing school curricula, intervention strategies and cultures, and which contexts and mechanisms are most likely to be relevant to the UK educational system.

### **Research question 2: implementation and process evaluations**

We synthesised data from implementation and process evidence to identify salient factors affecting intervention implementation and how these relate to context. We used thematic synthesis<sup>67</sup> to synthesise these data. Thematic synthesis is apposite as it is a flexible method that seeks to understand findings across studies, without necessarily imposing a theoretical framework or 'theorising' the data, while still preserving the value of reciprocal translation<sup>65</sup> in understanding patterns

of findings across studies rather than merely summarising these. Thematic synthesis includes both descriptive, or *in vivo*, themes that describe the specific content in included studies, and analytical themes that cut across findings from multiple studies.

Synthesis of implementation and process evidence began with a tabulation and descriptive coding phase. First, reviewers created a table of included studies reporting information on study type, methods, context and sample; interventions evaluated; and summarised findings. Two reviewers working in parallel then undertook pilot analysis of five reports that were appraised as being of high quality and relevance. The reviewers read and reread the results from these reports, applying line-by-line codes to capture the content of the data, and drafted memos explaining these codes. Coding then proceeded with descriptive codes which closely reflect the words used in findings sections. The reviewers then grouped and organised codes, applying analytical codes reflecting higher-order themes. Subsequently, the reviewers coded remaining studies drawing on the agreed set of codes but developing new descriptive and analytical codes as these arose from the analytical process, and again writing memos to explain these codes. At the end of this process, the two reviewers met to compare their sets of codes and memos. They identified commonalities, differences of emphasis and contradictions with the aim of developing an overall analysis. The outcome of this synthesis was an account of barriers and facilitators that intervention implementers could match to their own context. We sensitivity-analysed all findings by considering whether findings relate to high-income country (HIC) contexts or low-income and middle-income country (LMIC) contexts, in order to better understand the applicability of findings to the UK context.

In addition to routine auditing of findings by the investigator team, we presented findings to our advisory and youth stakeholder groups for feedback. The goal of this was to ensure that findings are relevant to intervention implementers in the UK, and to identify which findings are especially salient in the UK education context.

### ***Research question 3: intervention effectiveness, moderation, mediation and economics***

Included trials were organised by intervention type as informed by findings from RQ1, and relevant outcomes were hierarchically categorised by type of DRV or GBV behaviour or experience. We distinguished between different intervention follow-up times (up to 1 year from baseline; more than 1 year from baseline). All time points reported by a study were extracted.

We then meta-analysed study findings where possible. We first analysed within intervention and time-point categories to estimate specific intervention effects on broad outcome categories and then meta-analysed across intervention categories and across outcome categories. Meta-analyses then considered different behaviours, experiences and forms of perpetration and victimisation separately across all intervention types. We originally sought to meta-analyse by intervention type, specific outcome (e.g. physical DRV victimisation) and time point but found that this would be inappropriate given the small number of studies included in any one meta-analysis.

We sensitivity-analysed meta-analyses by whether evidence arose from high-income settings or low-income or middle-income settings, using a standard subgroup difference-based test to identify any statistical differences in effectiveness that may shape the relevance of overall results to the UK context.

The key metric for all meta-analyses was defined in the protocol to be odds ratios (ORs); we ultimately used a standardised mean difference to analyse knowledge and attitude outcomes (i.e. not victimisation or perpetration outcomes). Where victimisation and perpetration outcome measures were continuous, these were converted to ORs using the logistic transformation. We used a random-effects robust variance estimation meta-analysis model to synthesise effect sizes. This is because many outcome evaluations included multiple measures of conceptually related outcomes. Robust variance estimation meta-analysis improves on previous strategies for dealing with multiple relevant effect sizes per study,

such as multilevel meta-analysis, meta-analysing within studies or choosing one effect size, by including all relevant effect sizes but adjusting for inter-dependencies within studies.<sup>68</sup> Unlike multivariate meta-analysis, it does not require the variance–covariance matrix of included effect sizes to be known. Where meta-analyses were performed, we included pooled effect sizes in forest plots, with the individual study point estimates weighted by a function of their precision.

We checked that cluster randomised trials have accounted for unit of analysis issues. Prior to synthesis, we checked for correct analysis (where appropriate) by cluster and reported values of ICCs, cluster size, data for all participants or effect estimates and standard errors (SEs). Where proper account was not taken of data clustering, we corrected for this by inflating the SE by the square root of the design effect.<sup>58</sup> Where ICCs were not reported and were relevant to estimating an intervention effect, we imputed an ICC of 0.05, based on values reported in other studies. We planned to undertake sensitivity analysis with respect to ICC values, but this was ultimately unnecessary given that the ICC value was 'known' from included studies.

We then produced a narrative account of intervention effectiveness. This narrative account included a comprehensive table of included studies reporting information on study design, methods, context and sample, interventions evaluated and summarised findings, alongside forest plots to describe visually the range and precision of estimates of intervention effectiveness. We then examined the extent of heterogeneity among the studies (as determined both by a Cochran's Q test and by inspection of the  $I^2$ ). If an indication of substantial heterogeneity was determined (e.g. study-level  $I^2$  value > 50%) that cannot be explained through metaregressions, we investigated this further using subgroup and sensitivity analyses.

We used the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach as described in the Cochrane Handbook for Systematic Reviews of Interventions<sup>53</sup> to present the quality of evidence, focusing on primary victimisation and perpetration outcomes. The downgrading of the quality of a body of evidence for a specific outcome was based on five factors: limitations of study; indirectness of evidence; inconsistency of results; precision of results; and publication bias. The GRADE approach specifies four levels of quality (high, moderate, low and very low). If sufficient studies are found, we drew funnel plots to assess the presence of possible publication bias (trial effect vs. SE). While funnel plot asymmetry may indicate publication bias, this can be misleading with a small number of studies.<sup>69</sup> We considered possible explanations for any asymmetry in the review in light of our number of included studies. We assessed the impact of risk of bias in the included studies via restricting analyses to studies deemed to be at low risk of selection bias, performance bias and attrition bias; however, risk of bias in included studies was generally poor on these domains and thus a restricted analysis would have been uninformative.

Next, we planned to produce a narrative account of findings from economic evaluations by intervention type. However, as we did not identify any relevant economic evaluations, we instead provided a narrative summary of costing studies linked to included randomised trials. Measures of costs and indirect resource use were summarised using tables. These data were used to inform a narrative synthesis of economic analyses and applicability to the UK context. Because no full economic evaluations were identified, we did not appraise economic studies.

Subsequently, we narratively synthesised evidence of mediation of intervention impacts to demonstrate the degree to which DRV and GBV victimisation outcomes in interventions are mediated by intervening variables.

As a final synthesis step, we assessed equity effects by drawing on moderation analyses to illustrate how interventions impact health inequalities, focusing on ethnicity, socioeconomic position, gender, sexuality and age. Our analysis focused on priority outcomes of DRV and GBV victimisation and perpetration. Moderation analyses considering differential intervention effectiveness on these equity-relevant characteristics were organised by intervention type and outcome category and narratively

synthesised. Harvest plots were used to graphically depict how interventions ameliorate or worsen social gradients on specific outcomes by equity-relevant characteristics.<sup>70</sup> Where possible and where sufficient data existed, we extended our random-effects robust variance estimation meta-analyses using metaregression to estimate how equity-relevant characteristics of study populations relate to intervention effectiveness.

We presented the findings from meta-analyses to stakeholders in our advisory and youth groups to understand which interventions are 'best bets' in the UK context based on both effectiveness and relevance to the UK educational system.

### ***Research question 4: metaregression and qualitative comparative analysis***

The next step of the synthesis integrated findings across RQ1–3 to identify how components, mechanisms and implementation factors relate to effectiveness across DRV and GBV outcomes.

First, we used metaregression and pattern matching to understand which components are most effective across both DRV and GBV. We revisited meta-analyses undertaken in RQ3 and used components identified in RQ1 to estimate how inclusion or exclusion of a specific component is associated with intervention effectiveness. These analyses were undertaken by outcome type and time point, but not intervention type, and were estimated using random-effects robust variance estimation meta-analyses. We considered the magnitude, precision and direction of the regression coefficient associated with presence of an intervention component and the reduction in between-study variance resulting from using an intervention component as a metaregressor. We examined the same metaregressor across a range of outcomes and used pattern-matching<sup>71</sup> to identify how specific components are consistently associated with improved or reduced effectiveness across a range of DRV and GBV experiences and behaviours.

Second and where appropriate, we used qualitative comparative analysis (QCA) to identify how different implementation and intervention characteristics combine to form pathways to effectiveness, including on health inequalities. QCA focuses on understanding configurations of conditions that form pathways to effectiveness in interventions.<sup>72</sup> Given that interventions are likely to report diverse types of outcomes using different measurement approaches, QCA is especially apposite as it transforms numerical estimates into a calibrated measure of whether an intervention is effective. QCA focuses on how different implementation and intervention characteristics act in concert to 'unlock' pathways to effectiveness. Informed by findings from RQ1 to 3, we developed several candidate groups of implementation and intervention characteristics and considered how these characteristics form pathways to effectiveness. We coded included outcome evaluations as to the presence or absence of these implementation and intervention characteristics where this has not been done already, and classified outcome evaluations as to their effectiveness. We planned to describe effectiveness in several ways, each corresponding to a separate model:

1. interventions that are effective in reducing DRV versus interventions that are not effective in reducing DRV;
2. interventions that are effective in reducing GBV versus interventions that are not effective in reducing GBV;
3. interventions that are jointly effective on DRV and GBV versus interventions that are only effective in one or neither domain.

However, the overlap in interventions testing both DRV and GBV outcomes was insufficient to proceed with this analysis.

We generated truth tables to understand combinations of characteristics by each outcome, and sought to resolve any contradictory configurations (i.e. where combinations of characteristics span both effective and ineffective interventions). We used Boolean minimisation to identify common pathways



to effectiveness across each of the three models, and compared findings across models to note which pathways are specific to DRV or GBV outcomes and which pathways are common to effectiveness of DRV and GBV outcomes, and thus should be considered when seeking to address both types of outcomes jointly.

We presented the findings of this analysis to the advisory and youth stakeholder groups to consider how best bet components are relevant in the UK context, and how identified pathways from QCA may be especially important in shaping forward evaluation and implementation in UK schools.

### **Research question 5: network meta-analysis**

Drawing on randomised trials analysed in RQ3, we used the findings from our analysis of intervention typology and intervention components to estimate network meta-analyses (NMAs). For all NMAs, we considered as separate outcomes DRV victimisation, DRV perpetration, GBV victimisation, and GBV perpetration. We considered our intervention typology analyses as primary and our intervention component analyses as exploratory.

All NMAs were implemented in a frequentist framework drawing on methods proposed by Rücker *et al.* (2012)<sup>73</sup> and Melendez-Torres *et al.* (2015),<sup>74</sup> using random effects for all analyses. Implementation used – network – in Stata for typology NMAs and – netmeta – in R for component NMAs. Where trials reported multiple effect sizes for the same outcome (i.e. different estimates of DRV victimisation drawing on, e.g. emotional, physical, sexual DRV), we assumed that outcomes were correlated with  $\rho = 0.8$ , consistent with standard assumptions used in robust variance estimation meta-analyses. Analyses were checked for inconsistency statistically using standard design-by-treatment interaction models, and conceptually by considering transitivity with regard to key effect modifiers (principally age, sex, high-income vs. low- and middle-income context). Where necessary, we used network meta-regression to account for plausible effect modifiers unevenly distributed between network comparisons. NMAs were then presented using the standard pairwise comparison league table, with intervention types ranked using a resampling method and surface under the cumulative ranking curve (SUCRA).

### **Feasibility**

We first assessed feasibility for these NMAs, by examining the distribution of typology and components across different outcome domains, and examining where network metaregression may be necessary to balance effect modifiers between nodes in the network. We regarded the typology analyses as primary and the component analyses as exploratory given the number of trials available for each outcome, presenting component NMAs only where enough trials were available to provide a sufficiently rich data structure.

### **Intervention typology network meta-analyses**

Drawing on our findings from RQ1, we used our proposed intervention typology to estimate the relative effectiveness of different intervention types.

### **Intervention component network meta-analyses**

Drawing on our findings from our first research question, we explored where feasible the use of high-level component categories to estimate the impact of each component on intervention effectiveness. We planned to probe potential two-way interactions between components first by probing cross-level (e.g. structural by student) interactions and then considering within-level interactions, but ultimately explored strictly additive models where component NMAs were undertaken.

### **Research question 6: realist integration of findings**

We used implementation and process evidence, mediator and moderator analyses, and 'informal' evidence from included studies (e.g. discussion of how interventions were implemented in outcome evaluation discussion sections) to identify mechanisms by which interventions impact DRV and GBV outcomes and associated contextual contingencies. Mechanisms focused on understanding causal

chains by which interventions are likely effective, including proximal and antecedent steps (e.g. increasing engagement, addressing social norms, knowledge and attitudes) leading to distal effects on outcomes. Informed by realist synthesis<sup>75</sup> and best-fit framework synthesis,<sup>76</sup> reviewers used findings from RQ1 as a framework to infer and induce mechanisms from studies. Best-fit framework synthesis is a flexible form of framework analysis for the synthesis of diverse study types. Importantly, it accounts for the possibility that previously 'untheorised' findings may emerge as relevant from included study findings. Our approach is informed by realist synthesis in that we aim to understand the links between contexts and mechanisms in generating outcome patterns.

Using the context-mechanism findings from the theory synthesis in RQ1 as a template, reviewers analysed findings from the subsequent syntheses in RQ1–5, working independently and together to understand which causal propositions from the theory synthesis are supported, refuted or unevidenced. This synthesis unfolded collaboratively through meetings of the reviewer team, with an audit trail kept of synthesis decisions made. Analysis could also generate additional context-mechanism-outcome configurations that were not present in the theory synthesis but that reflected learning from other syntheses.

Findings were periodically discussed in meetings of the investigator group and were presented to the advisory and youth stakeholder groups. These groups provided additional insights in revisiting contexts and mechanisms previously discussed in RQ1, and in identifying which previously untheorised contexts and mechanisms are especially salient in the UK context.

### **Policy and practice consultation**

Our review benefited from a robust process of consultation, described in *Chapter 11*. We convened a policy and practice advisory group, which met three times over the course of the project, and also undertook consultation with different groups of young people in London and Cardiff. These consultations drew on a range of formats and helped in sensitising our analyses and identifying implications of our work.

### **Ethics approval**

The research involved no human participants and drew solely on evidence in the public realm, so ethics approval was not required.

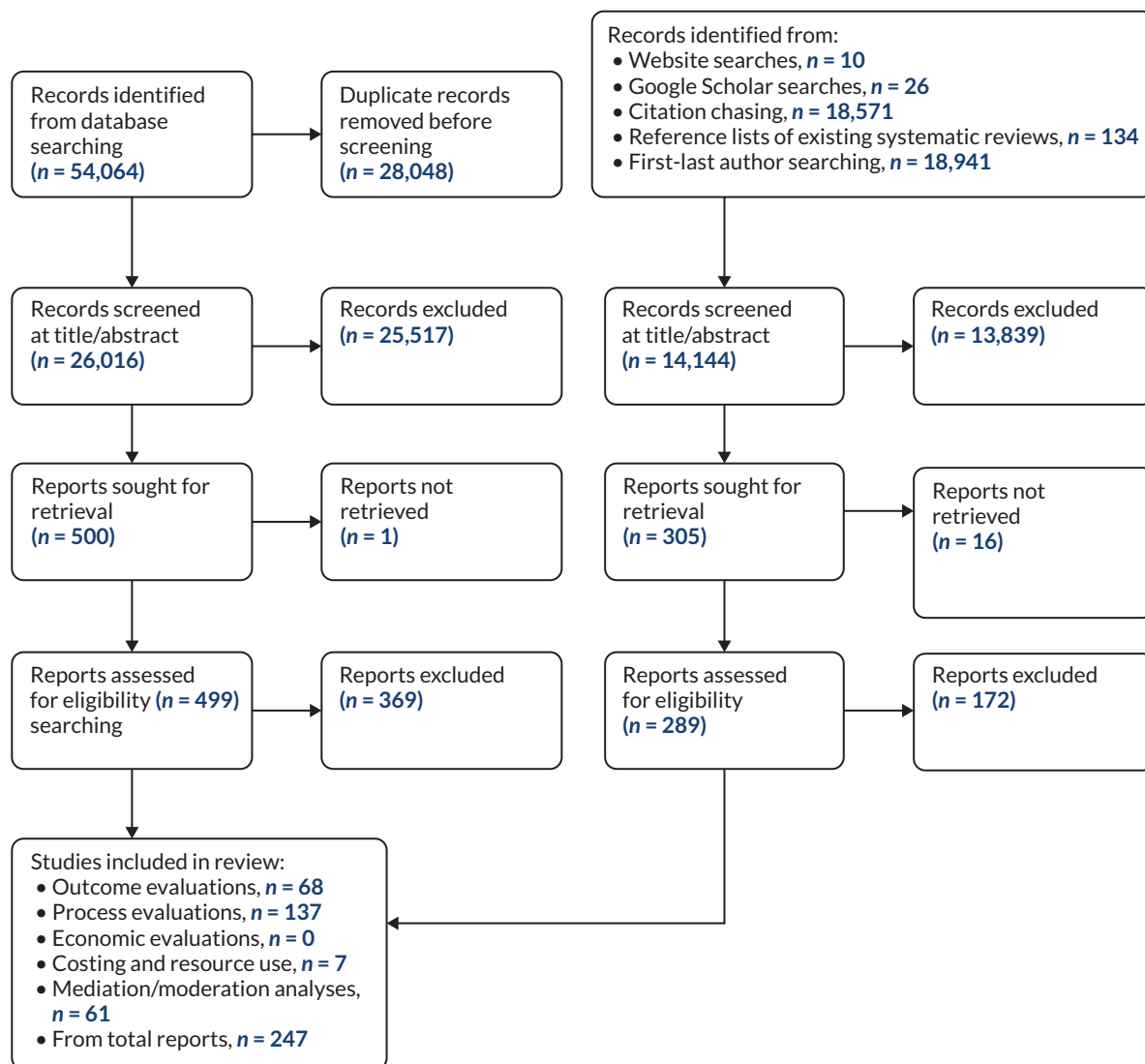
## Chapter 3 Overview of included studies

### Results of the search

Results of the search are depicted in [Figure 2](#). Up until June 2021, we included a total of 247 reports over 2 rounds of database searching, extensive citation chasing and first and last author searching. It was common for studies to be reported across multiple publications. Some studies were also included in more than one area of the review. An overview of studies included in the review and the type of evidence they reported is provided in [Appendix 2](#).

### Included studies and reports

The review identified 247 studies that evaluated 211 active interventions. An overview of the studies identified and the types of evidence they reported is provided in the appendix (see [Table 34](#) in [Appendix 2](#)), and can be summarised as follows:



**FIGURE 2** Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart.

- Fifty-one active interventions described across 75 publications were included in the theory synthesis.
- One hundred and sixty active interventions evaluated in 137 studies reported in 161 publications were included that reported process or implementation outcomes.
- Seventy-nine active interventions evaluated across 68 studies reported in 101 publications were included that reported effectiveness data.
- Twenty-five active interventions evaluated across 24 studies reported in 28 publications were included that reported moderation data.
- Five active interventions evaluated across five studies reported in seven publications were included that reported mediation data.
- No studies reporting a formal economic evaluation of a relevant intervention were identified. Instead, we identified eight active interventions evaluated across seven studies reported in seven publications that reported cost and resource use associated with interventions.

Excluded studies are detailed in [Report Supplementary Material 1](#).

## Study and intervention characteristics

### Study designs

#### Process and implementation evaluations (RQ2)

One hundred and thirty-seven studies reporting process or implementation outcomes were included. Twenty-nine (21.2%)<sup>49,77-101</sup> of the studies were associated with a RCT or cRCT, while others were non-randomised, naturalistic or quasi-experimental studies. Eighty-two studies reported quantitative data and 122 studies reported qualitative data.

Many of the included studies reported incomplete information about the sample included in process evaluations. Where sample size was reported or could be estimated, student sample size ranged from 6 to 5937 (median = 85). Studies evaluating interventions to target both DRV and GBV were generally larger (median = 307) than those targeting DRV only (median = 87) or GBV only (median = 44). Those studies including staff in their evaluations reported a sample size ranging from 1 to 1900 teaching staff or other professional stakeholders (median = 26).

Two studies (1.5%)<sup>102,103</sup> were conducted solely with children of primary school age (i.e. < 11 years), and an additional five studies (3.6%)<sup>104-108</sup> included students of primary school age or their teachers amongst a broader sample. The other studies were conducted in secondary/high school settings.

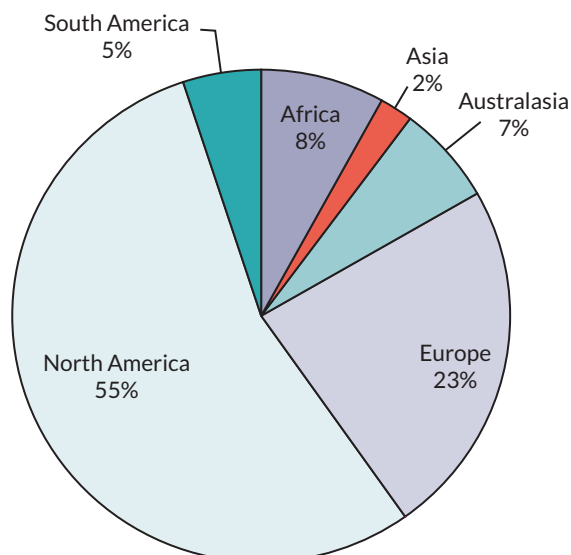
The majority of studies ( $n = 75$ , 54.7%) were conducted in North America, with the remaining split across Europe ( $n = 32$ , 23.4%),<sup>49,79,88,91,94,102,105-107,109-131</sup> Asia ( $n = 3$ , 2.2%),<sup>78,85,87</sup> Africa ( $n = 11$ , 8.0%),<sup>90,132-141</sup> South America ( $n = 7$ , 5.1%)<sup>84,104,142-146</sup> and Australasia ( $n = 9$ , 6.6%).<sup>103,147-154</sup> A breakdown of the specific continents and countries in which studies were conducted is provided in [Figure 3](#) and [Table 2](#).

#### Outcome evaluations (RQ3)

Of the 68 included studies, 14 were RCTs (20.6%) and 54 were cRCTs (79.4%). The majority of studies ( $n = 61$ , 89.7%) compared two intervention arms, though five (7.4%)<sup>78,94,155-157</sup> and two (2.9%)<sup>100,158</sup> studies evaluated three and four intervention arms, respectively. cRCTs ranged in size from 4 to 151 clusters (median 23).

Sample size (ITT population) of the studies ranged from 34 to 89,707 participants, including all groups (median 816). Trials targeting both DRV and GBV were generally larger (median = 2403) compared to trials targeting GBV only (median = 1306) or DRV only (median = 660).

School settings for the included studies are shown in [Table 3](#). Only one study (1.5%)<sup>159</sup> was conducted solely in a primary school setting; rather, the vast majority of included studies were conducted in secondary school settings.



**FIGURE 3** Location of process/implementation evaluations.

**TABLE 2** Continent/countries of included process evaluations

Location	Number of studies	% of total
<b>Africa</b>	<b>11</b>	<b>8.03</b>
Botswana	1	0.73
Botswana, Malawi and Mozambique	1	0.73
Cameroon, Senegal and Togo	1	0.73
Malawi	1	0.73
Namibia, South Africa, South Sudan, Eswatini, Tanzania, Zambia and Zimbabwe	1	0.73
South Africa	4	2.92
Tanzania	1	0.73
Zimbabwe	1	0.73
<b>Asia</b>	<b>3</b>	<b>2.19</b>
Bangladesh, India, Vietnam	1	0.73
India	1	0.73
Taiwan	1	0.73
<b>Australasia</b>	<b>9</b>	<b>6.57</b>
Australia	8	5.84
New Zealand	1	0.73
<b>Europe</b>	<b>32</b>	<b>23.36</b>
Bosnia and Herzegovina, Croatia, Serbia	1	0.73
England	10	7.30
England and Wales	1	0.73
Germany	1	0.73

continued

**TABLE 2** Continent/countries of included process evaluations (*continued*)

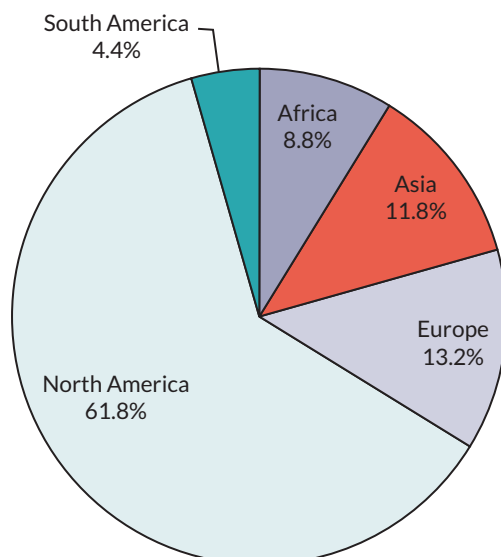
Location	Number of studies	% of total
Ireland	1	0.73
Netherlands	1	0.73
Portugal	1	0.73
Scotland	1	0.73
Serbia	1	0.73
Slovenia	1	0.73
Spain	7	5.11
Sweden	2	1.46
UK	3	2.19
UK, France, Spain and Malta	1	0.73
<b>North America</b>	<b>75</b>	<b>54.74</b>
Canada	14	10.22
Hawaii	2	1.46
Mexico	1	0.73
USA	58	42.34
<b>South America</b>	<b>7</b>	<b>5.11</b>
Brazil	5	3.65
Colombia	1	0.73
Peru	1	0.73

**TABLE 3** School settings of included outcome evaluations

School setting	Number of studies (%)
Primary school only (< 11 years)	1 (1.5) <sup>159</sup>
Primary and secondary school (≤ 16 years)	3 (4.4) <sup>160-162</sup>
Secondary school only (11–16 years)	57 (83.8)
Secondary school and sixth form (11–18 years)	4 (5.9) <sup>49,84,163,164</sup>
Sixth form only	2 (2.9) <sup>88,165</sup>
Unclear	1 (1.5) <sup>166</sup>

Where reported, recruitment or data collection for the included studies spanned 1994–2017; there was no clear difference in the age of studies evaluating DRV or GBV.

The majority of studies ( $n = 42$ , 61.8%) were conducted in North America, with the remaining split across Europe ( $n = 9$ , 13.2%),<sup>35,49,79,88,94,164,167-169</sup> Asia ( $n = 8$ , 11.8%),<sup>78,85,87,170-173</sup> Africa ( $n = 6$ , 8.8%)<sup>90,155,160,174-176</sup> and South America ( $n = 3$ , 4.4%).<sup>84,163</sup> A breakdown of the specific continents and countries in which studies were conducted is provided in [Figure 4](#) and [Table 4](#).



**FIGURE 4** Location of outcome evaluations.

**TABLE 4** Continent/country of included outcome evaluations

Location	Number of studies	% of total
<b>Africa</b>	6	8.82
Kenya	1	1.47
Malawi	1	1.47
South Africa	3	4.41
Uganda	1	1.47
<b>Asia</b>	8	11.76
India	2	2.94
Iran	2	2.94
South Korea	2	2.94
Taiwan	1	1.47
Vietnam	1	1.47
<b>Europe</b>	9	13.24
England	3	4.41
France	1	1.47
Germany	1	1.47
Spain	3	4.41
The Netherlands	1	1.47
<b>North America</b>	42	61.76
Barbados	1	1.47
El Salvador	1	1.47
Haiti	1	1.47
USA	39	57.35
<b>South America</b>	3	4.41
Brazil	3	4.41

## Intervention characteristics

### Process and implementation evaluations (RQ2)

One hundred and sixty active interventions were evaluated for process or implementation outcomes. A description of the interventions evaluated is provided in the appendix.

The majority ( $n = 104$ , 65.0%) of interventions were evaluated in single-arm studies, while 24 interventions (15.0%)<sup>78,80,94,100,105,123,177-179</sup> were compared with at least one other active intervention. Other comparative studies involved a control arm, typically no intervention or waitlist, but process and implementation outcomes were generally only reported for the active arm(s) without any controlled comparison.

### Intervention target

Included interventions intended to target DRV ( $n = 80$ , 50%), GBV ( $n = 53$ , 33.1%) or both ( $n = 27$ , 16.9%).

### Outcome evaluations (RQ3)

Seventy-nine active interventions were evaluated across 68 included studies. A description of the interventions identified is provided in [Report Supplementary Material 1](#).

Seven interventions were identified as having been evaluated across more than one study: Coaching Boys into Men (CBIM) ( $n = 2$ ),<sup>77,180</sup> Expect Respect ( $n = 2$ ),<sup>95,159</sup> IMPower ( $n = 2$ ),<sup>160,174</sup> Media Aware ( $n = 2$ ),<sup>97,98</sup> Safe Dates ( $n = 2$ ),<sup>36,166</sup> Shifting Boundaries ( $n = 2$ )<sup>100,156</sup> and TakeCARE ( $n = 2$ ).<sup>38,181</sup> Other interventions were found to share the same title (e.g. 'dating violence prevention program'), but were not clearly identified as being the same intervention.

The vast majority of studies ( $n = 65$ ) compared interventions against a control arm, namely an active control ( $n = 13$ ),<sup>36,38,87,92,160,162,165,169,170,176,181-183</sup> usual practice ( $n = 12$ ),<sup>35,83,98,99,174,184-190</sup> waitlist ( $n = 17$ ),<sup>49,79,84,88,89,94,95,97,167,168,172,175,191-195</sup> no intervention ( $n = 23$ )<sup>77,78,81,82,84,85,90,96,100,155,157-159,161,163,164,171,173,180,196-198</sup> or 'control' ( $n = 1$ ).<sup>199</sup> For the purpose of this review, one intervention was considered an active control, even though the content of the intervention was targeted towards DRV mechanisms.<sup>36</sup> This was because the intervention was almost entirely conducted outside the school setting, and therefore was not relevant for inclusion in the review. Limited information was typically reported about the nature of non-active control interventions, but often students received their usual classes (excluding topics related to DRV or GBV) over a comparable time period.

Only eight studies included a head-to-head comparison ([Table 5](#)). This included comparisons with different methods of delivering the same intervention (e.g. Sabella 1995),<sup>158</sup> and comparison of different intensity levels of the same intervention (e.g. Taylor *et al.* 2017).<sup>156</sup>

### Intervention target

Interventions were mostly intended to target DRV ( $n = 43$ , 54.4%). Fifteen<sup>92,96,157,158,160,168,171,174,175,180,184,192</sup> interventions targeted GBV (19.0%), and 14 interventions<sup>78,81,82,85,98,100,156,197</sup> were identified as targeting both (17.7%). The intended target was unclear for eight interventions<sup>79,94,170,176,191,195,198</sup> (10.1%), though these were included in the review as the intervention content included topics relevant to either DRV or GBV.

### Delivery of interventions

A breakdown of intervention facilitators is provided in [Table 6](#). The majority of interventions were facilitated by external staff only (35.4%) or teaching staff only (31.6%). A total of 38 interventions (48.1%) included at least some role for teaching staff in the delivery. A small number of interventions



TABLE 5 Head-to-head comparisons evaluated for outcome (efficacy)

Author, date	Comparison	Intervention 1	Intervention 2	Intervention 3
Achyut 2011 <sup>78</sup>	GEA + CAMPAIGN vs. CAMPAIGN vs. no intervention	GEA + CAMPAIGN	CAMPAIGN	-
Jewkes 2019 <sup>155</sup>	Skhokho vs. Skhokho + caregivers vs. no intervention	Skhokho	Skhokho + caregivers	-
Muck 2018 <sup>94</sup>	Scientist-Practitioner Program vs. Practitioner Program vs. control	Scientist-Practitioner Program	Practitioner Program	-
Niolon 2019 <sup>166</sup>	Dating Matters vs. Safe Dates	Dating Matters	Safe Dates	-
Sabella 1995 <sup>158</sup>	Sexual harassment classes: Peer-led vs. Adult-led vs. Self-led vs. control	Peer-led	Adult-led	Self-led
Taylor 2010 <sup>200</sup>	Interaction curriculum vs. law and justice curriculum vs. control	Interaction curriculum	Law and justice curriculum	-
Taylor 2011 <sup>100</sup>	Shifting boundaries: Building + classroom vs. building only vs. classroom only vs. control group	Shifting boundaries: Building + classroom	Shifting boundaries: Building only	Shifting boundaries: Classroom only
Taylor 2017 <sup>156</sup>	Shifting boundaries: 6th grade only vs. 6th and 7th grade vs. 6th, 7th and 8th grade	Shifting boundaries: 6th grade only	Shifting boundaries: 6th and 7th grade	Shifting boundaries: 6th, 7th and 8th grade

GEA, group education activities.

TABLE 6 Intervention facilitators (outcome evaluations)

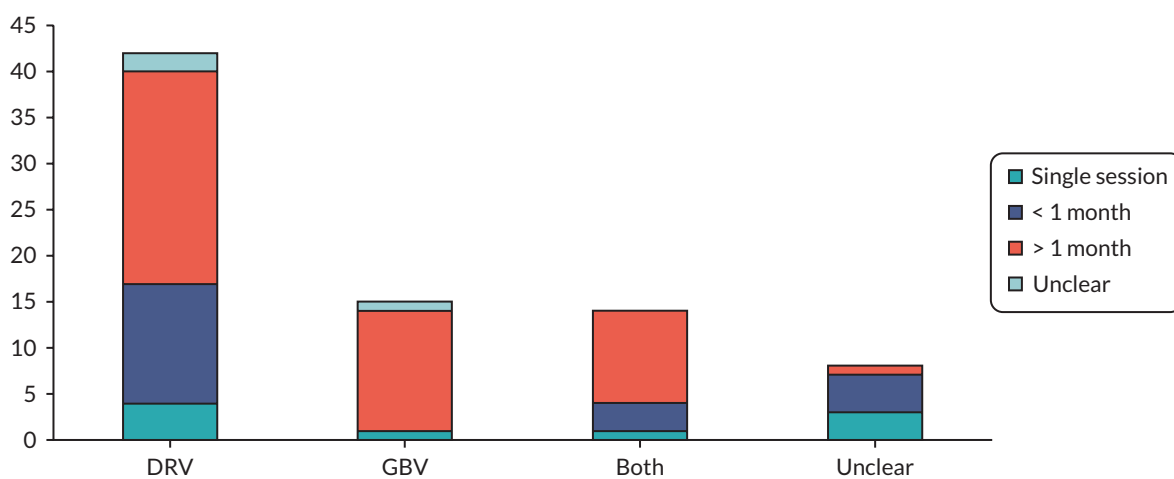
	Number of interventions (% of total)
External staff only	28 (35.4)
External staff and teaching staff	7 (8.9)
External staff and peer delivery	3 (3.8)
External staff, teaching staff and peer delivery	2 (2.5)
Teaching staff only	25 (31.6)
Teaching staff and peer delivery	4 (5.1)
Peer delivery only	2 (2.5)
No facilitator	3 (5.0)
Unclear	5 (6.3)

incorporated a peer-delivery component ( $n = 11$ , 13.9%),<sup>36,49,82,95,158,165,166,168,175,187,199</sup> or did not require a facilitator ( $n = 4$ , 5.0%).<sup>36,88,100,169</sup> There was no clear difference in choice of facilitator between DRV and GBV interventions.

The method through which interventions were delivered is outlined in [Table 7](#). The majority of interventions were delivered wholly (73.4%) or partially (16.5%) in person. A minority of interventions included a self-study (12.7%)<sup>88,158,163,169,171,184,188,189,195,198</sup> or digital (15.2%)<sup>49,79,96,97,161,166,170,171,173,182,188,189</sup> component (e.g. the use of virtual reality game format). One intervention<sup>100</sup> was delivered entirely as a community intervention. There was no clear difference in the method of delivery between DRV and GBV interventions.

**TABLE 7** Method for delivering the intervention (outcome evaluations)

	Number of interventions (% of total)
In-person delivery only	58 (73.4)
In-person with some self-study	4 (5.1)
In-person with a digital component	6 (7.6)
In-person with both a self-study and digital component	3 (3.8)
Self-study only	3 (3.8)
Digital delivery only	3 (3.8)
Community delivery	1 (1.3)
Unclear	1 (1.3)



**FIGURE 5** Duration of interventions according to intervention target (outcome evaluations).

The duration of the intervention was not clearly reported in many of the included studies. Overall, reviewers categorised nine interventions as being single sessions (11.4%), 20 interventions as lasting 1 month or less (25.3%), and 47 interventions as lasting longer than 1 month (59.5%). The duration of interventions according to intervention target is shown in [Figure 5](#); these were similar, though short interventions lasting < 1 month were more common for interventions targeting DRV only. Full details for the duration of interventions as reported by studies are provided in [Report Supplementary Material 1](#).

### Participant characteristics

#### Process and implementation evaluations (RQ2)

Sample characteristics for the included process evaluations are reported in [Report Supplementary Material 1](#). Across the 137 studies reporting process evaluation outcomes, data were reported for 109 independent samples of students. This includes one study that analysed three separate student samples (ICRW 2017<sup>85</sup> in India, Vietnam and Bangladesh), while 28 studies did not include students (i.e. the sample was school staff and teaching professionals only).

Demographic data for students were frequently not reported in included studies. In many instances, demographic data were only reported for the full sample of students included in the evaluation, and not reported separately for the intervention of interest, or specifically in those students relevant to process

evaluation outcomes. Demographic information was commonly not reported for students involved in qualitative research.

### Age

Where reported, students ranged in age from 7 to 19.3 years, though mean ages were all above secondary school age (mean age range 11.45–16.7 years, median mean age 14.5). No difference in age was noted between studies evaluating interventions for DRV and GBV.

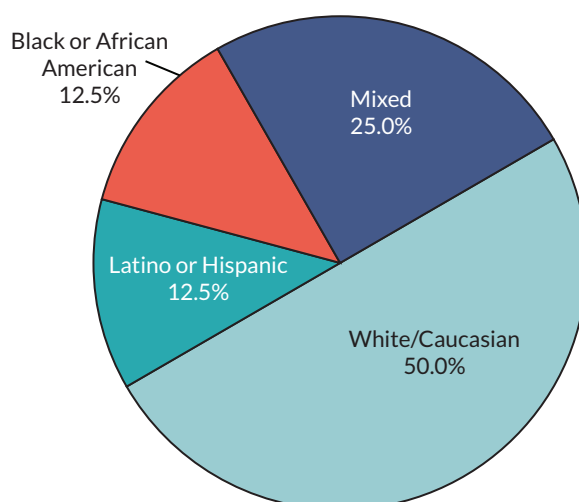
### Gender and sexuality

The vast majority of studies asked students to provide their sex or gender as a binary outcome. Five studies<sup>91,97,144,201,202</sup> permitted students to respond using alternative gender categories, although one of these (Scull *et al.* 2021)<sup>97</sup> did not report the number of students who identified with another gender category. In the remaining four studies, the number of students identifying with another gender ranged between 0.2% and 10%. One of these studies (Watson 2012)<sup>202</sup> evaluated an intervention intended to target GBV. [Note: when calculating sex/gender across studies, when a study reported data for a single sex only (e.g. 55% females), we assumed that the study offered a binary option for sex and lumped the remaining participants into the other sex category.]

Students reported their own sexuality in six of the included samples only (5.5%);<sup>91,202–206</sup> in four of these cases, the study was evaluating an intervention for GBV. Students not identifying as heterosexual ranged from 0% to 28% (median 15.8%). An additional study<sup>207</sup> reported that 100% of students were heterosexual, although this estimate was considered unreliable as it was based on staff judgement rather than self-reporting from students. There was no notable difference in the proportion of students not identifying as heterosexual between studies evaluating an intervention for GBV and DRV.

### Ethnicity

Race or ethnicity information was reported for students or the school setting in 33 samples (24.1%) and is displayed in [Figure 6](#). In 15 studies,<sup>97,98,123,147,201,204,206,208–212</sup> the majority student population was white or Caucasian. In 10 studies, ethnicity was mixed and no one race or ethnic group accounted for more than 50% of students.<sup>91,101,114,126,202,207,213–216</sup> The remaining studies reported a majority black or African American<sup>80,137,178,217</sup> or Hispanic or Latino<sup>96,218–220</sup> population (four studies each). Other ethnicity groups represented a minority of students in the included samples.



**FIGURE 6** Majority ethnicity/race group of process evaluation samples (where reported).

**Socioeconomic status**

Information about schools' or students' socioeconomic status (SES) was reported for 18 of the samples (16.5%),<sup>78,80,84,91,94,111,120,124-126,136,154,207,211,217,219,221,222</sup> although in some cases this was limited to a general statement about the district or school. Of these samples, nine samples<sup>80,124-126,136,154,207,217,221</sup> were identified as including a majority of students from lower SES backgrounds (e.g. free or subsidised school lunches, or in areas with high economic deprivation).

**Baseline risk**

Interventions in six of the included studies (4.4%)<sup>203,213,218,223</sup> were targeted towards students who had experienced or witnessed domestic violence or were considered to be at a high risk of DRV. One of these studies (Mgalla *et al.* 1998)<sup>138</sup> evaluated an intervention for GBV. An additional study (Kervin and Obinna 2010)<sup>224</sup> began with a universal sample, and then the intervention was moved to one school site where students were at high risk of not graduating and were considered to be at a higher than risk of domestic violence than the general population. All other samples were identified as universal, or abuse history was not reported (and therefore was considered likely to have included a universal sample).

**Outcome evaluations (RQ3)**

Cluster and sample characteristics for the included studies are reported in [Report Supplementary Material 1](#).

**Age**

Where reported, students in included studies ranged in age from 9 to 19 years, though all reported mean ages were above secondary school age (mean age range 12–16.5, median mean age 14.25). There was no clear difference in the age of students between studies evaluating interventions for DRV and GBV.

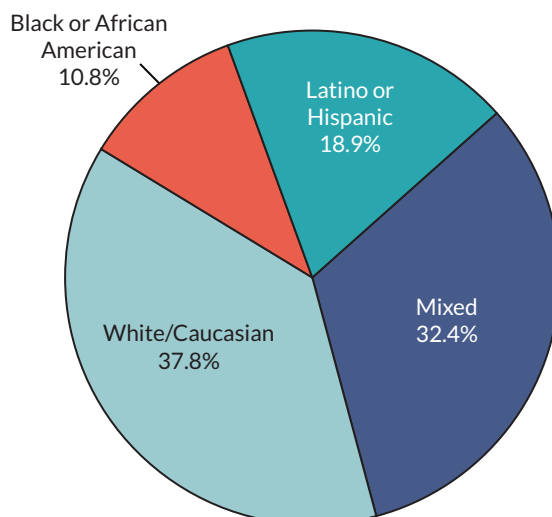
**Gender and sexuality**

The vast majority of studies asked students to provide their sex or gender as a binary outcome. Ten studies were conducted solely with a single-sex group: four<sup>77,170,180,183</sup> with males and six<sup>88,96,160,171,172,174</sup> with females. Where studies included mixed sex/gender samples, slightly more female students were included overall (mean 53.4%, range 41–76.3%). One study (Meiksin *et al.* 2020)<sup>91</sup> permitted students to record their gender as non-binary, or disclose that they were unsure about their gender, while another (Scull *et al.* 2021)<sup>97</sup> permitted students to record their gender as non-binary but numbers were not reported (numbers were lumped with students who did not disclose their gender). None of the studies that evaluated an intervention targeting GBV reported gender identities beyond the binary. [Note: when calculating sex/gender across studies, when a study reported data for a single sex only (e.g. 55% females), we assumed that the study offered a binary option for sex and lumped the remaining participants into the other sex category].

Only six studies<sup>35,49,81,167,182,197</sup> included in the review reported data for the sexuality of participants (8.8%). In five<sup>35,49,81,182,197</sup> of these studies where students self-reported their sexuality, the number of students not identifying as heterosexual ranged from 4.8% to 20% (mean 12.7%). In the sixth study,<sup>167</sup> it was reported that 3.6% of students had previously dated someone of the same sex. No studies that evaluated an intervention for GBV reported student sexuality.

**Ethnicity**

Race or ethnicity information was reported for students in 36 samples (52.9%). In a third of these studies, ethnicity was mixed and no one race or ethnic group accounted for more than 50% of students. In the remaining studies, more than 50% of students belonged to a white/Caucasian (37.8%),<sup>35,36,82,88,95,97,98,159,180,186,191,196,197,200</sup> Hispanic or Latino (18.9%),<sup>38,81,96,185,188,193,195</sup> or black or African



**FIGURE 7** Majority ethnicity/race group of outcome evaluation samples (where reported).

American (10.8%)<sup>89,92,155,166</sup> demographic (*Figure 7*). Other ethnicity groups accounted for a minority of samples only.

### **Socioeconomic status**

Information about schools' or students' SES was reported for 35 of the samples (51.5%), although in some cases this was limited to a general statement about the district or school. Of these samples, 11 studies<sup>38,81,85,166,175,180,181,187,189,192,195</sup> were identified as including more than 50% of students from lower SES backgrounds (e.g. free or subsidised school lunches, or in areas with high economic deprivation).

### **Baseline risk**

None of the identified studies exclusively included students who had experienced DRV or GBV. One study (Miller 1998)<sup>162</sup> included students with 'experience' of domestic violence, which may include DRV or experience of violence in the home (as victims, perpetrators or witnesses). A further study (Silverman 2000)<sup>99</sup> included students who were considered to be at high risk of DRV, which included students who had witnessed couple violence, experienced child abuse, reported current substance use, or expressed beliefs in the acceptability of couple violence or rape myths. All other studies included a universal sample of students regardless of their abuse history, although four studies<sup>166,169,184,186</sup> limited inclusion only to those students with 'dating experience'. The authors of these studies proposed this criterion either to identify those students who they considered more appropriate to receive a DRV intervention, or to best identify a change in DRV outcomes pre and post intervention.



## Chapter 4 Typology of intervention approaches and synthesis of theories of change (RQ1)

This chapter describes and reports on the qualitative syntheses used to produce an overall theory of change and summary components for the interventions included in this review, to address the research question:

*What are the theories of change (and components) of evaluated interventions for prevention of dating and relationship violence and gender-based violence?*

Drawing on intervention components analysis and a synthesis of theories of change, this chapter presents a typology of school-based interventions for the prevention of DRV and GBV, an analysis of the components used in these interventions, an overall theory of change for both DRV and GBV and the key themes that contributed to it, and a description of the overarching theory's key building blocks. Please note that this section contains extracts of text in the review protocol and a subsequent publication of the review findings<sup>225</sup> under the Creative Commons licence.

### Included studies

The syntheses reported in this chapter drew on all included studies. We present specific methods for the intervention components analysis and theory synthesis in turn.

#### *Intervention components analysis*

The intervention typology was informed by all included studies, while analysis of components focused on interventions reported in outcome evaluations ( $n = 68$ ). Drawing on intervention descriptions from the body of the text of each study, its supplementary appendices and, when necessary, additional information from the authors, we undertook initial inductive coding on a subset of 27 intervention descriptions, selected to capture a range of surface-level diversity. This generated an initial list of 143 different codes, organised around 11 preliminary groups that related to: content topics; skills discussed or practised; intervention activities; intended result of those activities; deliverer; location; the presence of others; how students are guided through the intervention; medium(s) of delivery; how school staff were involved with the student portions of the intervention and/or received their own material; and how the extended community (parents, families, local groups, etc.) were involved with the student portions of the intervention and/or received their own materials.

Through discussions between the co-investigators, and based on the principles of axial coding, these 11 groups were re-organised into an intervention typology:

- intervention type, the essential intention or 'spirit' of the programme
- intervention focus, the broad content included
- intervention domains of implementation, using the socioecological model to account for the levels over which the intervention was implemented.

We also described intervention components, defined as the activities in which students, school staff and the extended local community experience as part of the intervention.

**TABLE 8** Coding template

Themes	Codes and subcodes
<i>Inputs</i>	Curriculum, training of teachers, materials to deliver intervention
<i>Intervention goals</i>	Reduce bullying perpetration, peer victimisation, reduce sexual harassment victimisation and perpetration, increase bystander behaviour
<i>Key theoretical concepts</i>	Mid-range theories, for example social learning theory, theory of planned behaviour, theory of reasoned action
<i>Mechanisms of change</i>	Knowledge, attitudes, skills, behaviours
<i>Outcomes</i>	Reduce/prevent dating violence behaviours, reduce/prevent GBV behaviours

### **Synthesis of theories of change**

The synthesis of theories of change drew on different sets of included studies in cycles. After an initial coding round of 5 interventions described in 10 reports, we undertook exploratory coding between 2 reviewers informed by template analysis.<sup>226</sup> Using a coding template with codes exchanged by two reviewers in a subset of interventions ( $n = 14$ ) reported in two or more reports ( $n = 38$ ), focusing as well on breadth of intervention type. The template reflected the focus on programme theories and included intervention inputs, components, mechanisms, outcomes and key theoretical concepts ([Table 8](#)). An example of the coding template as applied to interventions is provided in [Report Supplementary Material 1](#).

We then synthesised programme theories using thematic synthesis.<sup>67</sup> A subset of interventions that were reported in two or more papers were prioritised: we identified similarities, differences of emphasis and contradictions between multiple papers and developed summaries of the theories of change for the interventions, describing inputs, mechanisms of change, underlying theoretical assumptions, and proximal and distal outcomes. For the remaining interventions, summaries of the programme theories were produced. The analysis proceeded with another 37 reports analysed, at which point thematic synthesis yielded no new insights. The remaining included reports were analysed using a confirmatory coding approach.

The findings of the thematic synthesis were transformed using meta-ethnography to develop a line of argument for an overarching theory of change. The key concepts identified during coding were treated as ‘first-order constructs’<sup>227</sup> which we summarised as logic models and compared between each intervention (see [Report Supplementary Material 1](#) for examples). This generated second-order constructs, which were our interpretations of the first-order concepts. Third-order constructs were then developed (based on key concepts and second-order interpretations) to produce a line of argument that aimed to describe the overall mechanisms of change.

### **Setting the scene: key theoretical frameworks**

In undertaking our synthesis for RQ1, three key frameworks emerged as central to the analysis. The first was a socioecological approach, which provides an account of multilevel influences in social systems. We primarily used five levels in our analysis:

- student intrapersonal, relating to students’ cognitions, understandings and internal processes;
- student interpersonal, relating to relationships between students;
- staff, relating to impacts on teachers and school staff;
- structural environmental, relating to physical aspects of school spaces;
- structural social, relating to school-level social contexts, norms, policies and practices.



An important distinction that arose in our analysis was between domain of implementation (i.e. the level at which components were delivered) and domain of mechanism (i.e. the level at which interventions were hypothesised to generate change). While often interrelated, these did not always coincide; for example, components implemented at structural social levels may have impacts on students' intrapersonal processes. We thus used domain of implementation and domain of mechanism flexibly. For example, our analysis of intervention type drew on both domain of implementation and domain of mechanism, whereas intervention components were described in terms of domain of implementation, and mechanisms in the theory synthesis were described in terms of domain of mechanism.

The second framework was the theory of human functioning and school organisation, originally described by Markham and Aveyard (2003).<sup>228</sup> We used this as it offered a framework for structuring the concepts and addresses the deeper structural influences on young people's behaviour.<sup>46</sup> The theory proposes that schools can assist students to develop 'essential human capacities': practical reasoning, or the ability to reason, critically reflect, and consider issues from a range of different perspectives; and affiliation, or concern for others and the ability to form relationships. Schools can play a role through their instructional order, which relates to the way in which they enable students to learn, and regulatory order, that is engendering a sense of belonging by encouraging students to internalise its values and adopt behavioural norms. Securing commitment from students depends on weakening the 'classification', that is school's institutional 'boundaries' and how rigidly they are set, as well as the 'framing', the degree to which a school's communication and pedagogic practices are student-centred. We describe the specific propositions we developed from this theory below.

The third was a systems thinking approach. Our analysis acknowledged that schools are 'bounded ecological systems' (Moore *et al.* 2019, p. 26<sup>47</sup>) into which interventions are introduced. A systems perspective recognises that schools comprise different subsystems, such as year-groups and form-groups, nested within bigger systems such as national education systems, the local community and other parallel systems, such as families.<sup>229</sup> Important interactions can occur within and between these systems.

## Intervention typology

Our intervention typology focused on intervention type, intervention focus and intervention domain of implementation. These are presented in tabular form in [Report Supplementary Material 1](#).

### Intervention type

The intervention 'type' captures the essential intention of the intervention. Categorisation depended on the intervention's domain, focus and components, as well as most importantly a consideration of the specific change mechanisms being targeted by the intervention. Programme type is a classification of the level of sophistication and complexity in how the intervention intends to accomplish the goal of reducing DRV/GBV. This classification allows us to look across the highly varied landscape of DRV/GBV prevention programmes and to group them in ways that make for more useful comparison.

### Single-component interventions

Single-component interventions accounted for about one-quarter (27%,  $n = 53$ ) of the programmes evaluated. Defining characteristics include that they tend to be shorter; have a key technology or novel activity as integral to their effectiveness; may or may not require any in-person facilitators to be delivered; and focus on a single or very narrow range of change mechanisms.

Single-component intervention *durations* range from one long session (e.g. watching and discussion a video; Durand 1997<sup>83</sup>) to a week of daily sessions (e.g. Joppa *et al.* 2016<sup>186</sup>). The *integral technology or novel activity* could include a specially designed video game, like *Jesse*, where students engage in role-play to consider intervening in different prosocial ways.<sup>161</sup> Alternatively, the novel component could be a special training day for senior athletes to teach them to have conversations with their younger

peers [e.g. Athletes As Leaders (AAL)<sup>216</sup>]. Some single-component programmes have lesson plans for their few sessions that can be delivered by teachers or external presenters (such as in Jaycox *et al.* 2006<sup>185</sup>), whereas others are completely self-guided by the students themselves (such as in Fernandez-González and Sánchez-Álvarez 2020<sup>169</sup>). Lastly, single-component interventions focus on a single change mechanism or a narrowly defined range of change mechanisms. A single change mechanism programme could take the form of having students only watch a video as a means of changing their individual beliefs relating to taking action as a bystander behaviour (e.g. Jouriles *et al.* 2019<sup>181</sup>). Or, a programme could focus on a narrow range of change mechanisms over several sessions, such as by having different sessions where students practice identifying warning signs of abuse, learn assertive response behaviours, and have false beliefs about seeking help corrected through information presentation; all of which target individual behaviour through different but closely related change mechanisms of skill acquisition, knowledge acquisition, and belief challenging about a single situation and behaviours (e.g. Levesque *et al.* 2016<sup>182</sup>).

### Curriculum-based interventions

Fewer evaluated programmes were curriculum-based interventions (13%,  $n = 25$ ). The key distinctions between curriculum-based interventions and the less complex interventions discussed above are their integration into existing school curricula, a longer programming duration, an increased amount of training is expected of programme facilitators, and these programmes focus on at least a narrow range of change mechanisms across one or two social ecological levels but not structural levels.

Curriculum-based interventions, as the label implies, have a distinguishing feature of usually being *integrated into the school schedule*. This usually takes the form of replacing the content of an existing health or wellness class that students already attend with DRV/GBV-focused material, or else special assembly times were added during the school day if students needed to be organised different than their normal class groups (i.e. separated by sex/gender).

To be classified as a curriculum-based intervention, the programme *duration* needed to last longer than a school week, meaning at least six sessions, which could be delivered on consecutive schooldays (e.g. Jung and Min 2013<sup>173</sup>) or over several months (Baiocchi *et al.* 2017<sup>160</sup>). Although programmes like the *Good School Toolkit*<sup>175</sup> included several years' worth of curricular lesson programming material, such programmes with that high amount of content tended to also contain additional elements which classified them as multicomponent or multilevel interventions (see following sections). Programme sessions are integrated into school classes typically in place of the health, wellness, sexual education or special elective class periods.

Curriculum-based interventions are *delivered* by extensively trained facilitators following manualised or scripted plans. Facilitators came from a variety of sources. The *IMPower* programme utilises staff from each school who prior to implementing the programme, 'attend a 3-week, 126-h[our] intensive training' (Decker *et al.* 2018, p. 1343<sup>174</sup>). Other curriculum-based programmes like *Expect Respect Support Group*<sup>223</sup> are delivered by external facilitators trained specifically by an external organisation or else are delivered by the programme designers and their team (e.g. *Connections: Dating and Emotions Curriculum*<sup>209</sup>).

A *narrow range of change mechanisms* at one or two levels of the socioecological model (SEM) can look very different from programme to programme, depending on which levels beyond the individual student that the programme targets. For example, the narrow focus could be only looking at a certain type of DRV/GBV, where the intervention sessions address sexual harassment prevention specifically through challenging myths about sexual harassment incidents, providing knowledge about actual sexual harassment statistics and situations, and practice behaviours for responding specifically to sexual harassment (e.g. Sabella 1995<sup>158</sup>). Other programmes might focus on a specific behaviour across many situations, such as how to communicate boundaries in both stranger and dating relationship situations (e.g. Lee *et al.* 2018<sup>87</sup>).

### Multicomponent interventions

About one-quarter (24%,  $n = 48$ ) of evaluated programmes are multicomponent interventions. These utilise not only a *variety of modes of intervention* (e.g. curriculum, videos, theatre activities, games, etc.) but also, in contrast to single-component or curriculum-based interventions, *address a limited range of change mechanisms across the student and/or staff levels of the SEM*. For example, in the *Green Dot High School* programme content for students, activities seek to address 'violence victimization, perpetration, and prosocial behaviours' (Coker *et al.* 2017, p. 567<sup>82</sup>) through motivational presentations that the entire student body receives, as well as more time-intensive bystander intervention trainings done with a selected group of student leaders who help to disseminate the message throughout the rest of the students.

Some multicomponent interventions also included change mechanisms with school staff. Interventions like this sought to provide information and activities *to teachers* rather than merely *through teachers* as facilitators. In some cases, trainings were designed for both goals, 'to increase the teachers' knowledge about dating violence, as well as prepare them to implement the curriculum' (Avery-Leaf *et al.* 1997, p. 13<sup>196</sup>). In other interventions a workshop or educational material was created specifically for the unique roles that school staff might play in addressing DRV/GBV. For example, the *Bringing in the Bystander - High School Curriculum* (BITB-HSC) describes their workshop for school staff as providing 'school personnel with specific behavioural strategies for how they can talk to teens about RA and SA, model healthy attitudes and behaviours for teens, and intervene safely and effectively in situations of teen RA and SA' (Edwards *et al.* 2020, p. 162<sup>230</sup>).

### Multilevel interventions

Approximately one-third of evaluated programmes (36%,  $n = 70$ ) are multilevel interventions. Multilevel programmes are complex interventions which use a *variety of modes of intervention*, are designed to target *multiple levels of the socioecological model*, and integrate explicit components that *target structural aspects of the school*. Structural aspects include the literal *environmental structure* of the school building and grounds. One way that programmes attempt to change school spaces is by providing posters/flyers to hang in common spaces (e.g. Bando *et al.* 2019;<sup>184</sup> Miller *et al.* 2015;<sup>93</sup> Niolon *et al.* 2019<sup>166</sup>). This alteration to the visual landscape of the school not only reinforces anti-DRV/GBV messages to students every day as they walk by, but it also communicates to newcomers and visitors what the school values. Other interventions target environmental change by improving the way that school space is monitored. Typically, they begin by asking students to identify high risk areas around school as 'hot spots' where violence frequently occurs; the school then creates safety plans or monitoring schedules among staff so that those spaces are no longer left unsupervised during the school day (e.g. Mathews *et al.* 2016;<sup>90</sup> Meiksin *et al.* 2020;<sup>35</sup> Taylor *et al.* 2011<sup>100</sup>).

Changing the *social structural* aspects of a school through targeting of the climate and norms is another way that interventions attempt to change structure. While posters can contribute to this, many programmes involve more direct activities. For example, in *Project Respect* key school staff receive training at the beginning of the programme, not only to learn how to implement the programme lessons but to help them 'review school rules and policies to help prevent and respond to gender-based harassment and DRV' (Meiksin *et al.* 2020, p. 3<sup>35</sup>). Others, like the *Gender Equity Movement in Schools (GEMS)*<sup>231</sup> or the *Good School Toolkit*,<sup>175</sup> foster the creation of school clubs so that students interested in anti-violence work can organise among themselves and create new school norms through sustained and self-directed action.

### Intervention focus

Intervention content addressed two broad areas: a focus on perpetrators and victims and/or a focus on bystanders. Most interventions (86%,  $n = 168$ ) focused on potential perpetration and victimisation but this content varied considerably. For example, one programme in Brazil *focused on perpetration* reduction by teaching content on coping skills, including 'practical aspects of how to behave in situations with aggressive potential in a non-violent way' (Filho 2017, p. 43<sup>84</sup>). Adolescents could use those coping

skills to replace their own violent tendencies or normalised violent reactions in those types of situations in future encounters. Other programmes, such as the *Fourth R* programme, addressed perpetration by targeting social norms about violence that students might see and 'counteracting pro-abuse messages from peer culture' (Cissner and Ayoub 2014, p. 32<sup>81</sup>). Such messages can normalise the idea that violence is an acceptable response in some situations, so some programmes address it directly by correcting myths and misperceptions about violence, the use of violence and the impact of violence on both perpetrators and victims (Knowles 1997;<sup>232</sup> Madsen 1994;<sup>233</sup> Maphosa 2017;<sup>136</sup> Rogers *et al.* 2018<sup>127</sup>). Content *focused on victimisation* also varied considerably. In some interventions this consisted of directly addressing adolescents as potential victims through teaching them active resistance or self-defence skills (e.g. Baiocchi *et al.* 2017;<sup>160</sup> Decker *et al.* 2018;<sup>174</sup> Ekhtiari *et al.* 2014;<sup>234</sup> Jordan and Mossman 2018<sup>147</sup>). Others focused on teaching adolescents to avoid such situations by talking about 'red flags' – an indication which might signal that a relationship or a situation is unhealthy or could potentially become dangerous (e.g. Busch-Armendariz *et al.* 2008;<sup>235</sup> Cascardi and Avery-Leaf 2014;<sup>80</sup> Genereux 2020;<sup>236</sup> Taylor *et al.* 2017<sup>156</sup>).

Fewer interventions (27%,  $n = 52$ ) *focused on bystanders* as a change pathway. Even within this smaller group the nature of how bystander helping is addressed was highly variable. Some programmes provided examples of bystander intervention in a variety of scenarios, stories or video clips (e.g. Cissner and Ayoub 2014;<sup>81</sup> Jouriles *et al.* 2019;<sup>181</sup> Merrell 2004<sup>92</sup>) so that students develop self-efficacy around using those concrete strategies in future situations they might encounter. Other interventions provided the scenarios but asked students to brainstorm or role-play different ways the situation might be risky and how to best respond those risks by overcoming obstacles to helping (Coker *et al.* 2017;<sup>82</sup> Lee *et al.* 2018;<sup>87</sup> Pacifici *et al.* 2001<sup>191</sup>). More general promotion of bystander behaviours was also included in many programmes. *Relationships Without Violence*<sup>214</sup> encouraged bystander behaviour with content that promoted 'pro-social behaviours, positive peer culture, and healthy masculinities/femininities' (p. 4). *Green Dot*<sup>82</sup> and *BITB*<sup>197</sup> both included direct skills practice for bystander intervention as well as reflective content addressing rape myths and attitudes supportive of or tolerant towards violence. Addressing myths and attitudes is an indirect way to increase bystander behaviour through reducing barriers to action. Bystander intervention can also take place after violence has occurred. The *Women's Self Defense Network-Wahine Toa* courses in New Zealand<sup>147</sup> not only taught skills for responding to attempt assaults but also focused on how to support others if they have experienced violence by helping them to seek support and access local resources.

### **Intervention domains of implementation**

Intervention implementation can be divided into student intrapersonal, student interpersonal, staff, structural-social and/or structural-environmental domains. The student individual level was addressed in two ways: the most common form was improving *students' intrapersonal traits* (86%,  $n = 168$ ), such as their own attitudes, opinions and behaviours. In addition, two-thirds of programmes focused on improving *students' interpersonal traits* (64%,  $n = 126$ ), such as incorporating skill-building to help students create healthier and more supportive relationships with their peers (e.g. Bando *et al.* 2019<sup>184</sup>).

About half of the interventions focused on *staff* (45%,  $n = 95$ ) as a domain of interest. This most often included preparing the school staff to deliver programme content to students as the key recipients. Additionally, some interventions targeted school staff specifically, to support them in how to respond to violence and improve their capacity for addressing violence that takes place in school settings. For example, the *BITB-HSC*<sup>197</sup> includes a 60-minute workshop designed to help school personnel learn skills to be positive bystanders when they see situations of adolescent interpersonal violence.

Interventions addressed the structural level in two ways. Just under a third of programmes (30%,  $n = 58$ ) focused on implementing components to change the *social* structure of a school by improving policies and influencing social norms in the school setting to be less tolerant of violence and more supportive of prosocial behaviours such as believing survivors and intervening in risky situations. Fewer programmes (13%,  $n = 25$ ) focused on changing the *environmental* structure of the school, including addressing things

in the physical landscape of the school and/or who is present in school spaces. This included alterations in staff practices (e.g. increasing hallway monitoring) or changing the visual landscape of the school (e.g. posters that list local survivor support resources).

## Intervention theory of change

The logic model (see [Figure 8](#)) depicts the overarching theory of change for the school-based interventions. We used the theory of human functioning and school organisation, with its focus on institutional processes, to theorise how interventions can modify institutional boundaries and improve the connections between the various systems. After describing our starting point, we present intervention inputs, including the analysis of intervention components; intervention mechanisms; and intervention outcomes, and conclude with several propositions that we explore in subsequent analyses.

### Starting theoretical propositions

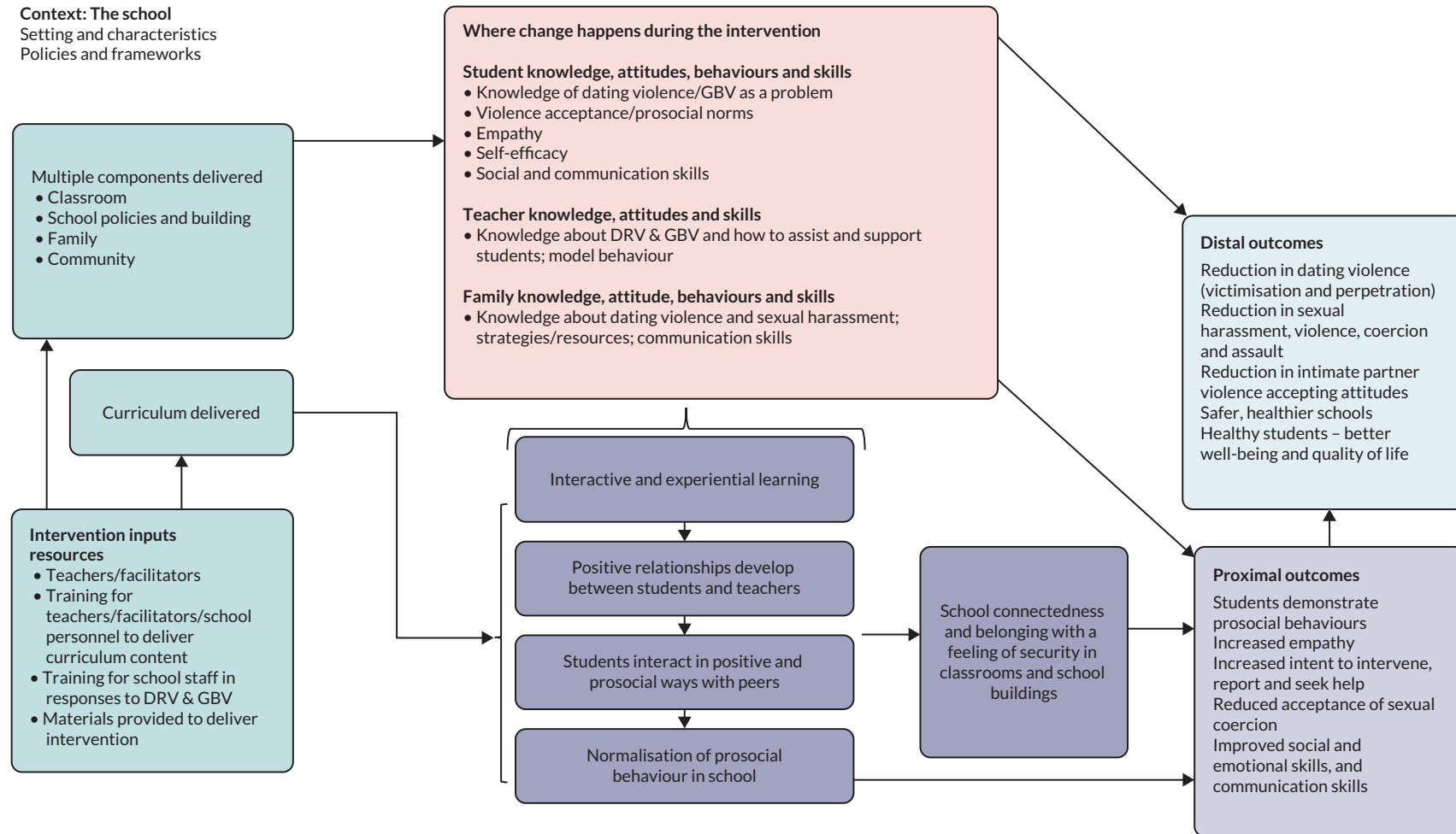
Shaped by Markham and Aveyard's (2003)<sup>228</sup> theory, we developed two starting hypotheses, broadly summarised as that school-based interventions designed to reduce DRV and GBV have the greatest potential to be effective when based on the principles of weakening classification and framing. Specifically, and first, we hypothesised that interventions aimed to weaken the institutional boundaries between the school and surrounding communities and strengthen relationships within the school. For example, several interventions with curricular and classroom-based components had additional components directed at the community and family/parents to support and reinforce change initiated in the classroom. Many of the interventions were delivered by teachers, some in collaboration with students, which opened opportunities for increased interaction and positive relations between teachers and students. In some cases, teachers and other school staff were trained to model positive behaviours to reinforce behaviour change among the students. Importantly, as our analysis progressed, our focus shifted from 'weakening boundaries' to 'improving commitment', based on an understanding of the diverse contexts in which interventions were implemented. As discussed below, weakening boundaries, for example with communities can be undesirable in the context of violent communities and harmful gender norms. Second, we hypothesised that interventions aimed to involve students in their learning and in school-level decisions and increase belonging, thereby weakening the framing of the schools. For example, many of the interventions utilised student-centred techniques such as group work and role-playing to encourage interactive and experiential learning with student input to their own learning, which in turn, promoted positive interactions and relations between students and teachers and among the students themselves. This facilitates students' capacity for practical reasoning and critical reflection.

### Intervention inputs

#### Characterising intervention inputs

Most interventions involved classroom curriculum components (see [Figure 8](#), light blue box) delivered within the school timetable or 'after school'.<sup>90,165,174,215</sup> Some interventions had additional components targeting multiple levels: for example, the Safe Dates intervention<sup>36</sup> had an out-of-school component with training of community service providers and a weekly support group for victims of dating violence; and the JOVEN intervention<sup>193</sup> had six group sessions for students and two for parents, with parents joining students in the last session to practise healthy communication skills in negotiating 'curfews' and dating. Others can be described as multicomponent or multilevel interventions, for example the Expect Respect project which targeted all groups within the school community:

*The Expect Respect Project utilized five programme components: a 12-week classroom curriculum for students, trainings for staff, education and support for parents, encouragement and guidance to administrators for policy development, and support services for students who had been affected by bullying, sexual harassment, or sexual or domestic violence.*<sup>237</sup>



**FIGURE 8** Integrated theory of change for school-based interventions. Light blue is intervention inputs; pink/orange is lower-level change mechanisms; dark blue is higher-level change mechanisms expected to facilitate outcomes; grey is proximal outcomes and light green is distal outcomes.

Many of the interventions involved training teachers to be able to deliver curriculum content (see [Figure 8](#), light blue box), while others did not require teacher input as the curriculum was delivered by external facilitators. External facilitators were often very experienced in prevention of GBV and DRV (e.g. Muck *et al.* 2018<sup>94</sup>) and could also receive extensive training: for example, the rape crisis educators who delivered the *Green Dot* curriculum had a 4-day in-depth training and received individualised feedback throughout the trial.<sup>238</sup> In another example, the 'instructors' delivering the curriculum for the *IMPower and 50 : 50* intervention were recruited through an 'intensive process' that ensured that they were passionate about preventing sexual violence and were respected members of their communities. They:

*... received extensive instruction by expert facilitators and participated in mock interviews and field-training exercises ... Trainers were required to pass a rigorous examination consisting of a written test, oral examination, and physical skills demonstration before becoming paid employees teaching the curriculum at intervention sites.*<sup>160</sup> (pp. 819–20)

In addition, throughout their first year of teaching, trainers were also supervised by an experienced trainer. There were those interventions that used both teachers and external facilitators such as the *Benzies and Batchies* intervention,<sup>168</sup> with teachers delivering the introductory and last classes and trained social-skills instructors delivering three lessons on student skills. Some interventions relied solely on teachers to deliver the intervention content such as *Beyond Victims and Villains (BVV)*, an audio-drama about gender and gender violence.<sup>134</sup> In some cases, teachers and other school staff also received training to enable them to be positive role models for students; for example, in the BITB-HSC intervention teachers and other school personnel were trained '... to be positive bystanders in situations of adolescent interpersonal violence' (Edwards *et al.* 2019, p. 489<sup>197</sup>). Several interventions did not require teachers or facilitators for implementation such as the healthy relationships text messaging campaign<sup>215</sup> and the online programme, *Teen Choices*.<sup>182</sup> It is cited as an advantage that it used computers and expert system technology to deliver the programme 'efficiently and 'cost-effectively' when compared to programmes that require significant teacher training or input from professional educators.

Topics and skills covered varied in terms of depth from those that had a specific focus such as the legal aspects of dating violence in the three-class-period *Ending Violence* intervention<sup>86,185</sup> to those that had a broader focus such as the social-emotional learning programme, *Second Step*, which was delivered in 41 lessons over 3 years.<sup>37,192,239</sup> Some of the interventions could best be described as 'add-ons' such as the *Katie Brown Educational Program (KBEP)*,<sup>186</sup> a 'freestanding' DRV programme that could be incorporated into any school curriculum by replacing five health class periods during 1 week. Other interventions integrated lessons into the existing school curriculum, often within health and physical education classes (e.g. Avery-Leaf *et al.* 1997,<sup>196</sup> Peskin *et al.* 2019<sup>188</sup>). In some cases, integrating interventions into existing curriculum requirements was an attempt to weave together violence-related and broader learning. For example, the *Fourth R* intervention integrated dating violence prevention with lessons on healthy relationships, sexual health and substance use; these topics were not addressed directly, rather healthy, non-violent relationship skills were an underlying theme that was 'woven' throughout the curriculum. For the authors, the advantages of integration were clear:

*[t]he focus on embedding the programme into curriculum that meets the guidelines for mandatory classes in high schools provides a vehicle for widespread dissemination and sustainability far beyond that which can be achieved by add-on programmes.*<sup>190</sup> (p. 698)

### Intervention components analysis

Our analysis drawing on outcome evaluations yielded 40 individual components as well as a superordinate hierarchy which categorised those components into 13 programme activities (see [Report Supplementary Material 1](#) for a full table). Nine of those activities (and their corresponding 25 programme components) describe intervention elements that target students directly, while the other 4

programme activities (and their 15 components) describe intervention elements that target non-student levels, including school personnel, family/parents of students, structural environmental and structural social aspects of school. A summary of these activity categories follows.

### **Student components**

**Guided practice activities:** Most evaluated interventions (83%,  $n = 63$ ) include activities where students practiced some behaviour or element of a behaviour and received feedback about their choices and reactions. This most frequently took the form of active *role-playing* (57%) where students act out how they would respond in a situation where they are the target or a bystander to DRV/GBV. For example, in the JOVEN intervention students engaged in role-playing activities ‘to act out healthy bystander behaviours with friends who were in risky situations and at risk for physical or sexual harm’ (Gonzalez-Guarda *et al.* 2015, p. 412<sup>193</sup>). Another method of guided practice included providing written *vignettes* (34%) describing DRV/GBV to which students would write their hypothetical reaction. For example, in the Rhode Island Teen Dating Violence Prevention Program (RITDVPP) intervention students consider a range of teen dating violence (TDV) vignettes and ‘brainstormed healthy behaviours to solve real life relationship dilemmas’ (Silverman 2000, p. 75<sup>99</sup>). About half of the programmes (43%) engaged in *guided practice not otherwise specified* in the programme descriptions. This included programmes that provided no detail beyond stating ‘skill building’ or ‘skill practice’ happened during the programme sessions.

**Discussion activities:** The majority of interventions (80%,  $n = 61$ ) included discussion-based activities where students would engage with each other in *groups* (78%), in *pairs* (9%) or through *question-and-answer* formats with their instructor (20%). Discussions are an interactive activity where students are required to think and apply knowledge as well as have the opportunity to ask questions about topics about which they are unsure. Importantly, most discussion formats allow for students to hear the ideas, reasons and explanations of their peers.

**Reflection/thinking activities:** Most evaluated programmes (78%,  $n = 59$ ) included activities in which students engaged in individual processing of material. This was most often through providing individual *reading material* (64%) or completing *workbooks/worksheets/writing assignments* (58%). A smaller number of programmes included *diary-keeping* (7%) as part of the sessions. For example, the *It's Your Game* intervention required students to complete ‘individualized journaling activities ... to help students personalize information’ (Peskin *et al.* 2014, p. 1472<sup>189</sup>).

**Visual/image-based activities:** More than half of evaluated programmes (58%,  $n = 44$ ) included activities centred around visual materials such as posters, films, video games or cartoons. The most common visual programme element was the use of a *film or video* (39%), which often serves as the basis for a following discussion or series of discussions among the students (e.g. Pacifici *et al.* 2001<sup>191</sup>). *Graphic novels or cartoon strips* were used by some programmes (16%), for example in *Let Us Protect Our Future*<sup>176</sup> a comic workbook was distributed to students to provide characters and a storyline they could follow along with throughout the intervention. Some programmes include *posters and other visuals* (14%) which are put up around the school and classrooms. For example, *Safe Dates*<sup>36,240-243</sup> not only provides visual materials with key programme messages but also engages students in creating more visual materials such as posters as a summative or concluding activity to have students shown what they learnt during the intervention. *Video games*, either via the internet, on a local personal computer, or a CD-ROM were also used in some programmes (13%). In a few cases, the entire programme consists of playing the video game (e.g. *PR:EPARe*<sup>79</sup>), whereas in other programmes playing the video was to supplement other activities (e.g. *Me and You*<sup>188</sup>).

**Narrative-based activities:** A third of evaluated programmes (33%,  $n = 25$ ) include activities that revolve around listening to the stories of others’ experiences with DRV/GBV or sharing one’s own stories of experiencing DRV/GBV. Often this includes asking students to *share personal stories* (21%), for example in *IMPower* adolescent women are encouraged to share their assault experiences as part of self-defence



training and also to inform resource dissemination to survivors of prior DRV/GBV in the class.<sup>160</sup> Other interventions involve *reading or listening to personal stories* (11%). This was frequently done by bringing in guest speakers to talk with students; or more simply students read stories from survivors, perpetrators or bystanders to violence. In *Expect Respect*<sup>95</sup> the entire third session involves students listening to and talking with a survivor of relationship abuse from the local community, to show what violence can look like in their own community. A small number of programmes (9%) incorporate *theatre shows, skits or plays* into the activities for students (such as in *Benzies and Batchies*<sup>168</sup>). Some programmes that used theatre incorporated students into the creation process of the performance. For example, in *Youth in Action Bystander Theatre*<sup>244</sup> students 'present an unsolved problem to the audience. The problem posed is a scene that comes from a real-life story of one of the participants, who then, drawing upon her real-life experiences, plays the protagonist in the scene' (p. 17). Most programmes provide a script for students to learn and perform for their peers (e.g. de Lijster *et al.* 2016;<sup>168</sup> Merrell 2004<sup>92</sup>).

**Competition activities:** A quarter of evaluated programmes (24%,  $n = 18$ ) include activities with a competitive component where two or more players or teams compete in giving answers, finishing a story, or some other game. Although programme descriptions often refer to these activities as *games* of some sort (22%), the content of these can be more serious than a typical child's game. For example, in *KBEP*<sup>186</sup> the gameplay involves students practicing assertive communication skills when confronted with dating violence (DV)-related behaviours. The game could also be a contest, as in the *You-Me-Us* programme<sup>199</sup> where students compete to design a healthy relationship message, with the winner's idea being used in the actual prevention campaign materials in their school community. In other programmes, the games had little relevance to violence but instead are meant to lighten the mood or break up the more serious content, as in the *GEMS* programme where the school-based campaign includes 'games and races' (ICRW 2017, p. 9<sup>85</sup>) alongside essay writing, speeches and pledges for equality and nonviolence. A few programmes mention *debate* (7%) as a form of competition among students during programme sessions, although none of those programmes provide examples or elaborate upon the nature of the debates nor the topics that students debate. One programme utilises a *game show* concept (*PR:EPARe*<sup>79</sup>) within a video game so that students can engage in competition virtually with the computer rather than against each other.

**Guidance:** The relationship between the student in the programme and the person facilitating or guiding the programme is coded to be inclusive rather than discrete. Portions of a programme could be facilitated by an adult, a peer or by the student themselves. The majority of evaluated programmes are *adult-facilitator-led* (86%,  $n = 65$ ). Adult facilitators could include external facilitators from a national or local agency, a trained team of presenters or performers from a local university, a single presenter (usually the creator of the programme) or school personnel who had received training or other materials to help them present lesson plans or other programme elements. Some evaluated programmes are *peer-facilitator-led* (8%,  $n = 6$ ). These students usually attended additional meetings and trainings prior to the programme implementation so as to learn how to facilitate the programme – in some programmes peer facilitators volunteer (e.g. Coyle *et al.* 2019<sup>199</sup>) and in other programmes peer facilitators are nominated by students or school staff because they are seen as leaders or influential in school (e.g. Miller *et al.* 2015<sup>187</sup>). Finally, some evaluated programmes are *self-paced* (16%,  $n = 12$ ) meaning the students usually work through material individually (or mostly on their own) and only come together as a group for a few topics. This can allow students to work at a pace that best suits them and allows them to fully comprehend material through rereading or reviewing work material additional times, etc., until they are ready to move to the next topic.

**Programme gender/sex composition:** The nature of whether classes were single-sex/gender or mixed is coded to be inclusive rather than discrete. Some programmes included sessions that were single-sex/gender early on in the programme but later sessions were mixed (e.g. Filho 2017<sup>84</sup>). Programmes that are entirely self-paced (i.e. only involve individual work) were excluded from this coding, as it is meant to capture the potential for interaction between students of the same or different sexes/genders, and no interaction at all takes place in these programmes. The majority of evaluated programmes use *mixed-sex/*

*gender groups* (72%,  $n = 55$ ) for at least one session or activity. Only a quarter of evaluated programmes use *single-sex/gender groups* (24%,  $n = 18$ ) for at least one session or activity.

**Miscellaneous:** A single programme utilised a component that did not fit into another category nor was anything similar used in any other programme. *Project Respect* in the UK<sup>35</sup> encouraged students to download and use the freely available *personal safety app* 'Circle of 6' ([www.circleof6app.com](http://www.circleof6app.com)). The app can be used on mobile devices and 'helps individuals discreetly request help from their preidentified contacts for support if threatened by/experiencing dating and relationship violence' (p. 4).

### **Non-student components**

**School personnel:** About half of evaluated programmes (45%,  $n = 34$ ) include activities or materials designed specifically for school personnel. Programmes which utilised school personnel as facilitators often provide *trainings and workshops* to prepare them to deliver the programme (32%). These trainings could not only cover the specific programme content of DRV/GBV information but could also include general skills and tips for facilitating discussion about sensitive issues with adolescents or positive classroom management strategies (e.g. Achyut *et al.* 2016;<sup>231</sup> Jewkes *et al.* 2019<sup>155</sup>). Some programmes had *manualised lesson plans or scripts* (33%) which school personnel received so as to make the programme delivery as standard and consistent across schools as possible (e.g. *Safe Dates*<sup>36,240-243</sup>). Those programme materials for school personnel sometimes included *progress tracking* checklists or other methods of charting fidelity (13%) concerning how the programme was delivered, the pace of the programme delivery and whether content was or was not included for any reason (e.g. Macgowan 1997<sup>89</sup>). A few programmes included specific access to expert support as a resource for school personnel (17%). For example, in the school coaches responsible for delivering the *CBIM* programme receive biweekly technical assistance from local advocates for the duration of the sports season, in order to review how content was progressing and address any challenging situations that might arise (Jaime *et al.* 2018<sup>245</sup>). A few programmes targeted school personnel as the recipients of the programme directly, rather than as agents to help deliver the student programme. This could include specific *trainings for school personnel* (25%) such as in the *Expect Respect* programme where teachers receive training 'to raise awareness of bullying and sexual harassment and to prepare school personnel to respond effectively to witnessed or reported incidents' (Rosenbluth *et al.* 2004, p. 214<sup>159</sup>). This could also include *materials shared with school personnel* (16%), such as in the *Me and You* programme in which a newsletter was e-mailed to all school staff with content that included 'DV types, unhealthy relationship behaviours, importance of addressing DV in schools, and the "recognize-respond-refer" model' (Peskin *et al.* 2019, p. 1421<sup>188</sup>). This type of passive material could be reviewed by school staff on their own time and saved as reference material for responding to situations in the future.

**Parent/guardian/family member:** A fifth of evaluated programmes (20%,  $n = 15$ ) include activities or materials designed specifically for parents, guardians and family members of students participating in the programme. Some programmes involved *family members as co-participants* (11%) with the students during some of the student programming. For example, the *JOVEN* curriculum<sup>193</sup> invites family members to attend the final session where students practice their newly learnt healthy communication skills in negotiating topics like curfews and dating. Similar to school personnel, family members could also be the recipients of the programme directly. A few programmes held special *activities for family members* (7%), such as an orientation held with the local parent teacher association to explain the programme to parents, encourage them to ask their students about the programme topics, and provide community resources if they need additional support.<sup>89</sup> Other programmes simply sent *materials for family members* (12%), like pamphlets, booklets, brochures or weekly newsletters, home with students for their family members to read if they choose (e.g. Ekhtiari *et al.* 2014;<sup>234</sup> Wolfe *et al.* 2009<sup>190</sup>).

**Structural environmental features:** A small number of evaluated programmes (16%,  $n = 12$ ) include specific activities and tasks that address structural-environmental elements of the school setting. These activities are concerned with improving the physical features of the school building itself to address

DRV/GBV. Programmes addressed structural environmental features in two ways. Several programmes included *visual materials in common spaces* (13%) such as posters or other artwork that could be placed in hallways, on bulletin boards, as screensavers in the computer labs, etc. Often students were involved in creating these visual materials so that they resonated more strongly with the specific school context. These materials help to alter the school climate by communicating the values of the school (anti-violence, pro-helping, etc.) to anyone who walks through school common spaces. A few programmes also included efforts to *track and monitor school spaces* (8%). Spaces with low or no adult supervision often rated as the most violence-prone and riskiest areas of a school campus.<sup>246</sup> To address this issue, programmes asked students to identify these areas and then created plans and staffing changes to support greater monitoring and patrolling of those areas (e.g. Taylor *et al.* 2013<sup>40</sup>).

**Structural social:** Just over a third of evaluated programmes (38%,  $n = 29$ ) include specific activities and tasks that seek to change the structural-social climate of the school setting. This could include a number of different efforts. Some programmes sought to improve the school climate by facilitating *service-in-reach* from local resources (16%) who could speak about the anti-violence work going on in their own community and promote local support services for survivors (e.g. Mathews *et al.* 2016;<sup>90</sup> Muck *et al.* 2018<sup>94</sup>). Other programmes required the school board or school staff to *review or create policies* (13%) for how they respond to DRV/GBV. The goal of this policy review in the *Project Respect* programme<sup>35</sup> is to strengthen or create policies that hold perpetrators accountable and support survivors in seeking help after violence. A few programmes started by having local staff and community members review the existing programme content and work together to make changes to *customise the programme material* (13%) to the specific school, community, region or country context where it was being implemented. For example, prior to the implementation of the *Jesse* prosocial video game,<sup>161</sup> a diverse group of school and community stakeholders reviewed the content to ensure it was culturally relevant and appropriate for their Caribbean context. Finally, a few programmes included the *creation of school clubs* or after-school activity groups (9%). These groups of students could sometimes be part of delivering programme materials, such as with the high school-aged 'brand ambassadors' in the *Dating Matters* programme.<sup>166</sup> Additionally, such clubs also help ensure that the programme persists beyond the time frame of the study itself by incorporating the programme aims into a more stable organisational structure at the school (i.e. special interest clubs and organisations for students).

### Intervention mechanisms of change

In *Figure 8*, we have depicted mechanisms as lower and higher level: the lower-level mechanisms are where individuals react or respond to the resources provided by the interventions that trigger change (e.g. the knowledge gained from lessons about sexual violence), and higher-level mechanisms are where mechanisms interact with one another and evolve within the school system of relationships (e.g. students interacting in positive and prosocial ways with peers).

Two key themes were generated from the theory synthesis mechanisms leading to student commitment to prosocial values and behaviours:

1. strengthening relationships between the school and surrounding communities, and within the school (between students and teachers, and between students)
2. increasing belonging in the school by student-centred approaches.

The higher-level mechanisms of change through which this happens are depicted in the logic model. These occur through interactive and experiential learning, positive and prosocial student interaction with peers, positive relationships between students and teachers, normalisation of pro-social behaviour and increased school connectedness (see *Figure 8* dark blue boxes). Lower-level mechanisms facilitating change occur through the changes in student, teacher and family knowledge, attitudes, behaviours and skills (see *Figure 8* pink/orange boxes). These are discussed below through the broader themes of 'strengthening relationships' and 'increasing belonging'.

## Strengthening relationships between school and the outside world

### *School and community*

Some multilevel interventions included community components. While these did not explicitly theorise strengthening school-community relations, they were conceived as influencing students' beliefs about the need for help and where help could be sought within their communities. Some interventions provided training for community service providers, for example *Safe Dates* offered 20 training workshops to service providers in the community. Other interventions, such as *Expect Respect*, aimed to help students and families link in with community resources and/or co-ordinated with local free counselling services (*PPM-Based Intervention for Domestic Violence*<sup>171,234</sup>). In the latter case, introducing the student to free community services was considered an enabling factor that would facilitate change by providing a resource that was considered 'essential' for the behaviour change (Ekhtiari *et al.* 2013, p. 23<sup>171</sup>).

From Markham and Aveyard (2003, p. 1217),<sup>228</sup> schools encourage students to become 'committed' by promoting 'cultural congruence between the school and the wider community'. However, strengthening relationships with the community may not be an option in some schools, particularly where the values of the school and community diverge widely. Arguably, in those communities where students experience a lack of community safety and are exposed to many forms of violence, it is conceivable that schools could seek to raise their institutional boundaries with the community. By problematising and addressing DRV and GBV, schools differentiate themselves from community norms supportive of such violence. This is illustrated by Decker *et al.*'s (2018)<sup>174</sup> discussion of sexual violence in Malawi, where the social norms that tolerate and expect silence and compliance, especially from young women, have resulted in sexual violence being endemic.

### *School and parents*

Some multicomponent and multilevel interventions had family or parent components: some provided information about the intervention via educational booklets, newsletters and presentations, but also linked parents to resources and included topic information, tips and activities. For example, the *It's Your Game* intervention's<sup>188</sup> parent component comprised two parent newsletters. These covered definitions of dating violence, signs of unhealthy relationships, strategies for enhancing parent-child communication and information on online safety and resources. In addition, parent-child homework activities<sup>188</sup> aimed to promote parent-child communication on dating expectations and relationships, and on healthy and unhealthy relationships. Some interventions provided training specifically for parents; for example, in the *Dating Matters* intervention, there were different training programmes for parents according to student grade: 'Each parenting programme taught participants skills for positive parenting and communicating effectively with their children about healthy relationships'.<sup>166</sup> (p. 16).

The JOVEN intervention was offered to student-parent dyads and parents received two group sessions focusing on effective parenting and communication. The content also addressed prosocial behaviour relating to dating violence, and understanding and working with the school system. Similarly, the Skhokho intervention had a 4-day workshop held at weekends for parents and their children to strengthen their relations and promote better communication.<sup>155</sup>

The family components of interventions were theorised as strengthening the relationships between schools and parents and aligning norms within the family to that of the school, thereby supporting and reinforcing the change and regulatory order instigated by the intervention. This is highlighted by Ekhtiari *et al.* 'Educational booklets were distributed among parents to involve them, especially mothers, in violence prevention education to their daughters and *reinforce messages* learned at the school'. (reviewer emphasis) p. 989.<sup>234</sup>

### **Strengthening relationships across the school beyond the classroom**

Some interventions aimed to strengthen relationships across the whole school system; this was particularly evident in the multicomponent, multilevel interventions such as *Respectful Relationships Education in Schools (RREIS)*<sup>148</sup> and *Expect Respect*,<sup>159,237</sup> which aimed to have strategies to build a consistent response at the individual, classroom and school-wide levels for responding effectively to bullying and sexual harassment. In addition to training the teachers delivering the curriculum and counsellors, training was provided once a semester for *all* school personnel, including bus drivers, cafeteria workers, hall monitors and office staff.<sup>159,237</sup> Preparing all school staff to respond effectively to witnessed or reported incidents of bullying and sexual harassment supported the aim of the prescribed curriculum in reducing the social acceptance of bullying and sexual harassment. In another example, school personnel received two sessions in the JOVEN intervention on mentoring youth on relationships which included a focus on recognising teen dating violence and promoting prosocial behaviours in the context of relationships.<sup>193</sup> As Espelage *et al.* (2015, p. 477)<sup>239</sup> observed '... lessons need to be reinforced outside of the classroom and outside of the lesson. All school staff need to be reinforcing the content of these programmes'.

Involving all school staff in interventions to bring about change worked towards strengthening the relations between different professional roles so that all were focused on student well-being. Arguably for students, it also increased the connection between their academic learning and broader development. Some interventions, for example *Expect Respect*, also targeted school policies and encouraged school principals and administrators to develop policy to ensure consistency in responses to incidents and reports of bullying and sexual harassment. School principals were urged to create a policy document with input from staff and present this document to school staff, students and parents.<sup>159,237</sup> However, there did not seem to be an opportunity for students to contribute to the creation of school policy documents, rather than passively 'receive' these policies.

Other interventions targeted change at the physical infrastructure of the school environment with the aim of limiting places that were conducive to violent behaviours such as bullying. In these interventions, the process of how this was achieved – consulting and listening to students – was possibly as important as the outcome, a safe physical space. In the multilevel *Shifting Boundaries* intervention,<sup>40,100</sup> the classroom intervention was matched by a building-based intervention which involved restraining orders, higher levels of security presence in safe/unsafe 'hot spots', and placement of posters in school buildings to increase awareness and reporting of dating violence to school personnel. Schools were involved in working with their students to identify the unsafe areas through 'hot spot mapping' which enabled students to communicate where potential danger spots were in the school building.

Similarly, another intervention called *PREPARE* engaged students in the development of a school safety programme via a photovoice programme.<sup>90</sup> Its aim was to '... empower students to be the driving force in improving physical, emotional and sexual safety at school and to influence school safety policy' (Mathews *et al.* 2016, p. 1824,<sup>90</sup> reviewer emphasis), and in this programme students mapped and photographed unsafe situations and places in the school. Arguably, involving students outside the classroom in issues such as school safety presented students with opportunities to contribute to school policy and have a voice in the school's decision-making process. This, in turn, could give both students and staff insights into 'each other's realities' (Markham and Aveyard 2003, p. 1216<sup>228</sup>), and potentially enhance students' attachment or connectedness to the school. It was also an opportunity for the school to reframe some of its decision-making to focus more on student perceptions.

### **Strengthening relationships between teachers and students**

Teachers were key to implementing many of the interventions with opportunities to improve the quality of their relationships and strengthen bonds with their students. This had the potential to increase student connectedness to teachers and student commitment to the school and its values.

Some of the interventions dealt with changing teachers' violent behaviours towards their students in order to achieve a safer learning environment and improve teacher–student interactions and relations. Teachers received training as part of an intervention in order to change their attitudes, beliefs, social norms and behaviour. In some cases, the school itself could be the main location where students were exposed to violence (most notably in Devries *et al.* 2017;<sup>175</sup> Jewkes *et al.* 2019<sup>155</sup>) and in South Africa, schools are often the places where students feel least safe.<sup>155</sup> Part of this can be explained by teachers using physical, sexual and emotional violence against students. The *Skhokho* intervention<sup>155</sup> trained teachers in positive discipline and classroom management as did the *Connect with Respect* intervention.<sup>132</sup> The *Good School Toolkit* also focused on improving teaching techniques, and training in non-violent methods of discipline, as part of a change of 'operational culture at the school level'.<sup>175</sup> In this context, strengthening the relationships between teachers and students was fundamentally one about changing the quality of the relationships, with teachers held accountable for respecting students' rights not to be assaulted.

Teachers also received training in some interventions to enable them to model prosocial behaviours and act as positive role models for students. This extended the teacher's role from a focus on academic education to include emotional education which, arguably, could impact the teacher–student relationship. It may not necessarily have helped strengthen relations between teachers and students but positive role-modelling by supportive teachers may have acted as a positive reinforcement for those students following prosocial behaviours. In the CBIM intervention, the coaches served as positive role models for the male athletes, modelling bystander skills of speaking up and intervening when witnessing harmful and disrespectful behaviours. Using coaches in this intervention was based on the recognition of their '... role as influential, nonparental role models render[ing] them uniquely posed to positively impact how young men think and behave' (Miller *et al.* 2012, p. 432<sup>77</sup>).

Many of the interventions were dependent on teachers delivering the specific intervention curricula, which then created greater opportunities for positive interaction. In programmes such as *Second Step*, delivered by teachers weekly or semiweekly over a period of 2 years, teachers could develop sustained and supportive relationships with students which could '... promote greater understanding of each other's values' (Markham and Aveyard 2003, p. 1216<sup>228</sup>), 'which might not occur when the majority of instructional time is spent on only academics' (Espelage *et al.* 2015, p. 466<sup>239</sup>).

This is considered in more detail in the discussion below of increasing belonging. Other interventions that involved students in aspects of the delivery of the intervention could potentially create opportunities for positive relations between students and teachers by giving students a greater insight into the teachers' realities. For example, in the *Safe Dates* intervention<sup>165</sup> students were trained to lead peer support sessions, complementing the work of the teachers; in the *Dat-e Adolescence* intervention,<sup>247</sup> two students implemented the final two sessions of a seven session programme; and in the *Good School Toolkit*, activities at each school were delivered by two staff and two student 'protagonists'.<sup>175</sup> Where teacher and student relationships became more collaborative, the distinction between teacher and student roles became more blurred. Furthermore, increasing student involvement in the curriculum, as illustrated by these examples, gave students a voice and input into their own learning. Opportunities to engage may also have increased student affiliation with the school and its values.

### **Strengthening relationships between students**

A strong subtheme of strengthening relationships was change aimed at encouraging better communication and co-operation, and positive peer bonding between students. Many of the interventions were focused on the nature of healthy and unhealthy relationships both in the context of dating and friendship and aimed to show students how to have clearer rules for their interactions and establish appropriate boundaries with one another. This concept of 'boundaries' in personal relationships, as norms guiding behaviours, was a recurring theme within the interventions and focused on provoking students to consider their relations with others. Recognising and setting

boundaries (e.g. naming harmful behaviours and warning about consequences) was central to the *IMPower* intervention.<sup>160,174</sup> The *Shifting Boundaries* intervention<sup>248</sup> had an 'interaction-based curriculum' which addressed the setting and communication of boundaries within relationships, forming healthy relationships and friendships (and the continuum of friendship and intimacy), and wanted/unwanted behaviours within relationships. The importance of setting personal limits and respecting others' limits in relationships were covered in various interventions, such as *It's Your Game ... Keeping It Real*,<sup>189</sup> *Me and You*,<sup>188</sup> the *KBEP*,<sup>186</sup> *Teen Choices*,<sup>182</sup> *Let Us Protect Our Future*<sup>176</sup> and other interventions (Lee *et al.* 2018;<sup>87</sup> Macgowan 1997<sup>89</sup>). Intervention curricula also considered the negative consequences of unhealthy relationships<sup>36,242</sup> and the legal penalties of perpetuating sexual violence.<sup>185,248</sup>

Recognising and respecting personal limits in their personal and peer relations is arguably fundamental to improving student relationships, increasing respect and encouraging prosocial behaviours. Bystander interventions also targeted changing the relationships between students and aimed to '... change the peer context, removing the bystander support that is such a critical driver of bullying and other violent behaviours'.<sup>37</sup> (p. 54)

Bystander interventions aim to change the norms that foster violence acceptance and seek to increase students' knowledge of the range of violent behaviours and how to distinguish between, for example, sexual harassment and flirting. They also seek to increase student empathy for the victim and self-efficacy to use appropriate bystander responses when students encounter situations and behaviours that could lead to violence (e.g. *Second Step* and *Expect Respect*).

The *Green Dot* bystander intervention illustrates how students themselves can play an important role in normalising prosocial behaviour and norms across student groups. Student leaders – those who were respected and emulated – were selected to receive intensive bystander training on the basis that they would diffuse prosocial behaviour and norms to other students through peer networks, '... changes in norms can be diffused rapidly within social networks through peer pressures'.<sup>238</sup> (p. 154)

The AAL intervention was similar in that it aimed to help female athletes use their influence (linked to their social status as athletes) to promote '... social norms of respect, safety and equity for all' (Harrington 2019, p. 2<sup>216</sup>). There were other interventions that recognised the role of students in diffusing behaviours to other students. For example, there was potential for 'peer-to-peer diffusion' through the peer mentoring and role-modelling that was part of the Fourth R curriculum.<sup>81,190</sup> In a similar vein, the *Benzie's and Batchie's* intervention used peer-educators to perform a play and lead a group discussion to influence students' perceptions of other people's behaviour.<sup>168</sup> However, student facilitators involved in delivering a healthy relationship programme (Chesterfield Relate) reported frustration that they were not taken seriously by their peers and experienced their disruptive behaviours in the classroom.<sup>249</sup>

### ***Increasing belonging with student-centred learning***

The second important theme related to mechanisms of action was that interventions sought to increase students' sense of belonging by involving them in the learning process using interactive and experiential learning, which encouraged a 'student-centred' framing of teaching.<sup>228</sup> However, enabling students to input to their own learning does not detract from the important role of the teacher in delivering the curricula; teachers received training in the curriculum specific to the intervention and could also receive resources to support their delivery. For example, in the *Fourth R* intervention, there were detailed lesson plans, handouts for lessons and exercises but also scope for teachers to choose activities and exercises they preferred. Training sessions explored how teachers could adapt materials, indicating that they were empowered to make lessons relevant to context. In the *Shifting Boundaries* intervention, Taylor *et al.* (2008)<sup>248</sup> report that teachers were involved in the curriculum design; their ideas were incorporated into the lessons, such as varying the pedagogy to avoid relying solely on didactic pedagogy. Thus, while the

curriculum for a particular intervention was designed to be delivered by teachers and ‘teacher-led’, this did not preclude student interactive and experiential learning.

Interactive learning was achieved with a variety of techniques such as discussions, activities, exercises, worksheets and games, all offering varying degrees of student participation. An emphasis on student interaction in the Second Step intervention was highlighted by Espelage *et al.* ‘Lessons are highly interactive, incorporating small-group discussions and activities, dyadic exercises, whole-class instruction, and individual work’ (2013, p. 181<sup>192</sup>).

And in the GEMS programme by Achyut *et al.* (2011, p. 11<sup>78</sup>) ‘the GEA [group education activities] use participatory methodologies such as role play, games, debates and discussions to engage students in meaningful and relevant interactions and reflection about key issues’.

Interactive media, videos, decision-making games, puppet shows, films, music and art projects featured and supported curricula content. Elements such as plays performed by students,<sup>250</sup> posters – with poster contests (*Safe Dates*, *Benzies and Batchies*) – and social marketing campaigns to enhance learning were ways of enabling students to use their own knowledge and skills in the learning process. Through learning to work with each other, students also had opportunities to interact in positive and prosocial ways and improve their social and communication skills. This, in turn, contributed to strengthening the relationships between students by reinforcing those aspects that aimed to alter relationships between students.

Student acquisition of skills was important for many of the interventions and different interventions had different skills foci, such as relationships, problem-solving, assertive skills, conflict management resolution and bystander skills to intervene safely and effectively. Lessons could incorporate skill-based exercises and video demonstration of skills, designed to increase acquisition and to enable students to practise new skills. The use of group and collaborative work also encouraged students to practise skills in a supportive environment. In some interventions such as *Expect Respect*, *Second Step*, and *Fourth R*, there was extensive use of scenarios and role-playing; for example, the bystander components of these interventions aimed to not only educate students on their responsibility as bystanders, but to enable them to practise the skills of intervening to help others and stop bullying behaviours.

While role-playing encouraged students to take the perspective of others and engender empathy for students victimised by bullying and violent behaviour, it was important that students were comfortable with the scenario content.<sup>105</sup> Understanding the perspective of others was also helped by reflection opportunities and ‘journaling activities’ being part of the learning process.<sup>189</sup> Facilitating student understanding of ‘multiple realities’ was embedded in the curricula of many interventions, as highlighted by Taylor *et al.* (2008, p. 10<sup>248</sup>) in relation to *Shifting Boundaries*: ‘[t]he lessons ... did not provide simple answers, or in some cases the answers at all, but rather made students struggle with subjectivity ... and ambiguity’.

Student-centred approaches had the potential to engage students in the learning process and develop their self-reflective skills ‘... to identify and understand the origins of their own and their classmates’ orientations to meaning, values, interests and expectations’ (Markham and Aveyard 2003, p. 1217<sup>228</sup>).

There were a number of interventions such as *TakeCARE* (video),<sup>38,181</sup> *Jesse* (prosocial video game),<sup>161</sup> *Green Acres High* (digital game),<sup>251</sup> and CD-ROM Educational Programme on Sexual Knowledge<sup>170</sup> that purportedly used technology to engage students to increase bystander behaviour, increase empathy for female victims of intimate partner violence, reduce dating violence and increase knowledge and attitudes towards sexual violence. The videos demonstrated helpful bystander responses and how to use healthy relationship skills, while the *Jesse* intervention allowed the student to role-play a variety of characters in the game, and so could be either experiencing or perpetrating physical and emotional violence. However, in these interventions, there was no provision for students to interact with each



other. Arguably, embedding these interventions in a context of active learning by using, for example, structured discussions facilitated by teachers and associated exercises and assignments, would have increased student engagement. The *My Voice, My Choice* intervention<sup>96</sup> is an example where technology, in the form of an immersive virtual environment, was combined with input from facilitators who modelled and demonstrated assertive resistance skills through role-play. The aim was that girls should have 'multiple' opportunities to practise skills and receive feedback on use of skills: each participant completed three virtual simulations in which the verbal sexual coercion increased in severity and after each they received feedback from the facilitator and other participants, with the possibility of repeating sessions until the participant demonstrated assertive resistance.

Where technology was used, authors often claimed that it enhanced interactivity and enjoyment (e.g. Yom and Eun 2005<sup>170</sup>), enabled immediate feedback and a 'tailored' experience. For example, in the *Me and You* intervention,<sup>188</sup> some of the computer activities such as the quizzes provided tailored answers depending on students' answers. Levesque *et al.* (2016, 2017)<sup>182,211</sup> have argued that one of the problems with existing interventions is that they are 'one-size-fits-all' and neglect important factors such as differences in dating history and history of dating violence victimisation and perpetration. Their *Teen Choices* intervention was an online programme that provided assessments and individualised guidance matched to dating history, dating violence experiences and stages of readiness for using healthy relationship skills. It had five intervention tracks to meet the unique needs of (1) high-risk victims, (2) high-risk daters, (3) low-risk daters, (4) high-risk non-daters and (5) low-risk daters. Clearly, tailored feedback to students according to dating status and risk level is responsive to student needs but arguably, there was scope to enhance *Teen Choices* further, as acknowledged by Levesque *et al.* (2016),<sup>182</sup> with teachers facilitating class discussions and student activities to encourage active learning.

### **Hypothesised intervention outcomes**

Most interventions aimed to reduce violence: some focused on DRV victimisation and perpetration, some on GBV and some on both interventions. For many of the interventions, proximal outcomes (see [Figure 1](#), light grey box) led to reduced dating-based violence and GBV through the development of prosocial skills, increased empathy, increased intent to intervene, report and seek help, and increased social and communication skills. These changes were realised through the mechanisms discussed in the previous section.

## **Discussion and propositions**

The overall theory of change for the school-based interventions suggests that where higher-level mechanisms of change (see [Figure 8](#) dark blue boxes) are targeted, there is greater potential for an intervention's outcomes to be achieved. Within the literature, there seems to be an assumption that interventions with multiple components aimed at multiple levels are more likely to be effective.<sup>252</sup> However, according to Hawe *et al.* (2009, p. 269),<sup>252</sup> there is still little understanding about how these levels 'impact the unfolding of the intervention' and affect its outcomes. In this chapter, we have illustrated how the two main lines of argument, strengthening relationships and increasing belonging, can be theorised as interacting with each other. For example, positive relations could develop between teachers and students, using student centred-learning techniques and through students collaborating in the delivery of aspects of the interventions, which increase student input to their learning as well as giving them a voice in the school system.

To better understand some of the challenges of introducing and sustaining change in schools, we can view interventions as 'events'<sup>252</sup> that aim to change a system's dynamics. Depending on how the system's dynamics are harnessed, the effects of the intervention will either be 'dissipated' or 'enhanced' (Hawe 2015, p. 310<sup>253</sup>). Drawing on insights from the theory synthesis and two example interventions, we offer some propositions in relation to the interventions targeting change at different levels and their potential for 'disrupting' the school's system.

The school-based interventions included in this review targeted change across multiple domains; we identified the student intrapersonal (apart, individual); student interpersonal (together, group); staff; structural environmental and structural social domains of mechanisms, and how these dovetailed with domains of implementation and intervention type (Table 9). Interventions that focused on the individual level – targeting students’ knowledge, skills and attitudes – were usually informed by an individual-level theory, such as a psychological theory of behaviour change. For example, the *TakeCARE* intervention,<sup>38,181</sup> a brief video bystander programme of < 30 minutes, was informed by the concept of self-efficacy and aimed to increase student knowledge and confidence to engage in bystander behaviour. There was no need for training of teachers to deliver the intervention which, according to the authors, made *TakeCARE* an ‘inexpensive’ and ‘easily scalable’ intervention. The rationale for the *TakeCARE* intervention was that the time and training demands of multilevel bystander interventions (such as *Green Dot*) ‘put them beyond the reach of many schools and school districts’ (Jouriles *et al.* 2019, p. 4<sup>181</sup>). Arguably, this intrapersonal, single level intervention was ‘not disruptive enough’<sup>254</sup> to bring about lasting change, in that only lower-level mechanisms were activated (see Figure 8 pink/orange box), meaning the intervention was likely to ‘wash out’ (Hawe *et al.* 2009, p. 270<sup>252</sup>) of the whole school system.

In contrast, the *Green Dot* bystander intervention<sup>82,238,255</sup> was a multicomponent intervention that aimed to reduce sexual violence, and related forms of interpersonal violence, as well as acceptance of sexual violence. A subset of students received intensive training on violence victimisation, perpetration, sexual harassment, stalking, partner violence and prosocial behaviours, and teachers and school administration were also trained in bystander behaviour. These students played a key role in diffusing prosocial behaviour and norms to other students through their networks, and by modelling bystander behaviour, teachers contributed to its normalisation. The change mechanisms underpinning the intervention were about developing relationships between students, and between students and teachers, to change social norms supporting violence. The authors argue that time is required for changes in norms to lead to changes in actions and behaviours.<sup>82</sup> This intervention activated higher-level change mechanisms by altering the networks of social relationships within the school system. Arguably, the *Green Dot* intervention was more likely to ‘leave a lasting footprint’ (Hawe *et al.* 2009, p. 270<sup>252</sup>) and had time to become embedded within the schools.

TABLE 9 Mechanisms of change by ecological level

Ecological levels	Mechanisms of change	Which key theme?	Intervention type
Intrapersonal (individual, apart)	Aspects of knowledge, attitudes and skills changes in individuals (students, staff or family members) that are the direct focus of the intervention and its activities (see Figure 8 pink/orange box).	Strengthening relationships	Single-component, curriculum, multicomponent, multilevel
Interpersonal (together, in a group)	Increased student-centred learning that draws on students’ own knowledge and skills to encourage experiential and interactive learning. Positive and prosocial relations develop between students with improved communication (see Figure 8 dark blue box).	Strengthening relationships	Single-component, curriculum, multicomponent, multilevel
Staff	Positive relations develop between students and teachers through curriculum delivery; teachers and school staff model prosocial behaviour within the school (see Figure 8 dark blue box).	Strengthening relationships/ increasing belonging	Curriculum, multicomponent, multilevel
Structural/ social	Teachers and school staff normalise prosocial behaviour within the school (see Figure 8 dark blue box).	Increasing belonging	Multilevel
Structural/ environmental	Students and staff working together to ensure a safer physical environment with students feeling secure in the school building (see Figure 8 dark blue box).	Weakening framing	Multilevel

In conclusion, our analysis of intervention types, components and theories suggests several propositions that will be considered in the following chapters:

- Individual-level interventions that only trigger lower-level change mechanisms are less likely to be effective and sustainable.
- Multilevel interventions triggering mechanisms across multiple levels to address the same outcome are more likely to be effective and sustainable.
- Higher-level change mechanisms will work in some schools, while lower-level mechanisms will work in all schools.
- Intervention change mechanisms appear to be similarly hypothesised in design for DRV as for GBV.

Based on the syntheses presented in this chapter, there is a case for designing and committing resources to school-based interventions that will activate higher level mechanisms across multiple levels that have the best possibility of bringing about change. Arguably, developing those interventions that are 'too modest to make a difference' (Hawe 2015, p. 266<sup>254</sup>) is unlikely to be a good use of scarce resources.



## Chapter 5 Synthesis of process evaluations (RQ2)

This chapter presents the synthesis of process evaluations to identify factors salient for understanding the implementation of school-based interventions to prevent DRV and GBV. We first summarise the appraisal of process evaluations before presenting our syntheses. This analysis drew on 137 included process evaluations reported over 161 reports. Please note that this section contains quotes from included studies, which are reproduced under the Creative Commons licence.

### Included reports

Quality appraisal ratings for the included process evaluations are reported in full in [Report Supplementary Material 1](#). Overall, the reliability of findings was rated high for 49.6%, medium for 9.4%, and low for 41% of studies, and usefulness was rated high for 16.5%, medium for 33.1%, and low for 50.4% of studies. Eighteen studies (12.9%)<sup>81,101,102,125–127,129,135,146,154,177,205,213,218,222,233,256</sup> were rated as having both high reliability and high usefulness. A third of the studies (34.5%,  $n = 48$ ) were rated as having both low reliability and usefulness. Common issues with process evaluations that were downgraded included a lack of information about steps taken to minimise bias, over-reliance on summary quantitative data and not privileging the perspectives of stakeholders.

### Synthesis of process evaluation reports

Our analysis unfolded using a multilevel framework, where key factors were identified at school, student, facilitator and intervention level. Findings are summarised in [Table 10](#) as if-then statements that relate contextual or intervention features to key issues with intervention implementation. Where

**TABLE 10** Summary of implementation syntheses

Analytical themes	Cited studies	Descriptive themes
<b>School level</b>		
School resources and infrastructure → ease of implementation and determined necessity of modification	32,33	Physical infrastructure ( $n = 20$ ) School organisational infrastructure and management ( $n = 13, 14$ )
Time constraints → impacted dosage and fidelity	49,56	Competing priorities in schedule (academic or otherwise) ( $n = 16$ ) Limitations in staff time ( $n = 27, 28$ ) Scheduling blocks ( $n = 17, 18$ )
DRV/GBV comprehension and perception → influences school and community support	42,43	Acceptance as a problem ( $n = 14$ ) Staff stigma against GBV ( $n = 3$ ) and staff stigma against LGBTQ ( $n = 5$ ) Fear of negative community perception ( $n = 8$ ) Full school support of intervention ( $n = 22$ )
		continued

TABLE 10 Summary of implementation syntheses (continued)

Analytical themes	Cited studies	Descriptive themes
<b>Student level</b>		
Intervention interactivity → improved student engagement and acceptability	43,46	High engagement for role-play, scenarios, discussion and theatre (n = 26) Poor engagement for homework and parental handouts (n = 4) Student ownership of intervention (n = 6)
Poor student attendance among certain populations → reduced intervention reach	22	Dropping out (n = 6) Truancy (n = 3) Parental permission forms (n = 7) Voluntary programming (n = 4) After-school programming (n = 5)
Positive relationship formation → improves student comfort and engagement over time	38,77	Positive relationship formation with facilitators (n = 13) Positive relationship formation with peers (n = 13) Group bonding with a shared background (n = 19)
Programme fit to student population → student and staff acceptability, student engagement and programme fidelity	54	Cultural context (n = 21) Student age (n = 14) Academic background (n = 14) Sexual identity (n = 7) Prior exposure to trauma (n = 6) Disability (n = 2)
Student immaturity and disruptive behaviours → inhibited student engagement and intervention delivery	30,31	Immature behaviours (n = 22) Male student resistance to GBV (n = 15)
<b>Facilitator level</b>		
Facilitator content knowledge → intervention fidelity and student acceptability	30,34	Variance in school-based facilitator content knowledge (n = 21, 23) Greater consistency in outside facilitator content knowledge (n = 13)
Support from 'other side' → facilitate implementation and increase facilitator comfort	27,29	Support from outside organisation on content (n = 20, 22) Support from school personnel on school logistics (n = 10)
Embedding intervention into school curriculum → increase intervention sustainability	3	Teacher-facilitators and sustainability (n = 3)
<b>Intervention level</b>		
Ease of programme delivery → increased programme fidelity	28,30	Extent of program materials → eased implementation (n = 5, 6) Adequate facilitator training → critical for implementation (n = 26, 27)
Ease of program modification → eased program implementation	19	Flexibility in timing → eased implementation (n = 8) Support and guidance on adaptations → improved fidelity to function (n = 11)

*n* is reported, this refers to the number of studies, unless two numbers are reported (e.g. *n* = 32, 34), in which case numbers refer to studies and then reports. Tables with lists of study-level references are available in [Report Supplementary Material 1](#).

## Implementation factors at the school level

At the level of the school, three main themes emerged regarding factors impacting implementation. First, existing school resources and infrastructure were shown to impact both the ease of implementation and the necessity of programme modification (*n* = 32, 33). Unsurprisingly, limitations in space, materials needed for the intervention or school organisational capacity made it difficult to implement the intervention as planned. Second, time constraints caused by both the structure of the school day and competing priorities, such as academic performance, were shown to impact both the dose delivered and fidelity of the intervention (*n* = 54, 61). Lastly, school perceptions and comprehension of DRV and GBV were shown to affect how engaged and committed school personnel were to the intervention, along with impacting the likelihood of intervention adoption in the first place (*n* = 44, 48). In instances where the school believed that DRV/GBV were problems within their student body, the commitment of school personnel to DRV/GBV prevention was high, often so much so as to compensate for any challenges caused by time or resource limitations. Ultimately, while the barriers and facilitators to implementation are divided into descriptive and analytical themes, there was significant overlap between the issues. Each of these themes is explored below.

### **School resources and infrastructure → ease of implementation and determined necessity of modification**

In 32 studies, school and programme staff cited the absence of necessary school infrastructure as both a barrier to implementation and a cause for programme modification. School infrastructure in these instances was defined in regard to physical infrastructure (building space and materials and resources) (*n* = 20) along with school organisational infrastructure and management (*n* = 13, 14). One such physical limiting factor was room availability (*n* = 9, 9), with rooms frequently limited by either number or size. These limitations in space necessitated modifications in intervention delivery, with the extent of the modification ranging in accordance with the severity of the limitation. In one example, limited rooms (and teaching staff) required that a same-gender programme be adapted to a mixed-gender programme due to the school's inability to run two simultaneous programmes.<sup>217</sup> In another study, small room sizes led to the omission of a physical activity as the activity was unsuitable for the 30 students to complete in the available classroom.<sup>257</sup> Similar themes were observed with regard to missing or non-functional materials and resources, which led facilitators to modify or entirely drop the delivery of intervention components (*n* = 12, 12). In one such case, an audio drama intended to be played on an MP3 instead had its script read by facilitators when the MP3 players were found to be broken or missing.<sup>134</sup> In regard to organisational infrastructure and management, unsurprisingly, strong pre-existing health programmes or prior experience with similar interventions aided implementation and fidelity, while intervention components being outside of existing procedures hindered their delivery (*n* = 5). Furthermore, poor organisational infrastructure reduced the already limited time available for the interventions, with poor communication leading staff to be uninformed and unprepared for training<sup>91</sup> or delayed pupil authorisation slips, leading students to be late for classroom-based interventions.<sup>130</sup> Ultimately, although on-the-ground modifications were often done with relative success (as claimed by the studies), programme staff expressed hesitancy in implementing prevention interventions in resource-poor schools due to structural limitations:

*the schools that need us the most – like the C schools [rating on the NYC DOE Report Card] ... they're in such bad shape ... Because if they're a C or a D, they've got a lot of structural problems. Not just the students being poor, they just have poor management.*<sup>81</sup>

**Time constraints → dosage and fidelity**

Fifty-four studies discussed barriers relating to school-level time constraints, which were shown to lead to curriculum modifications, impacting intervention dosage and fidelity. Limited space in the curriculum restricted the class time that was available for DRV/GBV interventions ( $n = 22, 23$ ). The time instead was prioritised for academic performance in many cases ( $n = 16$ ), such as relating to standardised testing, with school administrators believing that time out of class would impact test scores as described by the following Green Dot facilitator:

*when I talked to our control school, you know, he [the principal] was like these [results] are wonderful but, you know, test scores ... If your test scores aren't here or you don't have this much class time then you're losing funding and you're not going to do this.<sup>258</sup>*

Additionally, limitations in the time of school staff ( $n = 27, 28$ ) impacted both the time they had available to facilitate interventions ( $n = 12, 13$ ), along with the time they had available to participate in staff training ( $n = 7$ ), ultimately reducing both the dosage of intervention delivered and the attendance at staff training. Additionally, class session lengths often had to be modified to fit the block schedule of the school, along with accommodating half-days or other schedule modifications ( $n = 17, 18$ ). In some cases, these time limitations were anticipated by the intervention designers and led to the planned modification of the intervention prior to implementation ( $n = 8$ ). In other cases, or even with the aforementioned anticipated curricular modifications, dosage was reduced while the intervention was implemented, reducing intervention fidelity to varying degrees as lessons were abridged or removed ( $n = 17$ ). These limitations ultimately reduced the willingness of school personnel to implement longer interventions, instead preferring one-time lessons in order to save time:

*I don't think we can commit that amount of curriculum time to it, particularly in year 9 ... I would say whoever's organizing the package, they need to remember that everybody, so drugs awareness, smoking, tobacco-awareness you know, all the resources you can get are about five, six weeks.<sup>91</sup>*

**Dating and relationship violence/gender-based violence comprehension and perception → influences school and community support**

In addition to time and resource limitations, the school staff's comprehension and perception of DRV/GBV, both as issues more broadly and in relation to their specific student body, were cited in 42 studies as factors impacting a school's willingness to implement an intervention. A belief among school personnel that DRV and/or GBV was not a problem within their school community ( $n = 14$ ) was a major barrier to school personnel engagement, along with programme adoption in the first place. For school personnel and particularly administrators, this lack of recognition was often tied to worries about school reputation ( $n = 8$ ), as discussed below by a facilitator of one such intervention:

*any school will say we don't have a problem with bullying, we don't have a problem with violence, and I think that's just to say, you don't want it out there that your school is violent. Who wants to go to a school that's full of violence?<sup>258</sup>*

While for teen dating violence focused interventions, this manifested as a refusal to admit that DRV or GBV was a problem for their students ( $n = 5$ ), for lesbian, gay, bisexual and transgender (LGBT)-bullying focused interventions, this was both an assertion that LGBT-bullying was not a problem in their school or even that there were no LGBT students in the school ( $n = 3$ ).<sup>259</sup> Furthermore, although interventions addressing DRV and GBV could both come with political connotations, interventions addressing GBV (in regard to both LGBT-bullying and more broadly) generated more pronounced controversy than DRV prevention education alone. Specifically, the term GBV and any teaching about the gendered component of DRV alienated some male staff ( $n = 3$ ) and, as discussed in a later section, male students ( $n = 15$ ), reducing their engagement and investment in the intervention. In one school, the researcher was pulled aside by the leader to say that they were struggling with delivering the intervention to make teaching staff:



*there was almost a divide between the male and female staff about their views on it ... there was a clear distinction between who got that it's an issue and who didn't, as in like the males, sort of didn't as much.<sup>91</sup>*

Furthermore, some school personnel (both male and female) were averse to teaching LGBT-bullying due to personal homophobic views ( $n = 5$ ) or due to a fear of a negative community response ( $n = 8$ ):

*You've got to remember, especially, I'm talking about the [name of the area] community. It's a very white, working class, not so intellectual group ... The staff fear that they will be accused [by parents/community] of promoting ['transgender'] ... or that ... they are exposing their child to notions of that. Of transgenderism. Big fear there.<sup>260</sup>*

On the other hand, while negative perceptions and misunderstandings of DRV/GBV were linked to reduced staff engagement and likelihood of programme adoption, an informed comprehension of the issue and its relevance to the student was shown to have the opposite effect. In fact, comprehension of the importance of prevention education on these issues was shown in some cases to overcome other factors, such as time limitations ( $n = 10, 10$ ), that may otherwise have dissuaded school personnel from engaging with the intervention:

*Often my attitude with MVP is 'pain in the arse, pain in the arse' when it comes to organizing things together because I've got so many other things to do. But there's not a hint of resentment in anything I'm saying ... I do not mind the time because it's worthwhile.<sup>131</sup>*

One potential method of increasing DRV/GBV awareness (and thus increasing staff engagement and buy-in) was staff training, both for facilitators and for school personnel more broadly ( $n = 9$ ):

*[The training for our staff] has significantly raised awareness in groups of staff ... there is a belief that we need to do something about this. That it's such an important issue for young women and we need to make sure everyone in the school gets to know more.<sup>109</sup>*

At the same time, however, staff who had a lower level of initial knowledge, and thus could potentially benefit more greatly from the training, were less engaged in them ( $n = 4$ ). Modifications of training based on pre-existing staff knowledge and/or additional outreach for hesitant staff may be one potential way of combatting variable staff investment.<sup>204</sup>

Ultimately, a parallel fourth theme illustrated a point that is already well-established in school-based intervention literature, namely, that whole-school investment was critical for both programme adoption and successful implementation. As such, a major factor impacting implementation was how all levels of school personnel, ranging from head teachers to janitors, viewed DRV/GBV prevention ( $n = 21, 22$ ). Investment in and understanding of these issues could thus be a deciding factor in an intervention's success and was even able to overcome other barriers to investment, such as time or resource limitations.

## Implementation factors at the student level

Ninety-seven reports cited implementation factors at the student level, with the most common themes being intervention interactivity ( $n = 48, 51$ ), poor student attendance among certain populations ( $n = 22$ ), positive relationship formation ( $n = 38, 43$ ), programme fit to the student population ( $n = 56, 62$ ) and student immaturity and disruptive behaviours ( $n = 30, 31$ ).

### **Intervention interactivity → improved student engagement and acceptability**

Forty-eight studies reported that interactive components within the intervention helped improve student engagement and acceptability. Students tended to be more engaged and rated activities more

positively that were in some way interactive, whether through discussion ( $n = 10, 11$ ), role-play ( $n = 12, 13$ ) or games ( $n = 11, 12$ ), as discussed by a teacher below:

*The participatory and interactive nature of the work – not just talking or preaching at them through theatre but really involving them. The moments where the pupils were allowed to stop the action and take over were particularly effective.*<sup>130</sup>

Some studies found that, within their samples, the favourite type of activity varied by gender, specifically demonstrating a preference for role-plays among boys and a preference for discussion among girls ( $n = 2$ ). However, the small number of papers supporting this claim suggests this merits further investigation. In contrast to more interactive intervention components, more 'boring' tasks, such as homework, were found to have a lower level of student engagement ( $n = 4, 5$ ). Critically, the level of engagement and acceptability seems to rely not only on the 'fun' nature of these tasks, but on the interaction with their peers and facilitators, as shown by the fact that other activities outside the realm of normal school programming, such as media clips<sup>257</sup> or video games,<sup>251</sup> were reported to have a higher student engagement and often acceptability when followed by a discussion. Similarly, student ownership of the intervention ( $n = 9$ ), which likewise involved student interaction, was shown to increase student engagement, as discussed by a student:

*We were all happy that we would be the ones to distribute the condoms ... It was as if we belong to this team and, like, it somehow left the best impression on me ... We are now, you know, like some collective that fights together for something important.*<sup>124</sup>

### **Poor student attendance among certain populations → reduced intervention reach**

While intervention attendance was reported to be relatively high across the reviewed studies, a small percentage of students had inconsistent attendance or did not attend at all, as directly discussed in 22 studies. Numerous barriers to intervention attendance were cited, including dropping out of school ( $n = 6$ ), truancy ( $n = 3$ ) and parental permission forms ( $n = 7$ ). For voluntary programmes, attendance was frequently far more variable and often lower ( $n = 4$ ). This was especially true for after-school programmes ( $n = 5$ ) due to competing evening responsibilities ( $n = 2$ ) or a long travel time from school to home ( $n = 3$ ). Critically, the students absent from the intervention may be among those who need it most due to the perceived correlations between dropping out/truancy and DRV/GBV, as stated by a nurse below:

*At first I thought they choose the best children to come to the project until I realised it was voluntary. Basically the children that came weren't the children I wanted to see. The children that we know needed the programme were not there.*<sup>261</sup>

While ultimately those students who do not attend the intervention may be a minority of the total student body, they may be particularly vulnerable to DRV/GBV and thus may require innovative methods of outreach.

### **Positive relationship formation → improves student comfort and engagement over time**

Another major factor that operated at the level of the student, both in an intrapersonal and interpersonal sense, was how positive relationship formation, both with peers and facilitators, improved student comfort and willingness to engage in the face of embarrassing or challenging material, as cited in 38 studies. Student embarrassment and discomfort when learning about DRV/GBV was consistently seen across student populations and is thus a barrier that teachers across any school may initially encounter. At the same time, a number of studies discussed how students who were initially hesitant to participate in activities due to the private and 'awkward' nature of the topic, became increasingly comfortable and engaged over time ( $n = 23$ ):

*At first when I started group, I was kind of scared to talk to people. I didn't know anyone in group, so I kind of had to work up a trust relationship with them ... Now I can talk to anyone in group.*<sup>203</sup>

One factor that may mediate this increase in student comfort is relationship formation, both with facilitators ( $n = 13$ ) and with peers ( $n = 14$ ):

*As time passed, well, they gave me confidence ... [that] I can ask them for help, it won't be a problem. At the beginning I felt a bit embarrassed [talking to them], but afterwards, no, I would feel relaxed.*<sup>205</sup>

As a result, the fostering of positive relationships may be a key mechanism that facilitators can use to increase both student comfort and engagement. This is particularly critical for facilitators from outside organisations who will generally not have a pre-established relationship with the student body. For example, two student participants discussed how an outside facilitator built a positive relationship with them:

*He learned how we act towards each other. He'd get out there, and he'd wrestle with us.*<sup>124</sup>

*He was in the weight room, and I was like, 'Who's this bulky guy? ... This is our weight room'. And then towards the end of the season, I was like, 'This is a cool guy'. The longer he was there, the more I listened.*<sup>124</sup>

Facilitator relationship formation was shown to be enabled by a recognition of student experience and identity, as well as a willingness to be vulnerable at times and share perspectives from their own lives:

*Several students also said that when the Mexfam health educators shared their own personal experiences during the sessions it helped the course participants to open up about their own lives.*<sup>205</sup>

Similar benefits stemming from positive relationship formation were seen in peer interaction. Increased comfort and engagement due to peer social relationships were especially apparent in the case of interventions that targeted students with a shared background ( $n = 18, 19$ ). In one such example, a group of students who all had lived experience of DRV/GBV victimisation found comfort in their shared experience:

*It helped me because I kept it in for that long, for like seven years. It was just like a wound that never healed. It was shocking to see that I wasn't the only person who's been through it. It made me feel more comfortable.*<sup>218</sup>

Additional bonding was found in other group compositions as well, including LGBT students ( $n = 2$ ) or in single gender groups ( $n = 15$ ), the latter of which is discussed below by an intervention facilitator:

*The conversation with the young ladies when you separate [boys from girls] is awesome. They can talk freely and know that they're not going to be teased because usually a fellow girl won't do something like that.*<sup>262</sup>

At the same time, the question of gender composition is not so clear cut, as many students likewise preferred mixed-gender environments ( $n = 8$ ).

### **Programme fit to student population → student and staff acceptability, student engagement and programme fidelity**

Fifty-nine reports examined the extent to which interventions fit the needs of the student population as a factor impacting implementation. Critically, the fit of the intervention to student needs was shown not only to impact student and staff acceptability, but was a major cause for programme modification, both prior to and during implementation. The literature discussed several different aspects of student background that may require intervention modification, namely, cultural context ( $n = 21$ ), student age ( $n = 14$ ), background knowledge ( $n = 14$ ), sexual identity ( $n = 7$ ), prior trauma expose ( $n = 6$ ) and disability ( $n = 2$ ). In many studies, students struggled to engage with material that was not relevant to them or their experiences. This could be due to factors such as age; in the quote below, some students found materials either too young or too old for them, as discussed by one high school student:

*I thought it was more towards adult relationships than towards teenagers. Or like early adult ... There were some text messages we couldn't really relate to ... Usually, when you're an adult relationship it's like kind of different because you're with each other more.<sup>215</sup>*

For students with prior exposure to trauma, the material could be triggering and cause discomfort, leading to a decrease in student engagement, as discussed by another student: 'I found that it stirred up a lot of memories and emotions and really invaded my comfort zone, I wanted to leave after a while.'<sup>177</sup>

As a result of such disconnections between student profile and programme contents, facilitators frequently chose to modify the intervention, for example, in one study where the facilitator translated and added cultural adaptations, specifically through changing the names in a role play, to a pre-existing intervention to better suit the needs of their Spanish-speaking students.<sup>263</sup> Ultimately, the modification of interventions to improve student fit was shown to be a fairly frequent phenomenon across studies and should be expected when an intervention is implemented in a new context or with a population group different to the original design/test.

### **Student immaturity and disruptive behaviours → inhibited student engagement and intervention delivery**

Thirty studies discussed student immaturity and disruptive behaviours as a barrier inhibiting student engagement and intervention delivery, as discussed by a teacher: 'Students [not] taking the scenarios seriously [was a challenge]. Some were laughing and making jokes because they had a hard time taking it seriously and being vulnerable to the rest of the class'.<sup>81</sup>

Immature behaviours were found in participants of a wide age range, ranging from elementary schoolers laughing at words such as 'private parts'<sup>264</sup> to high-schoolers bullying their peers.<sup>265</sup> One common challenge that facilitators faced was deciding how to properly respond to these behaviours, with one option being detention:

*The teacher tells me about the consent lesson. He relates that a male student responded that in such a case 'I'd give her a slap!'. The teacher says there were intakes of breath from other students and that he gave the student an hour's detention - 'I had to, it was unacceptable ...'<sup>257</sup>*

But in other cases, facilitators instead chose either to allow the student's peers to respond to their comments or to respond to it themselves.<sup>266</sup> Facilitators often had to balance zero-tolerance policies on bullying with the imperative to educate disruptive students who may need DRV/GBV prevention education more than their peers. While disruptions and joking around were described as occurring regardless of gender, some studies did link these behaviours particularly to male students, some of whom could be particularly disruptive in the case of discussions on GBV or the gendered component of DRV more broadly due to a fear of the intervention being perceived as 'anti-men' (n = 15):

*Initial reaction to the project was a lot of caution on, 'Is this a male-bashing kind of thing?', from some students ... I think it was very smart on our parts to promote it not specifically as violence against women because there's a defensiveness that happens with a lot of males.<sup>250</sup>*

As a result, several studies discussed the inclusion of a discussion on male victims in the curriculum in order to prevent the alienation of male students (n = 3).

## **Implementation factors at the facilitator level**

Sixty-seven studies discussed implementation factors operating at the level of the facilitator, including facilitator content knowledge (n = 30, 34), support from the 'other side' (n = 29), and sustainability and

school-based facilitators ( $n = 3$ ). School-based facilitators were found to have a much wider variation in content knowledge, while outside facilitators were found to have a more consistent knowledge level. Furthermore, having support from an outside violence prevention organisation as a school-based facilitator was found to aid implementation, while having support from a school contact as an outside facilitator was found to likewise aid implementation. Lastly, using school-based facilitators was identified as a way to increase the sustainability of interventions through embedding it into the school curriculum.

### **Facilitator content knowledge → intervention fidelity and student acceptability**

Thirty studies identified facilitator knowledge of DRV/GBV as a key factor impacting implementation, shown to influence both intervention fidelity and student acceptability. At the same time, facilitator content knowledge and delivery comfort were consistently cited as being key factors impacting implementation success for both school-based ( $n = 21, 23$ ) and outside facilitators ( $n = 13$ ). While both facilitators from outside organisations and facilitators from within the school were discussed as having sufficient and insufficient levels of content knowledge, variation in skill and preparation was mentioned far more frequently for school-based facilitators ( $n = 21, 23$ ), for example as discussed by the deputy head of a school:

*I think that there are some staff that are absolutely fabulous at delivering stuff like that. And then some others who should not be allowed anywhere near it. Because it can ... be quite damaging if it's not done the right way.<sup>91</sup>*

In contrast, outside-facilitators were generally described as having high content knowledge and high confidence when delivering the material ( $n = 11$ ): 'The advocate's like well-versed in all the topics that he's talking about, and our coach like would just be making it up as he went along'.<sup>178</sup>

For both school-based facilitators and outside facilitators ( $n = 5$ ), selecting the most appropriate individuals to facilitate the programme was just as critical as sufficient training of the facilitators after selection. This was particularly apparent for school-based facilitators who ranged greatly in their initial confidence and competence in delivering DRV/GBV prevention materials. Teacher type was discussed as playing a role in facilitator comfort and confidence. Generally, school-based facilitators with a health background were identified as being more comfortable with leading DRV/GBV prevention due to their increased knowledge both in terms of the material itself and the style of teaching required to teach it, (e.g. participatory and activity-based approaches) ( $n = 8, 9$ ), along with the increased ease of incorporating this material into the existing health curriculum ( $n = 2$ ), as opposed to one focused on English or Maths. On the other hand, for LGBT-bullying prevention, LGBT teachers themselves were unsurprisingly identified as being the most knowledgeable and prepared to teach the topic ( $n = 3$ ). At the same time, however, 'coming out' was discussed as being a challenge in certain work environments and may not be what is wanted by every LGBT teacher:

*[Even now,] if you're gay and you're a teacher you may still be on guard that someone's going to find out about you. I'm not in this situation now but other teachers in my school are, and in every other school ... Most gay teachers are so closeted. They're scared to death.<sup>232</sup>*

### **Support from the 'other side' → facilitate implementation and increase facilitator comfort**

For both school-based and outside facilitators, having a support structure from the 'other side' was viewed as essential in successfully implementing the intervention as discussed in 27 studies. For teachers and other school-personnel, this mainly involved support from violence prevention experts with whom they could discuss delivery and resources ( $n = 20, 22$ ). For outside facilitators who likely already had substantial knowledge on violence prevention, having support from teachers and/or other school contacts could assist in the logistics of implementation, through tasks such as managing student

behaviour or student referrals ( $n = 10$ ). On the other hand, having disengaged or unsupportive school contacts could make implementation more logistically difficult ( $n = 2$ ), as discussed in the example below in which teachers remained passive at the sidelines of the intervention rather than helping manage students as planned:

*We had to spend half our time and energy dealing with crowd control ... the whole thing we do best – talking to them on their level, not like teachers, not like they're used to – that was being totally undermined.*<sup>130</sup>

Ultimately, whether a school-based or outside facilitator was used to deliver the intervention, communication and support between the school staff and any other participating parties appeared critical for a well-implemented intervention.

### **Embedding intervention into school curriculum → increase intervention sustainability**

One less frequently mentioned but still critical theme is the importance of embedding DRV/GBV prevention into the school curriculum in order to increase sustainability long-term, a task that is more easily accomplished via the use of teacher facilitators ( $n = 3$ ) as discussed by the Start Strong Bronx staff:

*it wasn't just this extra, side thing that they did once in a while, but it was woven into the curriculum in a lot of ways. The big thing, though was, that if the teachers were trained, then it stays.*<sup>81</sup>

## **Implementation factors at the intervention level**

Twenty reports directly discussed factors operating at the level of the intervention itself, with the ease of implementation impacting both facilitator acceptability and overall intervention fidelity and efficacy. This can be divided into two subthemes, namely the ease of programme delivery ( $n = 28, 30$ ) and the ease of programme modification ( $n = 19$ ).

### **Ease of programme delivery → increased programme fidelity**

For facilitators, especially those who were school-based, the more straightforward and lesser amount of additional time the intervention took to deliver, the better the programme fidelity and staff acceptability. Because school-based facilitators often had competing demands on their time, they could have limited availability to invest in intervention delivery:

*As teachers, you're busy, you already have your own lessons, and it's like, 'geez, I got to have another one.' So, it gave you a lot of stuff to pick and choose from.*<sup>256</sup>

As such, making the process as straightforward as possible, such as through providing robust programme materials so the teachers do not need to spend time developing their own curriculum ( $n = 5, 6$ ), increased the likelihood that the intervention would be delivered as intended. Despite restrictions in teacher time, another factor found to facilitate programme implementation was teacher training, which helped ensure that teachers felt comfortable and ready to implement the intervention ( $n = 26, 27$ ).

### **Ease of programme modification → increased programme fidelity and effectiveness**

Another theme that emerged at the level of the intervention was the ease with which the programme could be modified, as discussed in 19 studies. Across the interventions, programme modifications were consistently made for all of the reasons listed above. As a result, flexibility in delivery timing ( $n = 8$ ) was highly acceptable to school personnel.<sup>106</sup> At the same time, while some modifications could greatly improve the fit of the intervention to the needs of the student population, informed by the knowledge of the facilitator, other modifications could occur with less thought, due instead,

for example, to running out of time at the end of the lesson.<sup>267</sup> As a result, providing support and guidance on modifications in the programme materials themselves was cited in several studies as a potential method of easing implementation and improving fidelity to function ( $n = 11$ ). Ultimately, programme designers needed to strike a balance between providing enough materials to ease the implementation of interventions for often overburdened facilitators and, at the same time, giving enough flexibility in the programme so that it could be appropriately modified to fit within a variety of different contexts.





## Chapter 6 Meta-analysis and narrative synthesis of intervention effectiveness (RQ3)

This chapter reports the results of our analyses of intervention effectiveness, including categorisation of outcomes, meta-analyses and tests for publication bias and moderation by country context. In this chapter, we present an overview of included outcomes in our systematic review, orient the reader to how syntheses will be presented, present the results of meta-analyses, undertake protocol-specified sensitivity and bias analyses, and summarise findings using GRADE. Summaries of the narrative syntheses are presented in this chapter, with in-depth narrative syntheses, including forest plots, in [Report Supplementary Material 1](#).

### Reports included in outcome evaluation synthesis

We included 68 randomised trials of school-based interventions to address DRV or GBV. Of included studies, 66 were randomised trials presenting comparisons of 1 or more active interventions against a control (i.e. treatment as usual or waitlist control) and 2<sup>156,166</sup> were randomised trials presenting only comparisons between active interventions. Six other trials<sup>78,94,100,155,158,200</sup> included comparisons between both active and control interventions. See [Table 34](#) in [Appendix 2](#) for a list of all reports included in the outcome evaluation synthesis and [Report Supplementary Material 1](#) for characteristics of included outcome evaluations.

### Quality assessment

Critical appraisals for the included outcome evaluations are reported in [Report Supplementary Material 1](#). Only one of the included trials<sup>35</sup> was found to be at overall low risk of bias (1.5%). The other trials were split between trials appraised as having 'some concerns' (54.4%) and those considered to be at a 'high' risk of bias (44.1%).

An overview of ratings across domains of the risk of appraisal tool is shown in [Figures 9](#) and [10](#). The main risk of bias issues in the included trials were:

- a lack of clarity as to whether allocation of clusters and participants was concealed, with most trials using simple randomisation procedures that can be open to manipulation (directional impact of bias unclear)

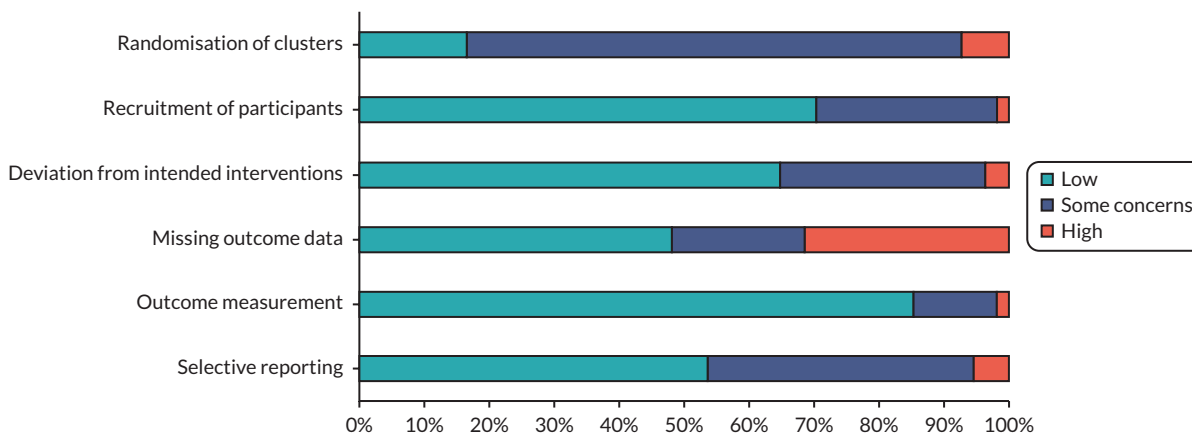
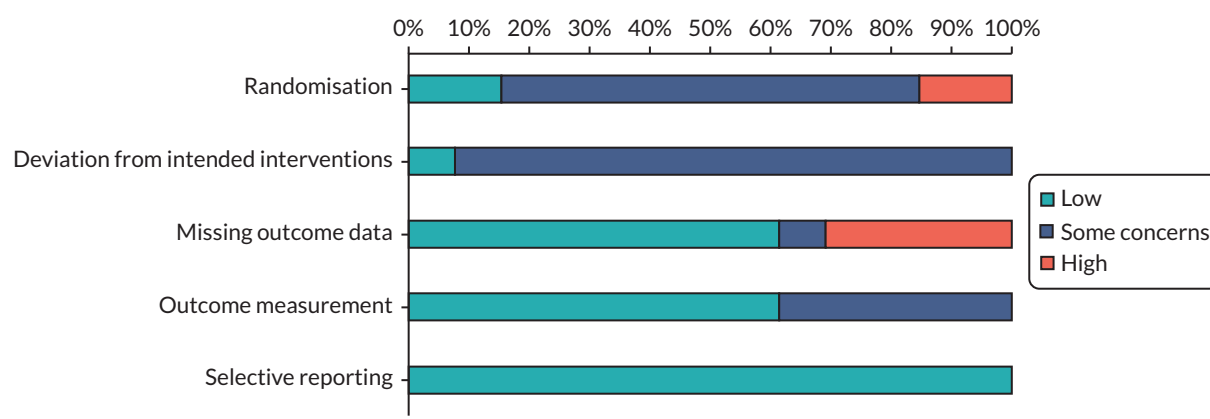


FIGURE 9 Critical appraisal ratings for included cRCTs.



**FIGURE 10** Critical appraisal ratings for included parallel RCTs.

- the potential for contamination in schools where students may mix with those in other intervention arms, or that teachers trained to deliver the intervention may alter their behaviour towards students in the control arm (bias towards null effect)
- the loss of clusters following randomisation without evidence that the reasons for drop-out were unrelated to the trial outcome (directional impact of bias unclear).

## Outcomes presented

As stated in our protocol, we inductively and hierarchically organised outcomes, starting from overarching domains of DRV perpetration and victimisation, GBV perpetration and victimisation, and knowledge, attitudes and behaviours linked to either DRV or GBV. All outcomes were grouped by time category, focusing on up to 1 year post baseline (short-term) or at 1 year or more post baseline (long-term). Our final grouping of outcomes reflected the evidence contained in the included randomised trials. Detailed extraction of effects is presented in [Report Supplementary Material 2](#).

### *Dating and relationship violence victimisation and perpetration outcomes*

Dating and relationship violence victimisation outcomes (i.e. victimisation by a dating partner, such as a boyfriend or girlfriend or intimate partner) were reported in 26 included trials (19 with short-term outcomes and 15 with long-term outcomes). Within this category, we distinguished between several DRV outcomes.

- First, we considered 'omnibus' measures of DRV victimisation, which included scales that captured multiple types of DRV victimisation. These were reported in 13 trials (11 with short-term outcomes and 6 with long-term outcomes).
- Second, we considered measures of emotional DRV victimisation, including behaviours such as insulting, demeaning or spreading rumours about a dating partner. These were reported in 14 trials (8 with short-term outcomes and 9 with long-term outcomes).
- Third, we considered measures of physical DRV victimisation, including behaviours such as hitting, kicking or slapping a dating partner, or threatening physical violence against a dating partner, but not sexual violence. These were reported in 11 trials (6 with short-term outcomes and 7 with long-term outcomes).
- Fourth, we considered measures of sexual DRV victimisation, including unwanted sexual touching, forced sex and other forms of sexual violence by a dating partner. These were reported in 12 trials (8 with short-term outcomes and 6 with long-term outcomes).
- Because several trials jointly measured physical and sexual DRV victimisation, we created a separate category for these outcomes. A total of six trials reported mixed physical and sexual DRV victimisation outcomes, of which four reported short-term and five reported long-term outcomes.

- Finally, a small number of trials reported cyber DRV victimisation outcomes, defined as dating violence via electronic means (e.g. threatening or demeaning behaviour towards a dating partner, disseminating intimate images of a dating partner). Of the four trials that reported these outcomes, three reported short-term outcomes and two reported long-term outcomes.

Classification of DRV perpetration outcomes followed a similar pattern. Of the 28 trials that reported DRV perpetration outcomes, 20 included short-term outcomes and 18 included long-term outcomes. Omnibus DRV perpetration outcomes were reported in 14 trials (12 with short-term outcomes and 8 with long-term outcomes), emotional DRV perpetration outcomes were reported in 15 trials (9 with short-term outcomes and 9 with long-term outcomes), physical DRV perpetration outcomes were reported in 14 trials (8 with short-term outcomes and 8 with long-term outcomes), sexual DRV perpetration outcomes were reported in 12 trials (9 with short-term outcomes and 5 with long-term outcomes), physical and sexual DRV perpetration outcomes were reported in 5 trials (3 with short-term outcomes and 5 with long-term outcomes) and 3 trials reported cyber DRV perpetration outcomes (2 with short-term outcomes and 2 with long-term outcomes).

### **Gender-based violence victimisation and perpetration outcomes**

Gender-based violence victimisation outcomes (i.e. violence on the basis of gender, sex or sexuality, whether physical, verbal or sexual) were reported in 20 trials. These included 15 trials reporting short-term outcomes and 13 trials reporting long-term outcomes. Within this category, we distinguished between several GBV outcomes.

- First, we considered 'omnibus' measures of GBV victimisation, which included scales that captured multiple forms of GBV victimisation. In total, 12 trials reported omnibus measures of GBV victimisation, of which 9 included short-term outcomes and 9 included long-term outcomes.
- Second, we considered measures of physical GBV victimisation, which included sexual violence of different forms, such as sexual touching and forced sex. Of the 14 trials reporting these outcomes, 11 reported short-term outcomes and 8 reported long-term outcomes.
- Third, we considered measures of verbal GBV victimisation, which included verbal sexual harassment but not homophobic harassment. Of four trials reporting this outcome, two reported short-term outcomes and three reported long-term outcomes.
- Finally, we considered measures of homophobic GBV victimisation, which included GBV on the basis of sexual orientation or gender identity. This was only reported in one trial, which provided both short-term and long-term outcomes.

GBV perpetration outcomes were similarly organised. Of the 16 trials reporting GBV perpetration outcomes, 13 reported short-term outcomes and 11 reported long-term outcomes. Of the 12 trials reporting omnibus measures of GBV perpetration, 11 reported short-term outcomes and 8 reported long-term outcomes. Four trials reported verbal GBV perpetration, of which all reported long-term outcomes, and two reported short-term outcomes as well. Two trials reported homophobic GBV perpetration, including both short-term and long-term outcomes.

### **Dating and relationship violence and gender-based violence knowledge, attitudes and behaviours**

Dating and relationship violence and GBV knowledge, attitudes and behaviours were classed into seven constructs measured for each of DRV and GBV, and one construct, bystander behaviours, that was measured across DRV and GBV.

Violence acceptance includes outcomes based on scales that measure participants' acceptance of DRV or GBV. The Illinois Rape Myth Acceptance Scale (Payne *et al.* 1999)<sup>268</sup> or the Attitudes Toward Dating Violence Scale (Price *et al.* 1999)<sup>269</sup> are archetypal scales for this construct. Of the 28 trials presenting DRV-related violence acceptance outcomes, 23 presented short-term outcomes and 8 presented

long-term outcomes. Of the 22 trials presenting GBV-related violence acceptance outcomes, 17 reported short-term outcomes and 7 reported long-term outcomes.

Knowledge includes outcomes based on knowledge or recognition of DRV or GBV, or resources to support victims or perpetrators of DRV or GBV. Scales used to measure this construct were often bespoke in nature. Of the 14 trials reporting DRV-related knowledge outcomes, all reported short-term outcomes, but only 3 reported long-term outcomes. Of the 12 trials reporting GBV-related knowledge outcomes, all but 1 reported short-term outcomes and reported long-term outcomes.

Attitudes to intervening includes outcomes based on participants' attitudes and self-efficacy to intervene when witnessing DRV or GBV; in some trials, these are described as 'bystander intentions'. Archetypal scales for this construct include the Readiness to Help Scale (Banyard *et al.* 2014)<sup>270</sup> and the Pros and Cons of Bystander Helping Scale (Edwards *et al.* 2017).<sup>271</sup> Eight trials presented DRV-related outcomes in this domain, of which all reported short-term outcomes and two reported long-term outcomes. Of the eight trials reporting GBV-related outcomes in this domain, all but one reported short-term outcomes and three reported long-term outcomes.

Attitudes to personal help-seeking includes outcomes based on participants' attitudes and intentions to seek help when experiencing DRV or GBV. While scales were primarily bespoke in nature, an indicative outcome measure is 'belief in need for help' from Safe Dates (Foshee *et al.* 1998).<sup>36</sup> Eight trials presented DRV-related outcomes in this domain, of which all reported short-term outcomes and three reported long-term outcomes. Four trials presented GBV-related outcomes in this domain, of which three reported short-term outcomes and one reported long-term outcomes.

Individual self-efficacy includes participants' intentions for, and confidence in, their ability to avoid DRV or GBV as either victim or perpetrator. Scales used to measure this construct were often bespoke in nature, but measured, for example, participants' responses in hypothetical DRV or GBV situations (Taylor *et al.* 2008, 2011).<sup>100,248</sup> Five trials reported DRV-related individual self-efficacy outcomes, and two trials reported GBV-related outcomes. All were short-term.

Social norms include outcomes related to participants' perceptions of injunctive norms and peers' behaviours relating to DRV and GBV. Scales were bespoke in nature but measured, for example, whether participants would '[believe] others would approve of responding assertively to pressure' (Arnab *et al.* 2013).<sup>272</sup> Three trials reported DRV-related social norm outcomes, of which two reported short-term outcomes and two reported long-term outcomes. Only one trial reported GBV-related social norm outcomes and these were short term.

A seventh construct that we did not meta-analyse relates to DRV and GBV-related behaviours. These related to help-seeking, harm reduction and preventive behaviours, and outcomes were heterogeneous in nature. Of the six trials reporting DRV-related behaviours in this category, five presented short-term outcomes and one presented long-term outcomes. Of the three trials reporting GBV-related behaviours in this category, two presented short-term outcomes and two presented long-term outcomes.

The eighth and final construct, bystander behaviours, was meta-analysed combining DRV and GBV estimates. This was because included bystander outcomes frequently did not distinguish between whether bystander behaviours were with respect to DRV or GBV. Bystander behaviours included proactive bystander behaviours, such as initiating conversations on DRV or GBV with peers; reactive bystander behaviours, such as intervening when witnessing abuse; and supportive bystander behaviours, such as offering support to victims of DRV or GBV. Of the nine trials presenting bystander outcomes, six presented short-term outcomes and six presented long-term outcomes.

## Syntheses presented

The results of meta-analyses undertaken in our review are summarised in this chapter and described in full detail alongside forest plots in [Report Supplementary Material 1](#). Findings are presented by outcome, intervention type and then by time category. We specifically draw on the four-category intervention typology developed in [Chapter 6](#), dividing interventions into single-component, curriculum, multicomponent and multilevel categories. Further, outcomes are organised into key domains: DRV victimisation; DRV perpetration; DRV knowledge, attitudes and behaviours; GBV victimisation; GBV perpetration; GBV knowledge, attitudes and behaviours; and bystander behaviours. Overall meta-analyses, including all trials regardless of intervention type, are presented for every outcome.

In our protocol, we stated that we would synthesise evidence by intervention, outcome and time categories. However, given the number of meta-analyses presented and studies included, and the need to improve intelligibility, we instead used a different order of categories. In addition, we undertook outcome-specific meta-analyses for DRV and GBV victimisation and perpetration, considering different forms of DRV and GBV victimisation and perpetration. We present these meta-analyses only overall rather than by intervention type given the small numbers of studies that would have been included in intervention-stratified meta-analyses. As stated in the protocol, the dominant approach to meta-analysis was robust variance estimation with random effects. However, because this method generates increasingly unstable estimates as degrees of freedom for the analysis approaches 1, we instead used a multilevel meta-analysis with between-study random effects and within-study correlations estimated at 0.8 for all meta-analyses including two studies and more than two effect sizes. For all meta-analyses with only one effect size per trial, we used a standard random-effects meta-analysis method.

Finally, the metric for all perpetration and victimisation outcomes is an OR, where values < 1 represent intervention benefits and values > 1 represent intervention harms. In contrast, the metric for all knowledge, attitude and behaviour outcomes is Cohen's *d*, where values greater than zero represent intervention benefits and values less than zero suggest intervention harms.

## Dating and relationship violence

### Dating and relationship violence victimisation

Meta-analysis results for DRV victimisation are presented in [Table 11](#). Findings collectively suggest clearer evidence of effectiveness in long-term outcomes [OR = 0.82, 95% confidence interval (CI) (0.68 to 0.99); 13 studies, 79 effect sizes] than in short-term outcomes [OR = 0.90, 95% CI (0.80 to 1.02); 17 studies, 118 effect sizes]. Evidence for effectiveness of specific intervention types or by type of DRV was generally unclear.

### Dating and relationship violence perpetration

Meta-analysis results for DRV perpetration are presented in [Table 12](#). Findings collectively suggest clearer evidence of effectiveness in long-term outcomes [OR = 0.78, 95% CI (0.64 to 0.94); 16 studies, 79 effect sizes] than in short-term outcomes [OR = 0.91, 95% CI (0.80 to 1.04); 18 studies, 118 effect sizes]. Evidence of effectiveness of specific intervention types or by type of DRV was generally unclear.

### Dating and relationship violence knowledge, attitudes and behaviours

Meta-analysis results for DRV knowledge, attitudes and behaviours are presented in [Table 13](#). Overall, interventions were effective at improving short-term violence acceptance, knowledge, attitudes to intervening and attitudes to personal help-seeking, but none of these effects persisted in long-term analyses.

## Summary of evidence for dating and relationship violence outcomes

Evidence was available for both victimisation and perpetration of multiple types of DRV. Interventions that were effective at reducing victimisation commonly showed effectiveness for perpetration

**TABLE 11** Meta-analysis results, DRV victimisation

Outcome	Intervention type	k	n	OR (95% CI)	I <sup>2</sup> (%)
<b>Up to 1 year post baseline</b>					
All outcomes	Overall	17	118	0.90 (0.80 to 1.02)	81
	Single-component	7	44	0.87 (0.64 to 1.18)	88
	Curriculum	1	8	Not estimable	
	Multicomponent	4	25	0.98 (0.78 to 1.25)	36
	Multilevel	6	41	0.87 (0.72 to 1.05)	74
Omnibus	Overall	10	45	0.88 (0.69 to 1.12)	84
Emotional	Overall	8	16	0.84 (0.55 to 1.27)	90
Physical	Overall	5	14	0.93 (0.69 to 1.25)	64
Sexual	Overall	7	29	0.97 (0.88 to 1.08)	76
Physical/sexual	Overall	4	8	0.85 (0.43 to 1.69)	76
Cyber	Overall	3	6	0.82 (0.31 to 2.16)	87
<b>At or more than 1 year post baseline</b>					
All outcomes	Overall	13	79	0.82 (0.68 to 0.99)	80
	Single-component	2	8	0.61 (0.49 to 0.77)	58
	Multicomponent	3	30	0.96 (0.85 to 1.09)	46
	Multilevel	8	41	0.83 (0.61 to 1.11)	71
Omnibus	Overall	5	12	0.85 (0.63 to 1.15)	52
Emotional	Overall	9	21	0.81 (0.59 to 1.12)	88
Physical	Overall	6	21	0.84 (0.61 to 1.16)	82
Sexual	Overall	5	13	0.88 (0.59 to 1.31)	78
Physical/sexual	Overall	5	9	0.90 (0.53 to 1.51)	73
Cyber	Overall	2	3	0.57 (0.45 to 0.72)	0
k, number of clusters; n, sample size.					

**TABLE 12** Meta-analysis results, DRV perpetration

Outcome	Intervention type	k	n	OR (95% CI)	I <sup>2</sup> (%)
<b>Up to 1 year post baseline</b>					
All outcomes	Overall	18	118	0.91 (0.80 to 1.04)	83
	Single-component	6	40	0.80 (0.47 to 1.37)	91
	Curriculum	2	5	0.83 (0.61 to 1.12)	79
	Multicomponent	6	38	0.96 (0.80 to 1.15)	46
	Multilevel	5	35	0.93 (0.88 to 0.98)	73
Omnibus	Overall	11	43	0.95 (0.85 to 1.07)	70
Emotional	Overall	9	19	0.77 (0.54 to 1.11)	90

Table 12 Meta-analysis results, DRV perpetration (continued)

Outcome	Intervention type	k	n	OR (95% CI)	I <sup>2</sup> (%)
Physical	Overall	7	16	0.91 (0.71 to 1.18)	83
Sexual	Overall	7	30	0.99 (0.86 to 1.13)	79
Physical/sexual	Overall	3	6	0.82 (0.13 to 5.29)	76
Cyber	Overall	2	4	0.96 (0.77 to 1.18)	71
<b>At or more than 1 year post baseline</b>					
All outcomes	Overall	16	79	0.78 (0.64 to 0.94)	79
	Single-component	2	8	0.51 (0.46 to 0.57)	44
	Multicomponent	5	29	0.81 (0.54 to 1.19)	82
	Multilevel	9	42	0.85 (0.66 to 1.09)	58
Omnibus	Overall	7	15	0.74 (0.52 to 1.06)	75
Emotional	Overall	9	21	0.77 (0.59 to 1.01)	85
Physical	Overall	7	22	0.83 (0.59 to 1.18)	80
Sexual	Overall	4	9	0.85 (0.37 to 1.92)	60
Physical/sexual	Overall	5	9	0.77 (0.42 to 1.43)	78
Cyber	Overall	2	3	0.49 (0.38 to 0.63)	50

k, number of clusters; n, sample size.

TABLE 13 Meta-analysis results, DRV knowledge, attitudes and behaviours

Outcome	Intervention type	k	n	d (95% CI)	I <sup>2</sup> (%)
<b>Up to 1 year post baseline</b>					
Violence acceptance	Overall	23	68	0.16 (0.08 to 0.24)	68
	Single-component	11	30	0.15 (-0.04 to 0.35)	73
	Curriculum	2	5	0.09 (-0.07 to 0.26)	0
	Multicomponent	7	21	0.19 (0.03 to 0.36)	57
	Multilevel	4	12	0.12 (-0.07 to 0.30)	61
Knowledge	Overall	14	34	0.69 (0.18 to 1.20)	96
	Single-component	6	10	0.63 (-0.41 to 1.66)	94
	Curriculum	2	12	0.96 (0.49 to 1.43)	87
	Multicomponent	3	4	0.07 (0.003 to 0.14)	26
	Multilevel	3	8	1.54 (-1.67 to 4.74)	98
Attitudes to intervening	Overall	8	31	0.14 (0.01 to 0.26)	42
	Single-component	4	16	0.25 (-0.19 to 0.70)	34
	Multicomponent	3	6	0.14 (-0.002 to 0.29)	0
	Multilevel	2	9	0.06 (-0.02 to 0.14)	0

continued

**TABLE 13** Meta-analysis results, DRV knowledge, attitudes and behaviours (*continued*)

Outcome	Intervention type	k	n	d (95% CI)	I <sup>2</sup> (%)
Attitudes to personal help-seeking	Overall	8	14	0.14 (0.06 to 0.22)	60
	Single-component	2	5	0.25 (-0.01 to 0.51)	71
	Curriculum	1	3	Not estimable	
	Multicomponent	2	3	0.19 (0.01 to 0.37)	38
	Multilevel	3	3	0.10 (0.01 to 0.20)	0
Individual self-efficacy	Overall	5	47	0.02 (-0.06 to 0.11)	60
	Single-component	3	32	0.03 (-0.26 to 0.32)	70
	Curriculum	1	9	Not estimable	
	Multicomponent	1	2	Not estimable	
	Multilevel	1	4	Not estimable	
Social norms	Overall	2	5	0.10 (-0.34 to 0.54)	74
	Single-component	1	4	Not estimable	
	Multicomponent	1	1	0.02 (-0.12 to 0.16)	
<b>At or more than 1 year post baseline</b>					
Violence acceptance	Overall	8	29	0.14 (-0.04 to 0.33)	86
	Single-component	1	1	0.22 (-0.10 to 0.54)	
	Multicomponent	1	10	Not estimable	
	Multilevel	6	18	0.20 (-0.16 to 0.57)	87
Knowledge	Overall	3	3	0.11 (-0.16 to 0.38)	85
	Multicomponent	2	2	0.04 (-0.07 to 0.14)	47
	Multilevel	1	1	2.39 (1.02 to 3.77)	
Attitudes to intervening	Overall	2	2	-0.02 (-0.18 to 0.15)	41
	Multicomponent	2	2	-0.02 (-0.18 to 0.15)	41
Attitudes to personal help-seeking	Overall	3	3	0.08 (-0.02 to 0.17)	12
	Multicomponent	1	1	0.06 (-0.15 to 0.27)	
	Multilevel	2	2	0.07 (-0.07 to 0.21)	55
Social norms	Overall	2	2	0.10 (-0.003 to 0.2)	0
	Multicomponent	1	1	0.07 (-0.09 to 0.22)	
	Multilevel	1	1	0.12 (-0.01 to 0.26)	

*d*, effect size; *k*, number of clusters; *n*, sample size.

outcomes. Effective interventions were also commonly effective across several types of DRV, though for some interventions this was not always the case. There was a great deal of heterogeneity in all analyses, with variation in effectiveness between interventions. Heterogeneity was explored and patterns are noted in the synthesis; however, in no case was heterogeneity explained satisfactorily to justify additional stratification. Factors explored for moderation of DRV outcomes are explored in [Chapter 9](#) and in [Report Supplementary Material 1](#).



Across subtypes of DRV, there was no consistent evidence that interventions were effective at reducing victimisation in the short or long term. Aside from emotional DRV, results were also unconvincing for DRV perpetration in the short term; there was a stronger signal that interventions were effective in the long term. However, all analyses showed moderate or substantial heterogeneity, and pooled effect estimates were not typically statistically significant. There was no clear evidence that one method for delivering interventions was superior for reducing DRV: effects reported for single and multilevel interventions for DRV victimisation were not repeated for perpetration and were associated with high imprecision. Overall, analyses suggest that there is variation in effectiveness across interventions that has not been resolved by intervention type.

Heterogeneity across analyses existed both between and within trials. Some variation in effect was noted across gender and abuse history, and these effects were explored further in moderation analyses (see [Chapter 9](#)). Heterogeneity was also often noted between the way outcomes were defined by studies, with reductions sometimes shown in the frequency of violent events but not in the number of students who were victimised (though this pattern of effect was not always consistent). Several interventions were shown to be effective across multiple types of DRV victimisation; however, it was noted that several interventions that were effective for physical, sexual or cyber DRV were not effective for emotional violence.

There was greater evidence of effectiveness of interventions for DRV at long-term follow-up as compared to short term. However, there was little evidence that the effectiveness of interventions changed over time. Few interventions reported both short- and long-term outcomes, and those interventions that did and showed effectiveness in the short term were also effective in the long term, and with similar effect sizes (e.g. *Teen Choices*, Levesque *et al.* 2016<sup>182</sup>). Short- and long-term analyses included different pools of interventions, and therefore it is unclear whether variation in effect between short- and long-term effects reflects an effect of time, or else heterogeneity in the evidence base. Some interventions showed variation in effect estimates across time points. However, the pattern of effect was not always in the same direction, and therefore did not suggest a clear effect of time. One exception to this is Green Dot (Coker *et al.* 2017<sup>82</sup>), where across analyses reductions in violence were found to be stable or increase in the first 2 years of follow-up, and only reduce at years 3 and 4.

## Gender-based violence

### Gender-based violence victimisation

Meta-analysis results for GBV victimisation are presented in [Table 14](#). There was no clear signal of effectiveness either in the short term [OR = 0.88, 95% CI (0.76 to 1.02); 13 studies, 72 effect sizes] or long term [OR = 0.93, 95% CI (0.80 to 1.08); 11 studies, 58 effect sizes], nor was there clear evidence of effectiveness by either intervention type or type of GBV.

### Gender-based violence perpetration

Meta-analysis results for GBV perpetration are presented in [Table 15](#). There was no clear signal of effectiveness either in the short term [OR = 0.95, 95% CI (0.85 to 1.07); 11 studies, 67 effect sizes] or the long term [OR = 0.90, 95% CI (0.73 to 1.12); 9 studies, 58 effect sizes]. There was no clear evidence of effectiveness by either intervention type or type of GBV.

### Gender-based violence knowledge, attitudes and behaviours

Meta-analysis results for GBV knowledge, attitudes and behaviours are presented in [Table 16](#). Interventions improved violence acceptance, knowledge and individual self-efficacy in the short term, with only knowledge and violence acceptance having a credible long-term effect. However, findings for long-term improvements in GBV knowledge relied on only two studies.

### Summary of evidence for gender-based violence outcomes

There was a smaller evidence base for GBV outcomes, and this prevented a more nuanced comparison of effects between types of violence. For example, outcomes pooled within verbal/psychological

**TABLE 14** Meta-analysis results, GBV victimisation

Outcome	Intervention type	<i>k</i>	<i>n</i>	OR (95% CI)	<i>I</i> <sup>2</sup> (%)
<b>Up to 1 year post baseline</b>					
All outcomes	Overall	13	72	0.88 (0.76 to 1.02)	75
	Single-component	3	31	0.85 (0.49 to 1.46)	74
	Curriculum	3	4	0.57 (0.27 to 1.22)	0
	Multicomponent	4	15	0.96 (0.90 to 1.03)	0
	Multilevel	4	22	1.00 (0.69 to 1.46)	74
Omnibus	Overall	7	29	1.00 (0.91 to 1.10)	60
Physical	Overall	9	40	0.76 (0.62 to 0.93)	78
Verbal	Overall	2	2	0.94 (0.82 to 1.08)	0
Homophobic	Overall	1	1	1.01 (0.77 to 1.33)	
<b>At or more than 1 year post baseline</b>					
All outcomes	Overall	11	58	0.93 (0.80 to 1.08)	66
	Curriculum	1	4	Not estimable	
	Multicomponent	3	37	0.86 (0.67 to 1.10)	80
	Multilevel	7	17	1.03 (0.77 to 1.36)	47
Omnibus	Overall	7	17	0.93 (0.79 to 1.10)	41
Physical	Overall	6	25	0.91 (0.68 to 1.23)	66
Verbal	Overall	3	11	0.92 (0.56 to 1.52)	76
Homophobic	Overall	1	5	Not estimable	
<i>k</i> , number of clusters; <i>n</i> , sample size.					

violence included sexual harassment, stalking, psychological victimisation and cyber abuse. Moreover, there was very little evidence for homophobic violence, and no evidence for transphobic violence. Heterogeneity was moderate or substantial across analyses; this was explored, but not resolved satisfactorily to allow for stratification. Potential factors that may moderate the effectiveness of GBV interventions are explored in [Chapter 9](#). There was no evidence for differential effectiveness according to the mode of delivery of the intervention.

Most analyses did not suggest an effect of interventions for GBV outcomes, in the short term or long term. While an effect was shown for physical GBV in the short term, this effect was not present in the long term. Effects appeared to be larger for perpetration than for victimisation for physical and verbal/psychological violence; however, these findings were also highly imprecise and therefore have a great deal of uncertainty.

There was little overlap in trials between the short- and long-term follow-up; however, interventions effective in the short term generally also showed effectiveness at long-term follow-up. A comparison of effects amongst trials with data at both short-term and long-term follow-up suggested that effects were small at later follow-up.

### **Role of knowledge and attitudes**

With regard to DRV, there was limited evidence that interventions led to improved knowledge and attitudes. Where evidence of an improvement was shown, this was in a subset of trials, and the evidence

TABLE 15 Meta-analysis results, GBV perpetration

Outcome	Intervention type	k	n	OR (95% CI)	I <sup>2</sup> (%)
<b>Up to 1 year post baseline</b>					
All outcomes	Overall	11	67	0.95 (0.85 to 1.07)	66
	Single-component	3	27	1.00 (0.91 to 1.09)	0
	Curriculum	1	2	Not estimable	0
	Multicomponent	5	19	0.93 (0.76 to 1.15)	27
	Multilevel	3	19	0.96 (0.40 to 2.28)	83
Omnibus	Overall	9	30	0.97 (0.88 to 1.06)	55
Physical	Overall	5	33	0.87 (0.62 to 1.23)	77
Verbal	Overall	2	2	0.85 (0.40 to 1.80)	76
Homophobic	Overall	2	2	1.06 (0.85 to 1.32)	0
<b>At or more than 1 year post baseline</b>					
All outcomes	Overall	9	58	0.90 (0.73 to 1.12)	67
	Curriculum	1	4	Not estimable	
	Multicomponent	4	44	0.88 (0.78 to 0.998)	60
	Multilevel	4	10	1.06 (0.54 to 2.06)	66
Omnibus	Overall	6	15	0.98 (0.73 to 1.30)	57
Physical	Overall	5	24	0.79 (0.48 to 1.28)	68
Verbal	Overall	4	12	0.86 (0.60 to 1.24)	63
Homophobic	Overall	2	7	0.95 (0.89 to 1.02)	38

*k*, number of clusters; *n*, sample size.

was greater at short- rather than long-term follow-up. The majority of attitudinal outcomes were related to perceived acceptability of violence, with fewer studies evaluating the impact of interventions on group norms, or attitudes towards help-seeking and intervening in risky situations.

Conversely, the majority of interventions for GBV showed effects in reducing violence acceptance and increasing student knowledge about GBV, though there was uncertainty in the magnitude of these effects. Effects for violence acceptance and knowledge were also greater in the short-term compared to long-term follow-up. There was limited evidence for other attitudinal outcomes, and the evidence available suggested that interventions were not effective at increasing willingness to intervene in GBV violence, or in personal help-seeking. However, there was some limited evidence that GBV interventions may increase individual self-efficacy in responding to GBV. Only one trial evaluated whether an intervention affected group norms to GBV, which did not show effectiveness.

Surprisingly, few interventions that were associated with a reduction in DRV and GBV victimisation or perpetration also reported knowledge and attitudinal outcomes. Two DRV interventions (*Teen Choices*,<sup>182</sup> Katie Brown<sup>186</sup>), each associated with reductions in DRV perpetration and victimisation, reported an improvement in attitudes towards DRV. Of these, only the evaluation of Katie Brown also measured knowledge about DRV, which was also found to increase following the intervention. Notably however, two other interventions with mixed or inconclusive evidence for DRV violence nevertheless reported an improvement in attitudes.<sup>36,49</sup> Furthermore, GBV interventions that showed improvements in attitudes to violence and in knowledge did not all also show a reduction in violence perpetration or

**TABLE 16** Meta-analysis results, GBV knowledge, attitudes and behaviours

Outcome	Intervention type	<i>k</i>	<i>n</i>	<i>d</i> (95% CI)	<i>I</i> <sup>2</sup> (%)
<b>Up to 1 year post baseline</b>					
Violence acceptance	Overall	17	57	0.29 (0.11 to 0.33)	77
	Single-component	9	28	0.20 (0.03 to 0.37)	74
	Curriculum	2	3	0.84 (0.54 to 1.14)	32
	Multicomponent	4	13	0.13 (-0.49 to 0.75)	84
	Multilevel	4	13	0.18 (0.12 to 0.24)	72
Knowledge	Overall	11	39	0.68 (0.26 to 1.11)	90
	Single-component	5	24	0.82 (-0.02 to 1.66)	94
	Curriculum	3	3	0.90 (-0.15 to 1.96)	90
	Multicomponent	2	4	0.16 (0.02 to 0.30)	0
	Multilevel	3	8	0.32 (-0.26 to 0.89)	65
Attitudes to intervening	Overall	7	33	0.10 (-0.05 to 0.24)	70
	Single-component	2	18	0.10 (0.05 to 0.15)	0
	Curriculum	2	5	0.54 (0.10 to 0.97)	39
	Multicomponent	2	3	0.00 (-0.09 to 0.09)	0
	Multilevel	2	7	0.11 (0.04 to 0.19)	48
Attitudes to personal help-seeking	Overall	3	44	0.04 (-1.36 to 1.44)	43
	Single-component	2	41	0.11 (-0.13 to 0.35)	9
	Curriculum	1	1	0.42 (-0.23 to 1.08)	
	Multicomponent	1	2	Not estimable	
Individual self-efficacy	Overall	2	9	0.16 (0.08 to 0.25)	0
	Curriculum	1	1	0.15 (-0.47 to 0.77)	
	Multicomponent	1	8	Not estimable	
Social norms	Overall	1	4	Not estimable	
	Multicomponent	1	4	Not estimable	
<b>At or more than 1 year post baseline</b>					
Violence acceptance	Overall	7	37	0.12 (0.06 to 0.19)	74
	Multicomponent	1	10	Not estimable	
	Multilevel	6	27	0.14 (0.05 to 0.22)	75
Knowledge	Overall	2	3	0.38 (0.22 to 0.54)	40
	Multilevel	2	3	0.38 (0.22 to 0.54)	40
Attitudes to intervening	Overall	3	6	0.08 (-0.30 to 0.46)	85
	Multicomponent	1	1	-0.05 (-0.15 to 0.05)	
	Multilevel	2	5	0.17 (-0.08 to 0.42)	79
Attitudes to personal help-seeking	Overall	1	2	Not estimable	
	Multilevel	1	2	Not estimable	

*d*, effect size; *k*, number of clusters; *n*, sample size.

victimisation. This suggests that changing knowledge or attitudes to violence may not be sufficient for affecting violent behaviour. The potential mediating role of attitudes and knowledge for DRV and GBV victimisation and perpetration is discussed in [Chapter 9](#).

### **Bystander outcomes**

Meta-analyses of bystander outcomes across DRV and GBV did not suggest a short-term effect [ $d = 0.17$ , 95% CI (-0.01 to 0.34),  $I^2 = 76%$ ] based on six studies and 23 effect sizes. This was similarly the case for long-term meta-analyses [ $d = 0.15$ , 95% CI (-0.32 to 0.61),  $I^2 = 89%$ ] based on six studies and 25 effect sizes.

### **Comparability of dating and relationship violence and gender-based violence evidence**

Few interventions reported both DRV and GBV outcomes. Those interventions associated with a reduction in DRV victimisation or perpetration did not report GBV outcomes, and therefore it was not possible to compare findings for GBV outcomes. Overall, current interventions were more effective at reducing DRV than GBV. This may suggest that the mechanisms involved in reducing GBV are less well understood, or else that GBV victimisation and perpetration may be less amenable to change than DRV.

## **Sensitivity and bias analyses**

Finally, the results for protocol-specific sensitivity and bias analyses are summarised in this chapter and presented in full in [Report Supplementary Material 1](#). We present tests for country contexts (HICs vs. LMICs) and publication bias for each of our priority outcome categories at the highest level (i.e. DRV victimisation, DRV perpetration, GBV victimisation, GBV perpetration; but not by intervention type or by type of DRV or GBV), and for categories of DRV-related and GBV-related knowledge, attitudes and behaviours where the number of included trials is > 10.

In our protocol, we proposed to undertake sensitivity and bias analyses considering (1) whether interventions were trialled in LMIC contexts versus HIC contexts, (2) publication bias tests and (3) tests stratifying interventions by risk of bias domains. We used a threshold of '10 studies or priority outcome category' to control multiple testing in light of the large number of meta-analyses undertaken and because subgroup analyses are only recommended when 10 or more trials are included in a meta-analysis.<sup>53</sup> This was not a criterion originally described in our protocol (which specified publication bias analyses would be undertaken if 'sufficient studies' were found, and that all meta-analyses would be tested for context differences), and is thus a clarification to the protocol. In a deviation from our protocol, we do not present tests stratifying interventions by risk of bias domains given lack of diversity in terms of risk of bias within any meta-analysis.

### **Stratification by country context**

Metaregressions comparing HIC contexts with LMIC contexts were estimated for DRV victimisation, DRV perpetration, GBV victimisation and GBV perpetration both in the short term and in the long term; and for DRV-related violence acceptance, DRV-related knowledge, GBV-related violence acceptance and GBV-related knowledge in the short term. Findings are presented in [Table 17](#), where intercept represents effectiveness in high-income contexts and coefficient represents the difference in effectiveness in low-income and middle-income contexts. Further description of findings is presented in [Report Supplementary Material 1](#). While effectiveness was comparable in the short term, the analyses showed variation in the effectiveness of interventions for long-term GBV outcomes: in the long term, interventions were associated with a reduction in GBV perpetration and victimisation in HIC contexts, but not in LMIC contexts. This may reflect the availability of resource to implement interventions, though gender inequality is also more pronounced in LMIC contexts.<sup>273</sup>

**TABLE 17** Metaregression results, country context

Outcome	Time	Term	k (n)	OR or d (95% CI)	Residual I <sup>2</sup> (%)
<b>Victimisation and perpetration outcomes</b>					
DRV victimisation	Up to 1 year post baseline	Coefficient	3 (15)	0.98 (0.6 to 1.59)	
		Intercept	14 (103)	0.9 (0.79 to 1.04)	83
	At or more than 1 year post baseline	Coefficient	3 (7)	1.24 (0.7 to 2.19)	
		Intercept	10 (72)	0.77 (0.61 to 0.98)	85
DRV perpetration	Up to 1 year post baseline	Coefficient	3 (10)	0.88 (0.21 to 3.66)	
		Intercept	15 (108)	0.92 (0.8 to 1.07)	86
	At or more than 1 year post baseline	Coefficient	3 (7)	1.5 (1.07 to 2.09)	
		Intercept	13 (72)	0.71 (0.57 to 0.89)	82
GBV victimisation	Up to 1 year post baseline	Coefficient	5 (8)	0.86 (0.53 to 1.39)	
		Intercept	8 (64)	0.93 (0.86 to 0.999)	82
	At or more than 1 year post baseline	Coefficient	6 (16)	1.24 (0.83 to 1.85)	
		Intercept	5 (42)	0.86 (0.78 to 0.94)	83
GBV perpetration	Up to 1 year post baseline	Coefficient	2 (3)	Not estimable	
		Intercept	9 (64)	0.95 (0.87 to 1.04)	59
	At or more than 1 year post baseline	Coefficient	4 (10)	1.13 (0.58 to 2.21)	
		Intercept	5 (48)	0.86 (0.77 to 0.96)	80
<b>Violence acceptance and knowledge</b>					
DRV violence acceptance	Up to 1 year post baseline	Coefficient	3 (10)	0.23 (-1.49 to 1.95)	
		Intercept	20 (58)	0.14 (0.06 to 0.22)	70
DRV knowledge	Up to 1 year post baseline	Coefficient	4 (9)	1.07 (-0.65 to 2.79)	
		Intercept	10 (25)	0.42 (0.09 to 0.76)	96
GBV violence acceptance	Up to 1 year post baseline	Coefficient	1 (2)	Not estimable	
		Intercept	16 (55)	0.19 (0.1 to 0.27)	74
GBV knowledge	Up to 1 year post baseline	Coefficient	2 (2)	0.68 (-4.29 to 5.64)	
		Intercept	9 (37)	0.57 (0.15 to 0.99)	92

d, effect size; k, number of clusters; n, sample size.

### Publication bias analyses

Publication bias analyses are undertaken for each of DRV victimisation, DRV perpetration, GBV victimisation and GBV perpetration both in the short term and in the long term, and for short-term DRV-related violence acceptance, short-term DRV-related knowledge, short-term GBV-related violence acceptance, and short-term GBV-related knowledge. Analyses for victimisation and perpetration are plotted on the log odds scale, whereas analyses for violence acceptance and knowledge are plotted on a continuous-scale parameterised by standardised mean differences. The findings are provided in [Report Supplementary Material 1](#). Overall, the analyses showed evidence of publication bias in short-term DRV victimisation, DRV perpetration and GBV victimisation, and in long-term DRV perpetration and GBV victimisation. In most cases, bias was towards publication of positive intervention effects by smaller trials, though for GBV victimisation smaller trials were more likely to report negative effects.

## GRADE analyses

Finally, we present GRADE analyses for our priority outcomes of DRV and GBV incidence to describe the strength of evidence underlying each analysis ([Table 18](#)). All outcomes were rated as low or very low certainty of evidence, due primarily to substantial unexplained heterogeneity and risk of publication bias.

## Economic evaluations

No economic evaluations were identified. As a result, we narratively synthesised cost and resource use studies identified through our searches. These are presented in [Report Supplementary Material 1](#).

TABLE 18 GRADE analyses for incidence of DRV and GBV

Certainty assessment									
No. of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Impact	Certainty	Importance
<b>DRV victimisation: short-term</b>									
17	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	Publication bias strongly suspected <sup>b</sup>	OR = 0.90 (0.80 to 1.02)	⊕○○○ Very low	Critical
<b>DRV victimisation: long-term</b>									
13	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	None	OR = 0.82 (0.68 to 0.99)	⊕⊕○○ Low	Critical
<b>DRV perpetration: short-term</b>									
18	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	Publication bias strongly suspected <sup>b</sup>	OR = 0.91 (0.80 to 1.04)	⊕○○○ Very low	Critical
<b>DRV perpetration: long-term</b>									
16	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	Publication bias strongly suspected <sup>b</sup>	OR = 0.78 (0.64 to 0.94)	⊕○○○ Very low	Critical
<b>GBV victimisation: short-term</b>									
13	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	Publication bias strongly suspected <sup>b</sup>	OR = 0.88 (0.76 to 1.02)	⊕○○○ Very low	Critical
<b>GBV victimisation: long-term</b>									
11	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	None	OR = 0.93 (0.80 to 1.08)	⊕⊕○○ Low	Critical
<b>GBV perpetration: short-term</b>									
11	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	None	OR = 0.95 (0.85 to 1.07)	⊕⊕○○ Low	Critical
<b>GBV perpetration: long-term</b>									
9	Randomised trials	Serious	Serious <sup>a</sup>	Not serious	Not serious	None	OR = 0.90 (0.73 to 1.12)	⊕⊕○○ Low	Critical

a Heterogeneity was very high, with  $I^2 > 60\%$ .

b A funnel plot suggested serious asymmetry.



## Chapter 7 Synthesis of mediation and moderation evidence (RQ3)

In this chapter, we present syntheses of mediation and moderation evidence in included outcome evaluations. Corresponding to our protocol, we used narrative synthesis to analyse mediation evidence, narrative synthesis with harvest plots where appropriate to analyse moderation evidence, and metaression to test moderation of intervention effects by equity-relevant characteristics where enough trial-level data relating to relevant characteristics were available.

As suggested in the protocol, mediation evidence was synthesised for all relevant outcomes, while moderation evidence was synthesised for priority outcomes of DRV and GBV victimisation and perpetration. Because many trials presented longitudinal tests of mediation and moderation spanning short-term and long-term time categories, we did not stratify narrative syntheses and harvest plots by time category. However, we did stratify by time category for any metaressions undertaken. In a deviation from our protocol, we organised our narrative synthesis of moderation evidence by outcome, without stratifying by intervention type. This was due to the sparseness of available evidence by intervention type, which would have made these comparisons unhelpful.

### Synthesis of mediation evidence

A total of seven papers<sup>36,181,186,238,239,241,274</sup> relating to five outcome evaluations presented mediation analyses for included outcomes.

#### *Characteristics of included mediation evidence*

Three analyses<sup>186,241,274</sup> relating to three outcome evaluations considered mediation for DRV victimisation, five<sup>36,186,238,241,274</sup> analyses relating to four outcome evaluations considered mediation for DRV perpetration, one analysis<sup>274</sup> considered mediation for GBV victimisation and three<sup>238,239,274</sup> analyses relating to two outcome evaluations considered mediation for GBV perpetration. One analysis<sup>181</sup> considered mediation for bystander outcomes.

Analytical methods used to test and estimate mediational relationships could be broadly classed as structural equation modelling-based (e.g. path analysis or conditional process analysis, used in five papers)<sup>181,186,238,239,274</sup> or causal steps-based (e.g. checking for attenuation of intervention effect on the outcome after inclusion of a mediator; used in three papers<sup>36,238,241</sup>). Three of the included papers<sup>238,241,274</sup> considered outcomes longitudinally (i.e. looking at mediation impacts over time), of which two<sup>238,274</sup> included only long-term follow-up periods and one<sup>241</sup> included both short-term and long-term follow-up periods. Five<sup>36,181,186,238,239</sup> of the included papers considered outcomes cross-sectionally (i.e. looking at end-point differences), of which three<sup>36,181,186</sup> tested mediation for short-term outcomes only and two<sup>238,239</sup> tested mediation for long-term outcomes only.

#### *Mediation of dating and relationship violence victimisation outcomes*

Of the three analyses<sup>36,186,274</sup> considering mediation of DRV victimisation, two<sup>186,274</sup> considered omnibus DRV victimisation outcomes and one<sup>36</sup> considered physical DRV victimisation. One analysis<sup>186</sup> considered only those reporting experience of dating at baseline.

Mediation of DRV victimisation outcomes via violence acceptance was assessed in two analyses.<sup>186,241</sup> In one analysis<sup>241</sup> considering longitudinal mediation in the Safe Dates trial, prescribed norms mediated the longitudinal impact of the intervention on physical DRV victimisation, based on a criterion of attenuating intervention effects to non-significance (intervention effect  $p = 0.01$  to  $p = 0.46$ ). In another analysis<sup>186</sup> considering short-term mediation in the KBEP, approval of retaliatory aggression [indirect effect  $-0.32$ ,

95% CI (-0.70 to -0.07)] and, marginally, attitudes towards dating violence [indirect effect -0.30, 95% CI (-0.67 to 0.00)] mediated impacts on total DRV victimisation.

Mediation of DRV victimisation outcomes via knowledge was also assessed in these analyses. In the longitudinal analysis of Safe Dates,<sup>36</sup> inclusion of awareness of community services as a mediator in regression models attenuated the intervention effect ( $p = 0.01$  to  $p = 0.08$ ). However, in the short-term analysis of the KBEP, the indirect effect of the intervention on DRV victimisation through knowledge outcomes was not significant [indirect effect -0.12, 95% CI (-0.94 to 0.28)].

The longitudinal analysis of Safe Dates<sup>36</sup> tested a number of additional mediators. Gender stereotyping (intervention effect  $p = 0.01$  to  $p = 0.29$ ), but not conflict management skills (intervention effect  $p = 0.01$ ) or belief in need for help (intervention effect  $p = 0.03$ ), mediated longitudinal impacts on physical DRV victimisation. Finally, the longitudinal analysis of Second Step<sup>274</sup> tested longitudinal change in school belonging over 3 years post baseline as a mediator of longitudinal change in DRV victimisation over 4–6 years post baseline, finding a non-significant indirect effect.

### **Mediation of dating and relationship violence perpetration outcomes**

Of the five analyses<sup>36,186,238,241,274</sup> considering mediation of DRV perpetration outcomes, two<sup>186,274</sup> considered omnibus DRV perpetration outcomes, three<sup>36,238,241</sup> considered emotional DRV perpetration outcomes, three<sup>36,238,241</sup> considered physical DRV perpetration outcomes and two<sup>36,241</sup> considered sexual DRV perpetration outcomes. Two<sup>36,186</sup> analyses considered subgroups defined by baseline dating experience, and one analysis<sup>36</sup> considered subgroups defined by baseline experience of DRV perpetration.

Four analyses<sup>36,186,238,241</sup> relating to three outcome evaluations tested mediation of impacts on DRV perpetration via violence acceptance. In one longitudinal analysis<sup>238</sup> testing mediation of DRV perpetration in Green Dot over 4 years, violence acceptance (defined as rape myth acceptance and acceptance of dating violence) did not attenuate intervention effects on physical DRV perpetration (intervention effect  $p < 0.0001$  to  $p = 0.0001$ ) but did attenuate intervention effects on emotional DRV perpetration (intervention effect  $p = 0.0005$  to  $p = 0.03$ ). In the short-term analysis of mediation in Safe Dates,<sup>36</sup> violence acceptance was assessed by prescribed norms, proscribed norms, perceived positive consequences of using DRV and perceived negative consequences of using DRV. Mediation by violence acceptance was assessed in the full sample for emotional DRV perpetration, physical DRV perpetration in the current relationship and sexual DRV perpetration using all mediators; for emotional DRV perpetration in the sample of baseline daters, using prescribed norms, proscribed norms and negative consequences of DRV as mediators; and for emotional DRV perpetration and sexual DRV perpetration in the sample of those with baseline experience of DRV perpetration, using negative consequences of DRV as a mediator. Prescribed norms attenuated intervention effects by more than 20%, meeting authors' criteria for mediation, in all DRV perpetration outcomes tested in the full sample. Perceived negative consequences of DRV mediated intervention impacts on sexual DRV perpetration in the sample with baseline experience of DRV perpetration. No other violence acceptance variables were found to mediate any of the DRV perpetration outcomes tested. In the longitudinal analysis of mediation in Safe Dates,<sup>241</sup> violence acceptance was measured via prescribed norms. Based on a criterion of attenuating intervention effects to non-significance, prescribed norms attenuated longitudinal impacts over 4 years post baseline for all DRV perpetration outcomes tested: emotional DRV perpetration (intervention effect  $p = 0.0005$  to  $p = 0.16$ ); physical DRV perpetration, both moderate ( $p = 0.02$  to  $p = 0.51$ ) and severe ( $p = 0.04$  to  $p = 0.09$ ) and sexual DRV perpetration ( $p = 0.04$  to  $p = 0.81$ ). Finally, in the short-term mediation analysis<sup>186</sup> of the KBEP, approval of retaliatory aggression [indirect effect -0.37, 95% CI (-0.78 to -0.10)] and attitudes towards dating violence [indirect effect -0.30, 95% CI (-0.70 to -0.02)] mediated intervention impacts on omnibus DRV perpetration.

Mediation of DRV perpetration outcomes by knowledge was assessed in three analyses<sup>36,186,241</sup> covering two outcome evaluations. In the short-term mediation analysis of Safe Dates, awareness of community services was tested as a mediator in the full sample for emotional DRV perpetration, physical DRV perpetration in the current relationship and sexual DRV perpetration; and in the sample of those with baseline DRV perpetration, for emotional DRV perpetration and sexual DRV perpetration. Awareness of community services was a mediator for all outcomes tested except for physical DRV perpetration in the current relationship. In the longitudinal mediation analysis of Safe Dates,<sup>241</sup> awareness of community services mediated intervention effects over 4 years on moderate physical DRV perpetration (intervention effect  $p = 0.02$  to  $p = 0.13$ ) and sexual DRV perpetration ( $p = 0.04$  to  $p = 0.15$ ) but not emotional DRV perpetration ( $p = 0.0005$  to  $p = 0.01$ ) or severe physical DRV perpetration ( $p = 0.04$  to  $p = 0.05$ ). In the short-term mediation analysis of the KBEP, DRV knowledge did not mediate intervention impacts on omnibus DRV perpetration [indirect effect  $-0.09$ , 95% CI  $(-0.63$  to  $0.26)$ ].

The short-term<sup>36</sup> and longitudinal<sup>241</sup> mediation analyses of Safe Dates considered several other mediators, specifically gender stereotyping, belief in the need for help and conflict management skills. Gender stereotyping was tested as a mediator in the short-term mediation analysis of Safe Dates for emotional DRV perpetration, physical DRV perpetration in the current relationship, and sexual DRV perpetration; and in the sample of those with baseline dating experience, for emotional DRV perpetration, mediating intervention effects on emotional DRV perpetration and sexual DRV perpetration only. In the longitudinal analysis of Safe Dates, gender stereotyping mediated intervention effects over 4 years for emotional DRV perpetration (intervention effect  $p = 0.0005$  to  $p = 0.06$ ), moderate physical DRV perpetration ( $p = 0.02$  to  $p = 0.19$ ) and sexual DRV perpetration ( $p = 0.04$  to  $p = 0.34$ ), but not severe physical DRV perpetration ( $p = 0.04$  to  $p = 0.05$ ). Belief in the need for help did not mediate any intervention effects in the short-term mediation analysis of Safe Dates, where it was tested in the full sample; in the longitudinal mediation analysis of Safe Dates, belief in the need for help mediated intervention effects for sexual DRV perpetration (intervention effect  $p = 0.04$  to  $p = 0.07$ ) but not emotional DRV perpetration ( $p = 0.0005$  to  $p = 0.001$ ), moderate physical DRV perpetration ( $p = 0.02$  to  $p = 0.001$ ) or severe physical DRV perpetration ( $p = 0.04$  to  $p = 0.05$ ). Finally, conflict management skills did not mediate any full-sample intervention effects in either the short-term mediation analysis or in the longitudinal mediation analysis, where intervention effects were not attenuated in any analysis (all  $p$ -values unchanged).

Bystander actions were tested as mediators in the longitudinal analysis of Green Dot.<sup>238</sup> Bystander actions, both proactive and reactive, did not mediate intervention impacts on physical DRV perpetration (intervention effect  $p < 0.0001$  to  $p = 0.0001$ ) or emotional DRV perpetration (intervention effect  $p = 0.0005$  to  $p = 0.001$ ).

Finally, the mediational analysis of Second Step<sup>274</sup> tested longitudinal change in school belonging over 3 years post baseline as a mediator of longitudinal change in DRV perpetration over 4–6 years post baseline, finding a non-significant indirect effect.

### **Mediation of gender-based violence victimisation outcomes**

Only one analysis,<sup>274</sup> relating to the trial of Second Step,<sup>192</sup> tested mediational impacts on GBV victimisation. In this analysis, longitudinal change in school belonging over 3 years post baseline was tested as a mediator of longitudinal change in sexual harassment victimisation and homophobic teasing victimisation over 4–6 years post baseline. In both cases, an indirect effect was significant at  $p < 0.10$  (indirect effect  $-0.004$ ,  $p = 0.088$  for sexual harassment victimisation; indirect effect  $-0.004$ ,  $p = 0.074$  for homophobic teasing victimisation). This indicated that the intervention reduced longitudinal increases in GBV victimisation by causing longitudinal improvements in school belonging.

### **Mediation of gender-based violence perpetration outcomes**

Of three analyses<sup>238,239,274</sup> considering mediation of GBV perpetration outcomes, all considered omnibus GBV perpetration outcomes, two<sup>239,274</sup> considered homophobic GBV and one analysis<sup>238</sup> considered both verbal GBV victimisation and physical GBV victimisation. Mediators were disparate across included analyses and no two analyses considered the same set of mediators.

A mediational analysis of Green Dot<sup>238</sup> included both longitudinal analyses and end-point difference analyses of long-term outcomes. Longitudinal analyses over 4 years considered the mediating impacts of violence acceptance (defined as rape myth acceptance and acceptance of dating violence) and bystander actions on various forms of GBV perpetration. Violence acceptance mediated all longitudinal intervention effects, defined as attenuation of intervention significance to a significance level of  $p > 0.01$ , specifically for physical GBV perpetration (defined as sexual violence, intervention effect  $p = 0.0003$  to  $p = 0.02$ ), two forms of verbal GBV perpetration (sexual harassment:  $p = 0.002$  to  $p = 0.05$ ; stalking:  $p = 0.009$  to  $p = 0.06$ ) and an omnibus measure of GBV perpetration ( $p = 0.0003$  to  $p = 0.02$ ). However, bystander actions did not mediate physical GBV perpetration (to  $p = 0.003$ ), emotional GBV perpetration defined as sexual harassment (to  $p = 0.01$ ), or an omnibus measure of GBV perpetration (to  $p = 0.001$ ). Bystander actions did mediate longitudinal effects on stalking as a form of emotional GBV perpetration (to  $p = 0.06$ ). A subsequent analysis considered mediation of end-point differences in physical GBV perpetration defined as sexual violence, examining long-term follow-up at the end of the second and fourth intervention years. None of these analyses produced significant indirect effects for violence acceptance or bystander actions.

A mediational analysis of Second Step<sup>239</sup> considering outcomes after 3 years considered whether intervention impacts on an omnibus measure of GBV perpetration and on homophobic GBV perpetration were mediated by change in delinquency behaviours. In both cases, indirect effects were substantial, and were significant for omnibus GBV perpetration (indirect effect  $-0.011$ , SE = 0.006) and marginally significant for homophobic GBV perpetration (indirect effect  $-0.041$ , SE = 0.018,  $p = 0.065$ ).

Finally, the mediational analysis of Second Step<sup>274</sup> tested longitudinal change in school belonging over 3 years post baseline as a mediator of longitudinal change in GBV perpetration over 4–6 years post baseline, finding a non-significant indirect effect for an omnibus measure of GBV perpetration (indirect effect  $-0.001$ ,  $p = 0.347$ ) but a significant indirect effect at  $p < 0.10$  for homophobic GBV perpetration (indirect effect  $-0.005$ ,  $p = 0.086$ ).

### **Mediation of bystander outcomes**

One analysis<sup>181</sup> considered mediation of short-term impacts of an intervention on bystander outcomes, measured as self-reported and observed bystander behaviours. Self-efficacy did not mediate intervention impacts on self-reported bystander behaviours but did generate a significant indirect effect for observed bystander behaviours [indirect effect. 93, 95% CI (0.46 to 19.06)].

## **Synthesis of moderation evidence: narrative synthesis and harvest plots**

A total of 26 reports of 23 outcome evaluations comparing active versus control were included in our synthesis of moderation evidence. One additional outcome evaluation presented in two reports<sup>166,275</sup> presented moderation evidence in respect of an active versus active comparison.

### **Characteristics of included moderation evidence**

Moderation analyses were classified by equity-relevant characteristics. These are presented in descending order of frequency, focusing on active versus control comparisons.

The largest category of moderator analyses covered sex, comparing boys versus girls as a binary variable (of note is that no included moderation analyses consider gender variation as opposed to sex). These

moderator analyses were reported in 22 publications<sup>36,81,82,85,94,157,164,168,175,176,182,185-187,189,190,193,240-242,276,277</sup> covering 20 outcome evaluations. Nine reports<sup>36,81,96,182,187,240-242,276</sup> from six outcome evaluations considered prior history of the outcome as a moderator (e.g. whether prior history of DRV perpetration moderates intervention impacts in reducing DRV perpetration).

Seven reports<sup>168,182,186,240-242,277</sup> from five outcome evaluations considered ethnicity as a moderator, contrasting majority populations against ethnic minority populations. A further moderator analysis<sup>189</sup> from an outcome evaluation presented stratified estimates for different ethnic minority groups, but because this was a 'majority-minority' population these analyses were incommensurate with other moderator analyses. Age, defined in various ways (e.g. biological age, grade level), was tested as a moderator in four reports<sup>168,174,182,277</sup> covering four outcome evaluations. Three reports<sup>36,188,240</sup> relating to two outcome evaluations tested dating history as a moderator of intervention impact. Two reports<sup>255,277</sup> relating to two outcome evaluations tested sexuality (i.e. sexual minority status) as a moderator of intervention impact. Acculturation<sup>185</sup> and poverty status<sup>277</sup> were tested as moderators in one outcome evaluation each.

Two reports<sup>166,275</sup> of one outcome evaluation comparing only active versus active interventions considered moderation of intervention impacts by sex.

A common issue in included analyses was the lack of a formal test for moderation, by significance of an interaction term, a Wald test or a likelihood ratio test comparing nested models. This affected 11 reports<sup>36,82,85,164,174,185,187-189,240,255</sup> of 10 outcome evaluations comparing active versus control comparisons. Of these, we were able to construct estimates of effect modification for six reports<sup>82,164,174,187,189,255</sup> covering five outcome evaluations. For the remaining five reports,<sup>36,85,185,188,240</sup> we were able to approximate direction and likely significance of interaction estimates based on additional within-trial information (e.g. information on precision and sample size, and similarity of effect estimates between groups).

### **Introducing harvest plots**

Harvest plots are used throughout this chapter to depict graphically the findings of moderation analyses. These are presented for moderation analyses of sex and of prior history of the outcome because these two moderators were the most populated within outcome categories. Remaining moderators were too sparse for harvest plots to be meaningful (e.g. because reporting of findings was 'non-significant' only).

Harvest plots are organised by moderator, with outcomes forming rows. In each row, individual bars represent trials. Bars are thus organised by whether trial evidence suggests greater impact for one group (e.g. boys), greater impact for another group (e.g. girls) or no gradient of impact. This means that analysis of each row is based on the collective weight of evidence based on the distribution of the bars as to whether the intervention is more effective for one of the two groups implied by the moderator, or whether there is on balance no evidence to suggest a definitive difference in effectiveness between groups.

Bars vary by height, colour and shading, and have a number that represents the number of relevant interaction tests for that moderator and outcome. To minimise double-counting, the number of relevant interaction tests is the number of tests for non-overlapping outcome constructs; that is where tests are undertaken for multiple types of violence and for a violence construct that is the sum of all of these types, overall tests are not reflected in the harvest plots. The *height* of bars represents the significance and direction of moderation presented by a trial. Full height bars represent significant moderation for a given outcome demonstrating greater intervention impact for one group. Three-quarters height bars represent a pattern of moderation estimates including some, but not consistent, significant evidence of greater intervention impact for one group. Half-height bars represent consistently non-significant estimates of moderation trending in one direction, and quarter-height bars are used for studies that only present non-significance as opposed to moderation estimates. The *colour* of bars represents the time frame of outcomes analysed. Blue bars represent moderation evidence from long-term outcomes, red

bars represent moderation evidence from short-term outcomes and purple bars represent moderation evidence from both short-term and long-term outcomes. Finally, the *shading* of bars represents the nature of the moderation analysis. Dark shaded bars represent formal interaction tests, whereas light shaded bars represent outcome evaluations where formal interaction tests were not presented.

### **Moderation of dating and relationship violence victimisation**

Moderation of intervention impacts on DRV victimisation was tested in 17<sup>36,81,82,94,157,182,185-189,193,240-242,255,276</sup> reports of 13 outcome evaluations. Moderators included sex, prior history of the outcome, dating history, age, ethnicity, acculturation and sexuality.

### **Sex as a moderator of dating and relationship violence victimisation outcomes**

In total, 15 reports<sup>36,81,82,94,157,182,185-187,189,193,240-242,276</sup> of 12 outcome evaluations reported moderation of DRV victimisation outcomes. Two trials<sup>82,157</sup> demonstrated a pattern of non-significant findings where moderation effects tended to favour boys. An analysis of Green Dot<sup>82</sup> presented evidence stratified by sex for long-term findings for physical and emotional DRV victimisation; we calculated interaction effects and did not find that the prevalence risk ratio for any outcome measure was significantly different between boys and girls, though the pattern of effects on balance favoured boys (five out of eight interaction tests). An analysis of subgroup effects from the Cleveland trial of law and justice and interaction-based programmes<sup>157</sup> drew on short-term evidence and included 24 tests, of which 16 were 'unique', covering omnibus measures of DRV victimisation and sexual DRV victimisation. No interaction tests were significant, but 18 of 24 tests (12-16 unique tests) suggested a greater impact on boys.

An additional 8 outcome evaluations in 11 reports<sup>36,94,182,185-187,189,193,240-242</sup> did not offer evidence of a gradient for DRV victimisation outcomes. Two of these reports<sup>189,193</sup> presented subgroup specific estimates. In the evaluation of JOVEN,<sup>193</sup> neither test of moderation over both short-term and long-term time categories yielded a significant result; however, stratified evidence suggested a greater impact of the intervention for girls on emotional DRV victimisation and a greater impact for boys on physical and sexual DRV victimisation. Nine reports<sup>36,94,182,185-187,240-242</sup> of six outcome evaluations only reported significance. Moderation analyses of Safe Dates were presented in four reports<sup>36,240-242</sup> covering short-term,<sup>36</sup> long-term,<sup>240,242</sup> and longitudinal<sup>241</sup> (both short-term and long-term) analyses of outcomes, representing 24 tests in total. In all reports, moderation tests were presented for physical DRV victimisation, sexual DRV victimisation and emotional DRV victimisation and all were reported as non-significant. A moderation test of Teen Choices<sup>182</sup> for long-term physical and sexual DRV victimisation and emotional DRV victimisation also reported both tests as non-significant. The remaining four moderation evaluations all presented evidence in respect of short-term outcomes. Moderation tests for the KBEP<sup>186</sup> included five tests, of which four were unique, covering an omnibus measure of DRV victimisation, emotional DRV victimisation and physical DRV victimisation. All tests were reported as non-significant. Three tests of moderation in the evaluation of Ending Violence,<sup>185</sup> all of which were omnibus measures of DRV victimisation, were reported as non-significant. Three tests of moderation (two unique) in the evaluation of the School Health Center Healthy Adolescent Relationships Program (SHARP)<sup>187</sup> related to cyber DRV victimisation, physical and sexual DRV victimisation, and an omnibus measure of DRV victimisation. All were reported as non-significant. Lastly, an evaluation of two programmes<sup>94</sup> against a control condition yielded two tests of sexual DRV victimisation, reporting non-significant results.

A final two evaluations<sup>81,276</sup> suggested greater impact for girls. The New York City evaluation of Fourth R<sup>81</sup> included two tests of moderation of omnibus measures of DRV victimisation. A short-term analysis yielded a numerical, but not significant, estimate of greater impact for boys, but a long-term analysis yielded a significant test ( $p < 0.05$ ) suggesting greater impact for girls. Finally, the moderation analysis for Shifting Boundaries included nine tests covering omnibus measures of DRV victimisation and sexual DRV victimisation. While all nine tests were reported as non-significant, five of six tests of sexual DRV victimisation suggested a numerically greater impact for girls. All three tests of omnibus measures were reported as non-significant only.

The resultant harvest plot for this outcome and moderator is shown in [Figure 11](#). Taking all outcomes together, the harvest plot does not provide evidence for a gradient in effectiveness on DRV victimisation outcomes by sex. Because of the size of the body of evidence, we undertook an exploratory analysis stratifying outcomes by type of DRV victimisation. This did not uncover any evidence of sex-specific moderation on different types of DRV victimisation.

### Prior history of the outcome as a moderator of dating and relationship violence victimisation outcomes

Eight reports<sup>36,81,182,187,240-242,276</sup> relating to five outcome evaluations reported prior history of DRV victimisation as a moderator of intervention effectiveness on DRV victimisation. The evaluation of Teen Choices<sup>182</sup> included two tests of long-term outcomes relating to physical and sexual DRV victimisation and emotional DRV victimisation. Both tests were significant at  $p < 0.01$  level, reflecting a substantially greater impact for those with prior history of DRV victimisation in reducing future DRV victimisation. The New York City evaluation of Fourth R<sup>81</sup> included two tests of moderation of omnibus measures of DRV victimisation. A short-term analysis yielded a numerical, but not significant, estimate of greater impact for those with no prior history of the outcome, but a long-term analysis yielded a significant test ( $p < 0.05$ ) suggesting greater impact for those with prior history. Three tests of moderation (two unique) in the evaluation of the SHARP<sup>187</sup> related to cyber DRV victimisation, physical and sexual DRV victimisation and an omnibus measure of DRV victimisation. All favoured those with prior history of DRV victimisation, with both the cyber DRV victimisation and the omnibus measure generating significant tests. Lastly, the evaluation of Shifting Boundaries<sup>276</sup> included nine tests covering omnibus measures of DRV victimisation and sexual DRV victimisation. While all nine tests were reported as non-significant, four of six tests of sexual DRV victimisation suggested a numerically greater impact for those with prior history of DRV victimisation. All three tests of omnibus measures were reported as non-significant only.

As noted above, moderation analyses of Safe Dates were presented in four reports<sup>36,240-242</sup> covering short-term,<sup>36</sup> long-term,<sup>240,242</sup> and longitudinal<sup>241</sup> (both short-term and long-term) analyses of outcomes, representing 24 tests in total. As for sex as a moderator, in all reports, moderation tests were presented for physical DRV victimisation, sexual DRV victimisation and emotional DRV victimisation. In short-term and longitudinal tests,<sup>36,241</sup> all results were reported as non-significant. However, long-term effects<sup>240</sup> at 1 year post intervention were all imputed as non-significant but favouring those with no prior history of DRV victimisation (four tests), while five out of eight tests at 4 years post intervention

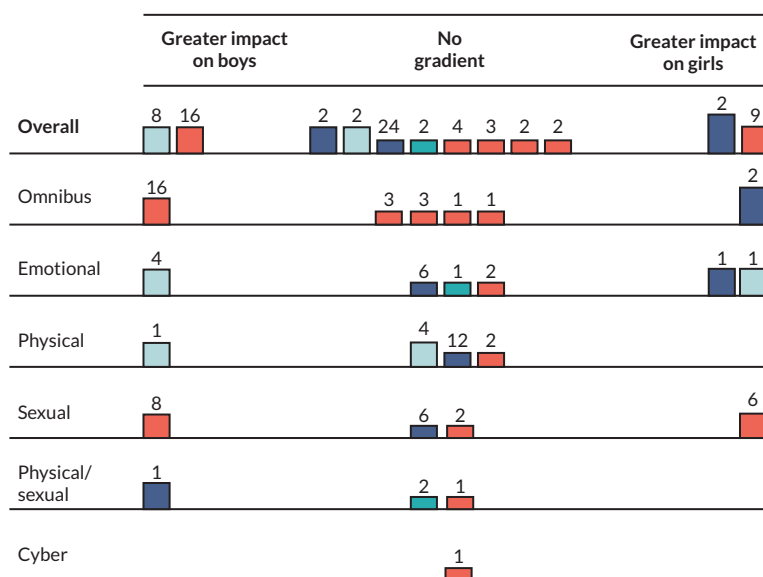
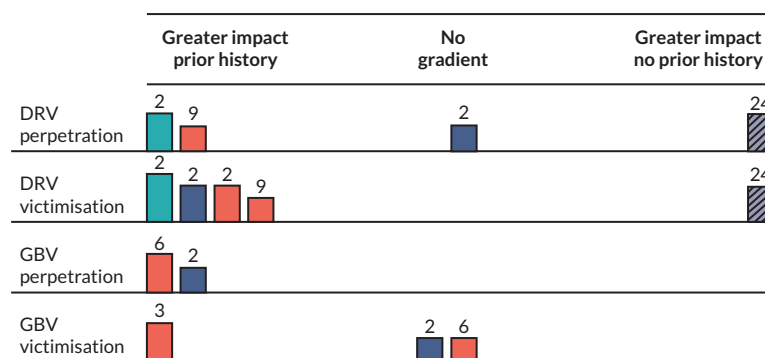


FIGURE 11 Harvest plot, DRV victimisation, boys vs. girls.



**FIGURE 12** Harvest plot, prior history of the outcome, all outcomes.

resulted in significant estimates demonstrating greater effectiveness for those with no prior history of DRV victimisation.

The resultant harvest plot for this outcome and moderator is shown in [Figure 12](#). Taking all outcomes together, the harvest plot does not provide evidence for a gradient in effectiveness on DRV victimisation outcomes by prior history of the outcome.

### Dating history as a moderator of dating and relationship violence victimisation outcomes

Three reports<sup>36,188,240</sup> relating to two outcome evaluations reported dating history as a moderator of DRV victimisation outcomes; in no report was moderation tested formally. Moderation by dating history was considered in short-term<sup>36</sup> and long-term<sup>240</sup> analyses of Safe Dates. All four short-term outcomes, covering physical DRV victimisation, sexual DRV victimisation and emotional DRV victimisation, were imputed as generating non-significant moderation. Of four long-term outcomes, three of four (physical DRV victimisation and sexual DRV victimisation) appeared to generate numerically but not statistically greater impacts on those with baseline dating history. The long-term evaluation of *It's Your Game*<sup>188</sup> compared estimates for a subsample of those with dating history at baseline to the whole sample for six outcomes (of which five were unique), covering emotional, physical, sexual and cyber DRV victimisation as well as an omnibus measure. Based on estimates of precision, there was no evidence of significant moderation, but all six outcomes showed attenuated effects for those with prior dating history, with the exception of cyber DRV.

### Age as a moderator of dating and relationship violence victimisation outcomes

The long-term evaluation<sup>182</sup> of *Teen Choices* evaluated grade level as a moderator of effects on physical DRV victimisation and emotional DRV victimisation. In both cases, results were reported as non-significant only.

### Ethnicity as a moderator of dating and relationship violence victimisation outcomes

Five reports<sup>182,186,240-242</sup> of three outcome evaluations considered ethnicity (white vs. other groups) as a moderator of DRV victimisation outcomes; all were reported as non-significant only. Ethnicity was considered as a moderator of DRV victimisation outcomes in the long-term<sup>240,242</sup> and longitudinal<sup>241</sup> evaluations of Safe Dates, across 20 tests covering emotional DRV victimisation, physical DRV victimisation and sexual DRV victimisation. The long-term evaluation<sup>182</sup> of *Teen Choices* evaluated ethnicity as a moderator of effects on physical DRV victimisation and emotional DRV victimisation. Finally, the short-term evaluation of the KBEP<sup>186</sup> evaluated five outcomes (four unique) covering an omnibus measure of DRV victimisation, emotional DRV victimisation and physical DRV victimisation.



An additional report<sup>189</sup> stratified DRV victimisation outcomes by ethnicity, comparing African American and Latinx students separately without a formal test. There was no evidence of significant moderation, but effects for physical DRV victimisation were greater for African American students and effects for emotional DRV victimisation were greater for Latinx students. In three short-term tests of omnibus measures of DRV victimisation, acculturation was not a significant moderator.

### **Acculturation as a moderator of dating and relationship violence victimisation outcomes**

One report<sup>185</sup> considered short-term moderation of DRV victimisation by acculturation in a sample of ethnic minority secondary school students. Two of three tests suggested numerically but not statistically greater effects for students with greater acculturation; the third was reported as non-significant only.

### **Sexuality as a moderator of dating and relationship violence victimisation outcomes**

The long-term evaluation of Green Dot<sup>255</sup> considered sexuality as a moderator of physical DRV victimisation and sexual DRV victimisation (specifically, reproductive coercion). While all tests were non-significant, six of eight tests suggested numerically but not statistically greater impacts for sexual majority students.

### ***Moderation of dating and relationship violence perpetration***

Moderation of intervention impacts on DRV perpetration was tested in 18 reports<sup>36,81,82,94,157,164,182,185,186,188,190,193,240-242,255,276</sup> of 14 outcome evaluations. Moderators included sex, prior history of the outcome, dating history, age, ethnicity, acculturation and sexuality.

### **Sex as a moderator of dating and relationship violence perpetration outcomes**

Sex was considered as a moderator of DRV perpetration outcomes in 16 reports<sup>36,81,82,94,157,164,182,185,186,189,190,193,240-242,276</sup> of 13 outcome evaluations. Four evaluations provided some evidence of a gradient in effectiveness favouring boys. A long-term test of the impact of Fourth R<sup>190</sup> on physical DRV perpetration found evidence of a greater impact on boys ( $p = 0.002$ ). Three further evaluations found non-significant evidence of a greater impact on boys. In JOVEN, two longitudinal tests of emotional DRV perpetration and physical and sexual DRV perpetration found numerical but not statistical evidence of a greater impact on boys. Across eight long-term tests of emotional DRV perpetration and physical DRV perpetration from the evaluation of Green Dot,<sup>82</sup> five suggested numerically greater impact on boys as compared to girls. Nine short-term tests from the evaluation of Shifting Boundaries<sup>276</sup> included three tests of an omnibus measure of DRV perpetration, all of which were reported as non-significant only, and six tests of sexual DRV perpetration, of which four numerically favoured boys, one numerically favoured girls and one did not favour either group.

A further seven evaluations<sup>36,81,94,182,185,186,189,240-242</sup> did not provide evidence of a gradient in effectiveness by sex. One evaluation<sup>189</sup> included two long-term tests: one of physical DRV perpetration, which numerically favoured girls, and one of emotional DRV perpetration, which numerically favoured boys. The other six evaluations<sup>36,81,94,182,185,186,240-242</sup> presented findings as non-significant only. The evaluation of Safe Dates<sup>36,240-242</sup> included 24 tests covering short-term,<sup>36</sup> long-term<sup>240,242</sup> and longitudinal<sup>241</sup> outcomes for emotional DRV perpetration, physical DRV perpetration and sexual DRV perpetration. The New York City evaluation of Fourth R<sup>81</sup> included one short-term and one long-term test of an omnibus measure of DRV perpetration. Two long-term tests from the evaluation of Teen Choices<sup>182</sup> included emotional DRV perpetration and physical DRV perpetration. The remaining three evaluations<sup>94,185,186</sup> presented short-term tests only: from the KBEP,<sup>186</sup> five tests, four unique, covering an omnibus measure of DRV perpetration, physical DRV perpetration and emotional DRV perpetration; from another evaluation, two short-term tests of sexual DRV perpetration;<sup>94</sup> and finally, from one evaluation,<sup>185</sup> a test of an omnibus measure of DRV perpetration.

A final two evaluations<sup>157,164</sup> provided short-term, non-significant evidence of a greater impact on girls. Across 24 tests (16 unique) including omnibus measure of DRV perpetration and sexual DRV perpetration, 21 tests indicated a numerically greater impact on girls, and three indicated a numerically greater impact on boys.<sup>157</sup> In another evaluation, intervention impacts on both emotional DRV perpetration and physical DRV perpetration showed numerically greater benefits for girls.

The resultant harvest plot for this outcome and moderator is shown in [Figure 13](#). Taking all outcomes together, the harvest plot provides some evidence of a gradient in effectiveness on DRV perpetration favouring greater impacts for boys. This appears to be primarily driven by long-term and longitudinal evidence. Because of the size of the body of evidence, we undertook an exploratory analysis stratifying outcomes by type of DRV perpetration. This suggested that greater impacts on boys were primarily driven by evidence for emotional and physical DRV perpetration.

**Prior history of the outcome as a moderator of dating and relationship violence perpetration outcomes**

Seven<sup>36,81,182,240-242,276</sup> reports of four outcome evaluations considered prior history of DRV perpetration as a moderator of DRV perpetration outcomes. Two of these reports provided some evidence of a gradient preferring those with prior history of DRV perpetration. Two long-term tests from the evaluation of Teen Choices<sup>182</sup> suggested numerically greater benefits of the intervention for those with prior history of DRV perpetration on physical DRV perpetration; these differences were significant for a test of emotional DRV perpetration. Nine short-term tests from the evaluation of Shifting Boundaries<sup>276</sup> related to omnibus measures of DRV perpetration and sexual DRV perpetration. None of the tests were significant. Three of the tests were reported as non-significant only, four were reported as showing numerically greater benefits for those with prior history of DRV perpetration, and two showed numerically greater benefits for those without a prior history of DRV perpetration. Finally, evidence from Safe Dates included short-term,<sup>36</sup> long-term<sup>240,242</sup> and longitudinal<sup>241</sup> tests relating to emotional DRV perpetration, physical DRV perpetration and sexual DRV perpetration. Only two<sup>241,242</sup> of the four reports included formal tests. In the two reports<sup>36,240</sup> that did not include formal tests, there was no clear evidence of significant moderation, though three of eight tests numerically preferred those with no prior history and five of eight tests numerically preferred those with prior history. In long-term and longitudinal evidence (16 tests), three tests were significant and indicated greater benefits for those with no prior history. The remaining 13 tests were reported as non-significant only. The resultant harvest plot is displayed in [Figure 12](#).

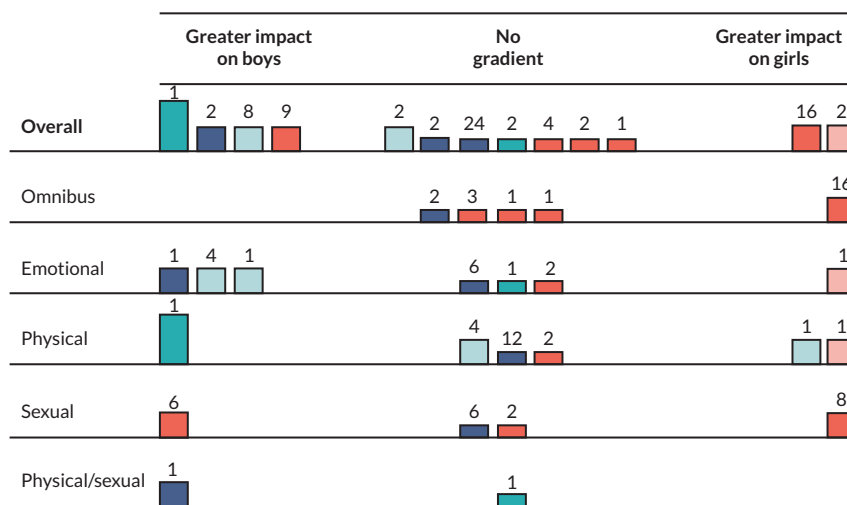


FIGURE 13 Harvest plot, DRV perpetration, boys vs. girls.

### Dating history as a moderator of dating and relationship violence perpetration outcomes

Three reports<sup>36,188,240</sup> of two outcome evaluations considered dating history as a moderator of DRV perpetration. None of the reports presented formal tests. In short-term<sup>36</sup> and long-term<sup>240</sup> evidence from Safe Dates, eight tests covered physical DRV perpetration, sexual DRV perpetration and emotional DRV perpetration. Numerical estimates for the dating subsample were similar to estimates for the entire population across all tests, suggesting no significant moderation; however, estimates numerically favoured those with dating history in five tests and did not favour either group in another test. In long-term evidence from another evaluation,<sup>188</sup> six tests (five unique) of physical DRV perpetration, sexual DRV perpetration, emotional DRV perpetration and cyber DRV perpetration, as well an omnibus measure of DRV perpetration, were presented. Evidence was equivocal, as comparing estimates for those with dating history to the entire population suggested little difference between groups. Numerically, three tests favoured those with dating history and three tests favoured those without dating history, though it was unlikely any difference was statistically significant.

### Age as a moderator of dating and relationship violence perpetration outcomes

The long-term evaluation<sup>182</sup> of *Teen Choices* evaluated grade level as a moderator of effects on physical DRV perpetration and emotional DRV perpetration. In both cases, results were reported as non-significant only.

### Ethnicity as a moderator of dating and relationship violence perpetration outcomes

Five reports<sup>182,186,240-242</sup> of three outcome evaluations considered ethnicity (white vs. other) as a moderator of DRV perpetration outcomes. All reported findings as non-significant only. This included long-term<sup>240,242</sup> and longitudinal<sup>241</sup> evidence from Safe Dates over 20 tests covering physical DRV perpetration, emotional DRV perpetration and sexual DRV perpetration; two long-term tests from Teen Choices<sup>182</sup> covering sexual DRV perpetration and emotional DRV perpetration; and five short-term tests (four unique) from the KBEP<sup>186</sup> covering emotional DRV perpetration, physical DRV perpetration and an omnibus measure of DRV perpetration.

An additional report<sup>189</sup> stratified DRV perpetration outcomes by ethnicity, comparing African American and Latinx students separately without a formal test. There was no evidence of significant moderation, but effects for physical DRV perpetration were greater for African American students and effects for emotional DRV perpetration were greater for Latinx students.

### Acculturation as a moderator of dating and relationship violence perpetration outcomes

One study<sup>185</sup> considered moderation of a short-term omnibus measure of DRV perpetration by level of acculturation. This was reported as non-significant only.

### Sexuality as a moderator of dating and relationship violence perpetration outcomes

The long-term evaluation of Green Dot<sup>255</sup> considered sexuality as a moderator of physical DRV perpetration over four tests. Two tests were non-significant but numerically favoured sexual minority students; one test was non-significant but numerically favoured sexual majority students; and one test generated significant evidence of greater impact on sexual majority students.

### Moderation of gender-based violence victimisation

Moderation of intervention impacts on GBV victimisation was tested in 13 reports<sup>81,82,85,94,96,157,168,174-176,255,276,277</sup> of 13 outcome evaluations. Moderators included sex, prior history of the outcome, age, ethnicity, sexuality and poverty status.

### Sex as a moderator of gender-based violence victimisation outcomes

Sex as a moderator of GBV victimisation outcomes was considered in 10 reports<sup>81,82,85,94,157,168,175,176,276,277</sup> of 11 outcome evaluations. Four outcome evaluations<sup>85,157,175,176</sup> suggested a greater impact for boys. One report of an intervention in South Africa<sup>176</sup> considered physical GBV victimisation (specifically, forced sex) over short-term and long-term time frames. All six tests indicated a greater impact on boys than girls, with three generating a significant result. Long-term evidence from an outcome evaluation in India<sup>85</sup> included four relevant tests of an omnibus measure of GBV. Three of four tests favoured boys, and two of these were imputed as significant. Long-term evidence from an outcome evaluation of the Good Schools Toolkit<sup>175</sup> included three tests of an omnibus measure of GBV victimisation. All were non-significant, but all three numerically favoured boys. Lastly, short-term evidence from the Cleveland evaluation of interaction-based and law and justice-based interventions<sup>157</sup> generated eight tests of moderation by sex. All were non-significant, but five numerically favoured boys.

A further five evaluations<sup>81,82,94,168,277</sup> did not offer evidence of a gradient in effectiveness by sex. The evaluation of Green Dot provided long-term evidence of moderation by sex over 24 tests relating to physical GBV victimisation and verbal GBV victimisation, 20 of which were unique. While all were non-significant, of the 20 unique tests, 2 demonstrated no numerical difference between groups, 9 suggested a numerical benefit for boys and 9 suggested a numerical benefit for girls. Of the remaining four non-unique tests, three favoured girls, though all were non-significant. A subsequent four evaluations<sup>81,94,168,277</sup> only presented evidence of non-significance. The New York City evaluation of Fourth R<sup>81</sup> included short-term and long-term tests of an omnibus measure of GBV victimisation. Another evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV victimisation and two were of verbal GBV victimisation. In the remaining two evaluations, all tests were short-term only. One evaluation<sup>168</sup> reported two tests of an omnibus measure of GBV victimisation, and one evaluation<sup>94</sup> reported two tests of physical GBV victimisation.

Finally, two evaluations suggested a greater impact on girls. One outcome evaluation<sup>85</sup> in Vietnam presented three long-term tests (two unique) of omnibus measures of GBV victimisation and of physical GBV victimisation. While all three tests numerically favoured girls, two were imputed as being statistically significant. Short-term tests from the New York City evaluation of Shifting Boundaries<sup>276</sup> included six tests of an omnibus measure of GBV victimisation. All were non-significant, but five of six yielded numerically greater benefits for girls.

The resultant harvest plot for this outcome and moderator can be seen in [Figure 14](#). On balance, the harvest plot suggests the plausibility of a gradient in intervention effects favouring boys, driven primarily by longer-term evidence. An exploratory analysis stratifying outcomes by type of GBV victimisation did not reveal a clear pattern explaining this distribution of study-level moderation findings, though longer-term evidence from omnibus measures of GBV victimisation did appear to favour greater impact on boys.

### Prior history of the outcome as a moderator of gender-based violence victimisation outcomes

Three reports<sup>81,96,276</sup> of two evaluations considered prior history of GBV victimisation as a moderator of intervention outcomes. Two evaluations were largely equivocal. The evaluation of Shifting Boundaries<sup>276</sup> included six tests of an omnibus measure of GBV victimisation, all short-term. While all tests were non-significant, three tests numerically favoured those with no prior history of GBV victimisation and three tests numerically favoured those with prior history. The New York City evaluation of Fourth R<sup>81</sup> included short-term and long-term tests of an omnibus measure of GBV victimisation. While both were non-significant, the short-term test numerically favoured those with prior history of GBV victimisation and the long-term test numerically favoured those with no prior history of the outcome. A final evaluation was delivered only to girls and included two short-term tests of physical GBV victimisation and one short-term test of verbal GBV victimisation. Both tests of physical GBV victimisation were non-significant but indicated greater impact for those with prior history of GBV victimisation, while the

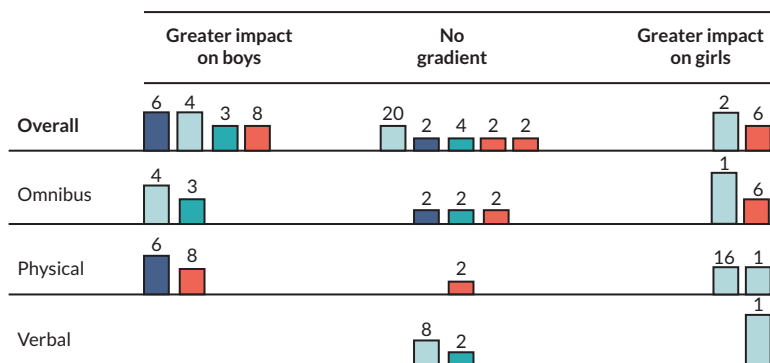


FIGURE 14 Harvest plot, GBV victimisation, boys vs. girls.

test verbal GBV victimisation was significant and indicated greater impacts for those with prior history. The resultant harvest plot is depicted in [Figure 12](#).

### Age as a moderator of gender-based violence victimisation outcomes

Three outcome evaluations<sup>168,174,277</sup> tested age as a moderator of GBV victimisation outcomes. One evaluation<sup>168</sup> considered two short-term omnibus measures of GBV victimisation by grade level and reported the moderation test as non-significant only. Another evaluation considered two short-term measures of physical GBV victimisation by school stage; while both tests were non-significant, one favoured younger students and the other favoured older students. A final evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV victimisation and two were of verbal GBV perpetration victimisation. All were reported as non-significant only.

### Ethnicity as a moderator of gender-based violence victimisation outcomes

Two outcome evaluations<sup>168,277</sup> tested ethnicity (white vs. other) as a moderator of GBV victimisation outcomes. One evaluation<sup>168</sup> considered two short-term omnibus measures of GBV victimisation and reported the moderation tests as non-significant only. Another evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV victimisation and two were of verbal GBV victimisation. All tests were reported as non-significant only.

### Sexuality as a moderator of gender-based violence victimisation outcomes

Two outcome evaluations<sup>255,277</sup> tested sexuality as a moderator of GBV victimisation outcomes. Both considered long-term effects only. The evaluation of Green Dot<sup>255</sup> tested sexuality as a moderator of physical GBV victimisation and verbal GBV victimisation for a total of 12 tests. Eight of 12 tests demonstrated numerically greater benefits for sexual majority students; of these, 1 test of a verbal GBV victimisation outcome was statistically significant. Another evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV victimisation and two were of verbal GBV victimisation. All tests were reported as non-significant only.

### Poverty status as a moderator of gender-based violence victimisation outcomes

One evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV victimisation and two were of verbal GBV victimisation. All tests were reported as non-significant only.

### Moderation of gender-based violence perpetration

Moderation of intervention impacts on GBV perpetration was tested in 10 reports<sup>81,82,85,94,157,168,176,255,276,277</sup> of 10 outcome evaluations. Moderators included sex, prior history of the outcome, age, ethnicity, sexuality and poverty status.

### Sex as a moderator of gender-based violence perpetration outcomes

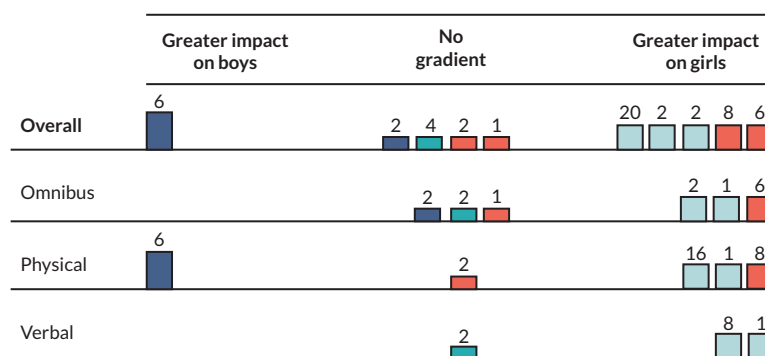
Nine reports<sup>81,82,85,94,157,168,176,276,277</sup> of 10 outcome evaluations considered sex as a moderator of GBV perpetration outcomes. One report of an intervention in South Africa<sup>176</sup> considered physical GBV

perpetration (specifically, forced sex) over short-term and long-term time frames. All six tests indicated a greater impact on boys than girls, with four generating a significant result. A further four outcome evaluations<sup>81,94,168,277</sup> reported only that results were not significant: one<sup>81</sup> including a short-term and a long-term test of an omnibus measure of GBV perpetration, one<sup>277</sup> including four long-term tests of an omnibus measure of GBV perpetration and of verbal GBV perpetration, one<sup>94</sup> including two short-term tests of physical GBV perpetration, and one<sup>168</sup> including one short-term test of an omnibus measure of GBV perpetration. A final set of five outcome evaluations<sup>82,85,157,276</sup> showed a pattern of greater effects for girls. In the long-term evaluation of Green Dot,<sup>82</sup> findings for verbal GBV perpetration (eight tests) and physical GBV perpetration (specifically, forced sex outcomes; 16 tests, of which 12 were unique) were stratified by sex. All tests were non-significant, but 5 of 8 tests for verbal GBV perpetration suggested greater impacts for girls and 9 of 16 tests for physical GBV perpetration suggested greater impacts for girls (1 test estimated identical effects). One report<sup>85</sup> covering included long-term outcomes for two outcome evaluations in Vietnam and India. Three tests (two unique) from the Vietnam outcome evaluation covered physical GBV perpetration, verbal GBV perpetration and an omnibus measure of GBV perpetration. While all three were imputed as non-significant, effects were numerically more beneficial for girls. Two tests from the India outcome evaluation covered an omnibus measure of GBV perpetration. Though both tests were imputed as non-significant, effects suggested the intervention was more beneficial for girls. Eight tests from the Cleveland evaluation<sup>157</sup> of law and justice and interaction-based interventions considered short-term physical GBV perpetration outcomes, and though all were non-significant, five of eight tests indicated numerically greater effects for girls. Finally, the evaluation of Shifting Boundaries<sup>276</sup> presented six tests of an omnibus measure of GBV perpetration; four of six tests indicated numerically greater effects for girls, though all were non-significant.

The resultant harvest plot is displayed in [Figure 15](#). A plausible interpretation of the collected evidence is that interventions may be more effective for girls than for boys, though this is countervailed by the evidence from one evaluation<sup>176</sup> featuring short-term and long-term follow-up and the relatively high proportion of evaluations reporting non-significance only. As an exploratory analysis, we stratified outcomes by type of GBV victimisation. This suggests that evidence for a gradient favouring greater impact on girls is strongest for omnibus measures and verbal GBV perpetration, albeit with relatively few studies supporting each conclusion.

### Prior history of the outcome as a moderator of gender-based violence perpetration outcomes

Two reports<sup>81,276</sup> of two evaluations considered prior history of GBV perpetration as a moderator of intervention outcomes. Both collectively indicated that interventions could be more effective for students with a prior history of GBV perpetration. The evaluation of Shifting Boundaries<sup>276</sup> included six short-term tests of an omnibus measure of GBV perpetration. One test suggested a greater impact in reducing GBV perpetration for students with prior history of GBV perpetration; a further two tests suggested a numerical, but not statistical, benefit greater for students with prior history; and a further



**FIGURE 15** Harvest plot, GBV perpetration, boys vs. girls.

three tests indicated a numerical, but not statistical, benefit greater for students without a prior history of the outcome. The New York City evaluation of Fourth R<sup>81</sup> included short-term and long-term tests of an omnibus measure of GBV perpetration. Both tests indicated a numerically, but not statistically, greater benefit for those with a prior history of GBV perpetration. The resultant harvest plot is depicted in [Figure 12](#).

### **Age as a moderator of gender-based violence perpetration outcomes**

Two outcome evaluations<sup>168,277</sup> tested age as a moderator of GBV perpetration outcomes. One evaluation<sup>168</sup> considered a short-term omnibus measure of GBV perpetration and reported the moderation test as non-significant only. Another evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV perpetration and two were of verbal GBV perpetration. Both tests of the omnibus measure were reported as non-significant only; one test of verbal GBV perpetration was reported as non-significant only, whereas one test of verbal GBV perpetration reported a significant result favouring greater impacts for younger students.

### **Ethnicity as a moderator of gender-based violence perpetration outcomes**

Two outcome evaluations<sup>168,277</sup> tested ethnicity (white vs. other) as a moderator of GBV perpetration outcomes. One evaluation<sup>168</sup> considered a short-term omnibus measure of GBV perpetration and reported the moderation test as non-significant only. Another evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV perpetration and two were of verbal GBV perpetration. All tests were reported as non-significant only.

### **Sexuality as a moderator of gender-based violence perpetration outcomes**

Two outcome evaluations<sup>255,277</sup> tested sexuality as a moderator of GBV perpetration outcomes. Both considered long-term effects only. The evaluation of Green Dot<sup>255</sup> tested sexuality as a moderator of physical GBV perpetration and verbal GBV perpetration for a total of 12 tests. None were statistically significant, but 8 of 12 tests demonstrated numerically greater benefits for sexual majority students. Another evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV perpetration and two were of verbal GBV perpetration. Three tests were reported as non-significant only; one test of an omnibus measure of GBV perpetration demonstrated statistically greater benefits for heterosexual students.

### **Poverty status as a moderator of gender-based violence perpetration outcomes**

One evaluation<sup>277</sup> presented four long-term tests, of which two were of an omnibus measure of GBV perpetration and two were of verbal GBV perpetration. All tests were reported as non-significant only.

## **Metaregression analyses**

Across all equity-relevant domains considered in our narrative synthesis, only sex was reported with enough detail to estimate metaregressions. We considered whether the percentage of girls in the sample moderated intervention effectiveness. Importantly, because this is a trial-level analysis, interpretation should be exploratory only. This is because the ecological fallacy can affect interpretation of trial-level metaregressions when the metaregressor is a compositional characteristic of the sample. A more probative test would have been a meta-analysis of within-trial comparisons of intervention effects by sex.

For each analysis, the metaregressor was the percentage of girls in each trial centred at the mean of the trials included in the metaregression and divided by 10. These linear transformations mean that the intercept for each regression is the effectiveness of interventions at the average of percentage of girls in included trials, and the coefficient is expressed as the change in effectiveness associated with a change in the percentage of girls of 10% points.

Findings from metaregressions are reported in [Table 19](#).

### **Dating and relationship violence victimisation**

For DRV victimisation up to 1 year post baseline, intervention effectiveness does not appear to be significantly moderated by the percentage of girls in the trial sample [OR = 1.02, 95% CI (0.67 to 1.56)]. However, there was some signal of moderation of intervention effectiveness on DRV victimisation at or more than 1 year post baseline. Each additional increase by 10% points of girls was linked at marginal significance ( $p = 0.073$ ) with increased effectiveness in reducing DRV victimisation [OR = 0.78, 95% CI (0.59 to 1.04)].

### **Dating and relationship violence perpetration**

For DRV perpetration up to 1 year post baseline, intervention effectiveness does not appear to be moderated by the percentage of girls in the trial sample [OR = 1.01, 95% CI (0.89 to 1.15)]. Similarly, intervention effectiveness at or more than 1 year post baseline does not appear to be moderated by the percentage of girls in the trial sample [OR = 1.05, 95% CI (0.70 to 1.57)].

### **Gender-based violence victimisation**

For GBV victimisation up to 1 year post baseline, the percentage of girls in trial samples was linked with increased effectiveness; with each additional 10% points of girls in the trial sample, the odds of GBV victimisation decreased by 9% [OR = 0.91, 95% CI (0.85 to 0.97)]. Of note is that percentage of girls appeared to be an important source of heterogeneity, as the intercept was statistically significant and indicative of a significant impact of interventions in reducing GBV victimisation [OR = 0.86, 95% CI (0.79 to 0.93)], in distinction to the overall pooled effect for this outcome [OR = 0.88, 95% CI (0.76 to 1.02)]. Interestingly, the relationship between percentage of girls and effectiveness was numerically reversed for outcomes at or more than 1 year post baseline, but the coefficient in this analysis was not statistically significant [OR = 1.21, 95% CI (0.71 to 2.06)].

**TABLE 19** Metaregression, percentage of girl children

Outcome	Time	Term	OR (95% CI)	Residual I <sup>2</sup> (%)
DRV victimisation	Up to 1 year post baseline	Coefficient	1.02 (0.67 to 1.56)	
		Intercept	0.9 (0.79 to 1.04)	83
	At or more than 1 year post baseline	Coefficient	0.78 (0.59 to 1.04)	
		Intercept	0.82 (0.69 to 0.98)	79
DRV perpetration	Up to 1 year post baseline	Coefficient	1.01 (0.89 to 1.15)	
		Intercept	0.9 (0.78 to 1.04)	87
	At or more than 1 year post baseline	Coefficient	1.05 (0.7 to 1.57)	
		Intercept	0.77 (0.62 to 0.95)	81
GBV victimisation	Up to 1 year post baseline	Coefficient	0.91 (0.85 to 0.97)	
		Intercept	0.86 (0.79 to 0.93)	68
	At or more than 1 year post baseline	Coefficient	1.21 (0.71 to 2.06)	
		Intercept	0.93 (0.8 to 1.09)	66
GBV perpetration	Up to 1 year post baseline	Coefficient	0.99 (0.69 to 1.41)	
		Intercept	0.96 (0.85 to 1.08)	74
	At or more than 1 year post baseline	Coefficient	1.01 (0.67 to 1.51)	
		Intercept	0.9 (0.71 to 1.14)	74



**Gender-based violence perpetration**

For GBV perpetration up to 1 year post baseline, the percentage of girls in trial samples was not linked with effectiveness [OR = 0.99, 95% CI (0.69 to 1.41)]. A similar finding was in evidence for outcomes at or more than 1 year post baseline [OR = 1.01, 95% CI (0.67 to 1.51)].



## Chapter 8 Analysis of heterogeneity in intervention effectiveness (RQ4, RQ5)

This chapter presents results from the metaregression of intervention components and QCA (RQ4). It also presents results from the intervention typology and component NMAs. As specified in the protocol, the outcomes analysed in this chapter relate to DRV and GBV victimisation and perpetration.

### Metaregression of intervention components

To undertake this analysis, we used the 13 higher-order component groups defined in [Chapter 6](#). We did not analyse 'miscellaneous' as a category, because of the sparsity of interventions with a miscellaneous component, and where variables were multicategorical and overlapping (i.e. self-guided, peer-guided, adult-guided; single-gender, mixed-gender), we considered these one at a time.

Metaregression results for DRV victimisation are presented in [Table 20](#), and for DRV perpetration in [Table 21](#). Metaregression results for GBV victimisation are presented in [Table 22](#), and for GBV perpetration in [Table 23](#). In all tables, 'coefficient' refers to the change in effectiveness associated with the component's presence, whereas 'intercept' refers to the effectiveness of interventions without the component. Because the goal of this analysis was to identify commonalities in the relationship between component presence and effectiveness across multiple outcomes (i.e. pattern-matching), we present results by specific component type. When there is a numerical (but not statistically significant) difference between groups defined by component inclusion, we refer to a component as 'numerically linked' to a difference in effectiveness, reserving the term 'association' for a statistically significant finding.

### Discussion components

The presence of a discussion-based component was in no analysis significantly associated with improved or worsened intervention effectiveness, though analyses were not estimable for long-term GBV victimisation or perpetration. Based on the direction of coefficients, the presence of a discussion-based component appeared linked with numerically worse effectiveness in all DRV analyses, though this was plausibly related to the relatively small number of interventions without these components. In contrast, the presence of a discussion-based component appeared linked with numerically better effectiveness in short-term analyses for GBV victimisation and perpetration.

### Visual components

The presence of a visual component was in no analysis significantly associated with improved or worsened intervention effectiveness, though analyses were not estimable for long-term GBV victimisation. The presence of a visual component appeared linked with numerically better effectiveness in DRV victimisation analyses, but numerically worse effectiveness in DRV perpetration analyses and short-term GBV victimisation analyses. In analyses for GBV perpetration, the inclusion of a visual component was linked to no difference in effect (OR = 1.00 for both analyses).

### Competition components

The presence of a competition component was in no analysis significantly associated with improved or worsened intervention effectiveness. In all analyses except for short-term DRV victimisation, the inclusion of a competition component was linked with numerically worse effectiveness.

### Guided practice components

The presence of a guided practice component was in no analysis significantly associated with improved or worsened intervention effectiveness, though analyses were not estimable for long-term DRV victimisation, long-term GBV victimisation or long-term GBV perpetration. The presence of

TABLE 20 Component meta-regression findings for DRV victimisation

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)
Overall	Intercept	17	118	0.90 (0.80 to 1.02)	81	13	79	0.82 (0.68 to 0.989)	80
<b>Content components</b>									
Discussion	Coefficient	13	94	1.28 (0.77 to 2.13)		10	69	1.08 (0.32 to 3.67)	
	Intercept	5	24	0.75 (0.40 to 1.40)	94	3	10	0.77 (0.13 to 4.68)	96
Visual	Coefficient	9	62	0.95 (0.71 to 1.28)		10	66	0.93 (0.63 to 1.36)	
	Intercept	8	56	0.94 (0.71 to 1.24)	91	3	13	0.87 (0.71 to 1.06)	96
Competition	Coefficient	4	19	0.93 (0.60 to 1.44)		2	13	1.07 (0.46 to 2.50)	
	Intercept	14	99	0.92 (0.79 to 1.07)	86	11	66	0.81 (0.63 to 1.03)	83
Guided practice	Coefficient	13	75	0.98 (0.62 to 1.55)		12	73	Not estimable	
	Intercept	7	43	0.92 (0.59 to 1.41)	94	1	6		
Reflection	Coefficient	14	93	0.78 (0.32 to 1.93)		11	72	Not estimable	
	Intercept	4	25	1.14 (0.43 to 3.04)	97	2	7		
Narrative-based	Coefficient	7	51	1.08 (0.86 to 1.35)		6	40	1.18 (0.81 to 1.73)	
	Intercept	10	67	0.87 (0.69 to 1.09)	87	7	39	0.77 (0.57 to 1.03)	87
School personnel	Coefficient	7	44	1.03 (0.78 to 1.35)		9	57	1.28 (0.82 to 2.01)	
	Intercept	11	74	0.89 (0.73 to 1.08)	88	4	22	0.70 (0.40 to 1.22)	94
<i>As facilitators</i>	Coefficient	4	21	0.87 (0.50 to 1.52)		5	25	1.06 (0.71 to 1.57)	
	Intercept	13	97	0.92 (0.80 to 1.06)	84	8	54	0.80 (0.60 to 1.07)	88
<i>As participants</i>	Coefficient	4	29	1.05 (0.72 to 1.53)		7	52	1.29 (0.92 to 1.81)	
	Intercept	14	89	0.89 (0.75 to 1.04)	85	6	27	0.72 (0.52 to 1.004)	90

**TABLE 20** Component meta-regression findings for DRV victimisation (*continued*)

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual I <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual I <sup>2</sup> (%)
Parent/family	Coefficient	3	14	0.82 (0.26 to 2.57)		7	38	0.71 (0.48 to 1.06)	
	Intercept	14	104	0.93 (0.84 to 1.04)	82	7	41	0.95 (0.79 to 1.14)	85
Structural environmental	Coefficient	4	27	0.96 (0.72 to 1.28)		4	25	1.23 (0.75 to 2.01)	
	Intercept	14	91	0.91 (0.76 to 1.09)	85	9	54	0.77 (0.61 to 0.98)	86
Structural social	Coefficient	7	55	0.85 (0.67 to 1.09)		6	35	0.82 (0.55 to 1.22)	
	Intercept	11	63	0.98 (0.83 to 1.16)	90	7	44	0.90 (0.68 to 1.18)	88
<b>Delivery components</b>									
Self-guided	Coefficient	3	10	0.82 (0.21 to 3.19)		4	16	0.62 (0.48 to 0.80)	
	Intercept	14	108	0.93 (0.85 to 1.02)	82	9	63	0.94 (0.83 to 1.07)	72
Peer-guided	Coefficient	2	18	0.93 (0.26 to 3.35)		0	0	Not applicable	
	Intercept	15	100	0.91 (0.80 to 1.04)	84	13	79		
Adult-guided	Coefficient	14	108	1.22 (0.31 to 4.71)		11	71	Not estimable	
	Intercept	3	10	0.76 (0.12 to 5.02)	97	2	8		
Single-gender	Coefficient	5	20	1.29 (0.87 to 1.92)		2	5	1.38 (0.19 to 9.87)	
	Intercept	13	98	0.88 (0.76 to 1.01)	85	11	74	0.78 (0.64 to 0.96)	83
Mixed-gender	Coefficient	15	110	Not estimable		10	67	0.85 (0.37 to 1.92)	
	Intercept	2	8			3	12	0.94 (0.32 to 2.77)	96

*k*, number of studies; *n*, number of effect sizes.

**TABLE 21** Component meta-regression findings for DRV perpetration

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)
Overall	Intercept	18	118	0.91 (0.80 to 1.04)	83	16	79	0.78 (0.64 to 0.94)	79
<b>Content components</b>									
Discussion	Coefficient	16	102	1.79 (0.53 to 6.00)		13	69	1.19 (0.43 to 3.34)	
	Intercept	3	16	0.55 (0.11 to 2.62)	96	3	10	0.68 (0.16 to 2.85)	96
Visual	Coefficient	8	54	1.03 (0.74 to 1.42)		11	63	1.10 (0.68 to 1.76)	
	Intercept	10	64	0.89 (0.65 to 1.22)	93	5	16	0.72 (0.44 to 1.18)	94
Competition	Coefficient	4	19	1.03 (0.64 to 1.64)		2	9	1.34 (0.69 to 2.63)	
	Intercept	15	99	0.90 (0.76 to 1.07)	87	14	70	0.73 (0.59 to 0.91)	80
Guided practice	Coefficient	14	75	1.11 (0.63 to 1.95)		13	70	1.28 (0.41 to 3.97)	
	Intercept	7	43	0.83 (0.46 to 1.49)	94	3	9	0.63 (0.15 to 2.67)	97
Reflection	Coefficient	13	80	0.95 (0.59 to 1.54)		12	69	1.32 (0.67 to 2.59)	
	Intercept	6	38	0.95 (0.61 to 1.49)	95	4	10	0.62 (0.28 to 1.39)	96
Narrative-based	Coefficient	7	51	1.08 (0.82 to 1.42)		6	40	1.30 (0.93 to 1.82)	
	Intercept	11	67	0.87 (0.67 to 1.14)	90	10	39	0.71 (0.55 to 0.92)	84
School personnel	Coefficient	8	51	1.08 (0.83 to 1.40)		12	61	1.09 (0.68 to 1.77)	
	Intercept	11	67	0.88 (0.70 to 1.10)	88	4	18	0.73 (0.42 to 1.25)	95
<i>As facilitators</i>	Coefficient	4	27	1.01 (0.43 to 2.37)		7	28	1.13 (0.77 to 1.65)	
	Intercept	14	91	0.91 (0.78 to 1.06)	85	9	51	0.74 (0.55 to 0.98)	88
<i>As participants</i>	Coefficient	5	36	1.10 (0.79 to 1.53)		10	56	0.97 (0.65 to 1.46)	
	Intercept	14	82	0.88 (0.73 to 1.07)	85	6	23	0.78 (0.55 to 1.13)	93

**TABLE 21** Component meta-regression findings for DRV perpetration (*continued*)

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual I <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual I <sup>2</sup> (%)
Parent/family	Coefficient	3	14	0.77 (0.18 to 3.31)		8	39	0.75 (0.51 to 1.09)	
	Intercept	15	104	0.96 (0.87 to 1.05)	82	9	40	0.88 (0.72 to 1.09)	85
Structural environmental	Coefficient	3	21	1.08 (0.76 to 1.53)		7	36	1.34 (0.89 to 2.00)	
	Intercept	16	97	0.88 (0.73 to 1.08)	85	9	43	0.72 (0.57 to 0.91)	84
Structural social	Coefficient	6	49	0.95 (0.67 to 1.33)		4	25	0.91 (0.60 to 1.37)	
	Intercept	13	69	0.93 (0.80 to 1.08)	90	12	54	0.81 (0.65 to 1.02)	88
<b>Delivery components</b>									
Self-guided	Coefficient	2	8	0.55 (0.13 to 2.31)		4	16	0.69 (0.45 to 1.07)	
	Intercept	16	110	0.97 (0.91 to 1.05)	79	12	63	0.87 (0.73 to 1.03)	77
Peer-guided	Coefficient	1	12	Not estimable		0	0	Not applicable	
	Intercept	17	106			16	79		
Adult-guided	Coefficient	16	110	Not estimable		14	71	Not estimable	
	Intercept	2	8			2	8		
Single-gender	Coefficient	6	26	0.87 (0.40 to 1.86)		5	9	0.80 (0.47 to 1.37)	
	Intercept	13	92	0.92 (0.80 to 1.07)	86	11	70	0.82 (0.66 to 1.02)	85
Mixed-gender	Coefficient	16	105	Not estimable		10	63	1.14 (0.68 to 1.91)	
	Intercept	2	13			6	16	0.71 (0.41 to 1.23)	93

*k*, number of studies; *n*, number of effect sizes.

TABLE 22 Component meta-regression findings for GBV victimisation

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)
Overall	Intercept	13	72	0.88 (0.76 to 1.02)	75	11	58	0.93 (0.80 to 1.08)	66
<b>Content components</b>									
Discussion	Coefficient	12	61	Not estimable		9	52	Not estimable	
	Intercept	2	11			2	6		
Visual	Coefficient	8	39	1.29 (0.90 to 1.84)		9	53	Not estimable	
	Intercept	5	33	0.74 (0.49 to 1.13)	89	2	5		
Competition	Coefficient	4	8	1.11 (0.51 to 2.43)		6	38	1.30 (0.92 to 1.83)	
	Intercept	10	64	0.86 (0.73 to 1.01)	81	5	20	0.83 (0.80 to 0.86)	81
Guided practice	Coefficient	13	50	0.84 (0.35 to 2.05)		11	58	Not estimable	
	Intercept	3	22	1.02 (0.40 to 2.58)		95	0	0	
Reflection	Coefficient	9	56	1.41 (0.99 to 2.02)		10	55	Not estimable	
	Intercept	5	16	0.68 (0.45 to 1.03)		90	1	3	
Narrative-based	Coefficient	6	33	1.24 (0.95 to 1.62)		5	20	1.39 (0.85 to 2.27)	
	Intercept	7	39	0.78 (0.61 to 0.996)	84	6	38	0.83 (0.72 to 0.95)	79
School personnel	Coefficient	6	27	1.27 (0.90 to 1.79)		9	30	Not estimable	
	Intercept	8	45	0.78 (0.57 to 1.06)		89	2	28	
<i>As facilitators</i>	Coefficient	3	6	1.21 (0.71 to 2.05)		7	17	1.29 (0.92 to 1.79)	
	Intercept	10	66	0.83 (0.70 to 0.999)		82	4	41	0.81 (0.61 to 1.07)
<i>As participants</i>	Coefficient	4	24	1.34 (0.90 to 1.99)		7	25	1.02 (0.69 to 1.52)	
	Intercept	10	48	0.81 (0.67 to 0.98)		83	4	33	0.92 (0.58 to 1.47)



**TABLE 22** Component meta-regression findings for GBV victimisation (*continued*)

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)
Parent/family	Coefficient	1	3	Not estimable		5	11	1.27 (0.80 to 2.02)	
	Intercept	12	69			7	47	0.87 (0.70 to 1.06)	79
Structural environmental	Coefficient	3	20	1.29 (0.60 to 2.74)		5	11	1.46 (0.95 to 2.23)	
	Intercept	11	52	0.83 (0.69 to 0.99)	81	6	47	0.83 (0.73 to 0.95)	77
Structural social	Coefficient	7	50	0.97 (0.70 to 1.35)		6	13	1.34 (0.94 to 1.90)	
	Intercept	7	22	0.89 (0.77 to 1.03)	89	5	45	0.82 (0.69 to 0.98)	83
<b>Delivery components</b>									
Self-guided	Coefficient	0	0	Not applicable		0	0	Not applicable	
	Intercept	13	72			11	58		
Peer-guided	Coefficient	0	0	Not applicable		1	3	Not estimable	
	Intercept	13	72			10	55		
Adult-guided	Coefficient	13	72	Not applicable		11	58	Not applicable	
	Intercept	0	0			0	0		
Single-gender	Coefficient	5	10	0.69 (0.49 to 0.98)		1	3	Not estimable	
	Intercept	9	62	0.97 (0.87 to 1.08)	76	10	55		
Mixed-gender	Coefficient	9	65	1.28 (0.78 to 2.11)		10	56	Not estimable	
	Intercept	4	7	0.74 (0.37 to 1.45)	94	1	2		

*k*, number of studies; *n*, number of effect sizes.

**TABLE 23** Component meta-regression findings for GBV perpetration

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual <i>I</i> <sup>2</sup> (%)
Overall	Intercept	11	67	0.95 (0.85 to 1.07)	66	9	58	0.90 (0.73 to 1.12)	67
<b>Content components</b>									
Discussion	Coefficient	10	58	0.69 (0.23 to 2.08)		7	54	Not estimable	
	Intercept	2	9	1.32 (0.23 to 7.55)	94	2	4		
Visual	Coefficient	7	33	1.00 (0.80 to 1.25)		6	48	1.00 (0.63 to 1.59)	
	Intercept	4	34	0.95 (0.80 to 1.13)	89	3	10	0.90 (0.59 to 1.37)	90
Competition	Coefficient	3	4	1.15 (0.24 to 5.56)		4	32	1.05 (0.56 to 1.99)	
	Intercept	9	63	0.93 (0.85 to 1.02)	72	5	26	0.88 (0.75 to 1.03)	81
Guided practice	Coefficient	10	41	0.85 (0.58 to 1.27)		8	53	Not estimable	
	Intercept	4	26	1.08 (0.86 to 1.35)	86	1	5		
Reflection	Coefficient	8	50	1.03 (0.76 to 1.38)		7	50	Not estimable	
	Intercept	4	17	0.93 (0.74 to 1.18)	92	2	8		
Narrative-based	Coefficient	4	28	1.18 (0.93 to 1.49)		3	16	1.33 (0.56 to 3.15)	
	Intercept	7	39	0.87 (0.72 to 1.05)	78	6	42	0.83 (0.66 to 1.03)	73
School personnel	Coefficient	7	30	1.08 (0.81 to 1.44)		7	30	Not estimable	
	Intercept	5	37	0.91 (0.70 to 1.20)	89	2	28		
<i>As facilitators</i>	Coefficient	2	3	1.10 (0.09 to 13.03)		4	10	1.29 (0.79 to 2.12)	
	Intercept	9	64	0.94 (0.86 to 1.03)	72	5	48	0.82 (0.64 to 1.05)	79
<i>As participants</i>	Coefficient	6	28	1.18 (0.90 to 1.54)		6	26	1.31 (0.78 to 2.19)	
	Intercept	6	39	0.88 (0.69 to 1.14)	86	3	32	0.76 (0.38 to 1.55)	90

**TABLE 23** Component meta-regression findings for GBV perpetration (*continued*)

		Short-term				Long-term			
		<i>k</i>	<i>n</i>	OR (95% CI)	Residual I <sup>2</sup> (%)	<i>k</i>	<i>n</i>	OR (95% CI)	Residual I <sup>2</sup> (%)
Parent/family	Coefficient	1	1	Not estimable		3	5	1.65 (0.89 to 3.07)	
	Intercept	10	66			7	53	0.82 (0.69 to 0.98)	62
Structural environmental	Coefficient	2	17	1.31 (0.36 to 4.81)		2	4	1.61 (0.37 to 7.00)	
	Intercept	10	50	0.92 (0.82 to 1.03)	71	7	54	0.85 (0.73 to 0.998)	59
Structural social	Coefficient	4	43	1.22 (0.94 to 1.59)		3	6	1.45 (0.70 to 3.02)	
	Intercept	8	24	0.87 (0.73 to 1.04)	76	6	52	0.82 (0.68 to 0.99)	69
<b>Delivery components</b>									
Self-guided	Coefficient	0	0	Not applicable		0	0	Not applicable	
	Intercept	11	67			9	58		
Peer-guided	Coefficient	0	0	Not applicable		0	0	Not applicable	
	Intercept	11	67			9	58		
Adult-guided	Coefficient	11	67	Not applicable		9	58	Not applicable	
	Intercept	0	0			0	0		
Single-gender	Coefficient	3	9	1.00 (0.60 to 1.66)		2	8	0.90 (0.37 to 2.17)	
	Intercept	9	58	0.95 (0.83 to 1.09)	73	7	50	0.92 (0.69 to 1.22)	75
Mixed-gender	Coefficient	9	61	Not estimable		7	51	Not estimable	
	Intercept	2	6			2	7		

*k*, number of studies; *n*, number of effect sizes.

a guided practice component was numerically linked with improved effectiveness in short-term DRV victimisation, GBV victimisation and GBV perpetration, but numerically linked with decreased effectiveness in short-term DRV perpetration.

### **Reflection components**

The presence of a reflection component was in no analysis significantly associated with improved or worsened intervention effectiveness, though analyses were not estimable for long-term DRV victimisation, long-term GBV victimisation or long-term GBV perpetration. Reflection components were linked with a numerically greater effectiveness for short-term DRV victimisation and short-term DRV perpetration, but with numerically worse effectiveness for long-term DRV perpetration, short-term GBV victimisation [though this was only marginally non-significant, with OR = 1.41, 95% CI (0.99 to 2.02)] and short-term GBV perpetration.

### **Narrative-based components**

The presence of a narrative component was in no analysis significantly associated with improved or worsened intervention effectiveness, though in every analysis it was numerically linked with worse effectiveness.

### **School personnel components**

Analysis of school personnel components was first undertaken overall, and then considering just school personnel as facilitators and just school personnel as participants. An overall analysis of school personnel components was not estimable for long-term GBV victimisation and perpetration, but in every other analysis presence of any school personnel component was numerically, but not statistically, linked to worse effectiveness.

Focusing on just components where school personnel were trained as facilitators, none of the eight analyses suggested a statistical association between this component and effectiveness; however, there was a numerical link with greater effectiveness for short-term DRV victimisation, no numerical difference for short-term DRV perpetration (OR = 1.01), and a highly imprecise CI for short-term GBV perpetration. In all other analyses, presence of a component with teachers as facilitators was linked with numerically worse effectiveness.

Focusing on just components where school personnel were participants, none of the eight analyses suggested a statistical association with effectiveness. All analyses except long-term DRV perpetration suggested a numerical worsening in effectiveness with the inclusion of a component where school personnel were participants.

### **Parent and family components**

The presence of a parent and family component was in no analysis significantly associated with improved or worsened intervention effectiveness, though analyses were not possible for short-term GBV victimisation or GBV perpetration. While the presence of this component was numerically linked with improved effectiveness for all DRV victimisation and perpetration analyses, the converse was demonstrated for long-term GBV victimisation and long-term GBV perpetration.

### **Structural environmental components**

The presence of a structural environmental component was in no analysis significantly associated with intervention effectiveness. In all analyses except for short-term DRV victimisation, the presence of this component was linked with numerically worse intervention effectiveness.

### **Structural social components**

The presence of a social structural component was in no analysis significantly associated with intervention effectiveness. In short-term and long-term DRV victimisation and perpetration and short-term GBV victimisation, presence of this component was linked to numerically better intervention

effectiveness; however, for long-term GBV victimisation and short-term and long-term GBV perpetration, this component was linked to numerically worse intervention effectiveness.

### **Guidance: self, peer, adult**

Due to limited variation within analyses, type of guidance could not be analysed at all for GBV victimisation or GBV perpetration. The presence of a peer-guided component could only be tested in short-term DRV victimisation. This suggested a numerical, but not statistical, benefit of including this component. In contrast, an adult-guided component was linked with numerically worse effectiveness for this outcome. Self-guided components were numerically linked with improved effectiveness in short-term DRV victimisation and DRV perpetration, and long-term DRV perpetration, and statistically associated with improved intervention effectiveness in long-term DRV victimisation.

### **Gender composition**

Single-gender components and mixed-gender components were considered separately. An analysis for single-gender components was not estimable for long-term GBV victimisation; however, in remaining analysis, a signal of effectiveness was unclear. While numerically linked with worse effectiveness for both short-term and long-term DRV victimisation, single-gender components were numerically linked with improved effectiveness for both short-term and long-term DRV perpetration. Single-gender components were statistically associated with improved effectiveness for short-term GBV victimisation, but with no difference in effect (OR = 1.00) for short-term GBV perpetration; and numerically linked with improved intervention effectiveness in long-term GBV perpetration.

In contrast, models for the inclusion of mixed-gender components were not estimable for short-term DRV victimisation, short-term DRV perpetration, long-term GBV victimisation or long-term and short-term GBV perpetration. While numerically linked with improved effectiveness in long-term DRV victimisation, mixed-gender components were numerically linked with worse effectiveness in long-term DRV perpetration and short-term GBV victimisation.

## **Qualitative comparative analysis**

We undertook QCA, or QCA, to understand pathways to intervention effectiveness. In addition to analysing pathways to effectiveness for DRV and GBV victimisation and perpetration outcomes distinctly, our protocol proposed that we considered pathways to 'joint' effectiveness on DRV and GBV victimisation or on DRV and GBV perpetration. However, the overlap between trials reporting both DRV and GBV victimisation outcomes or both DRV and GBV perpetration outcomes was too small to render these analyses meaningful.

Therefore, we started with a pool of eight possible models: DRV victimisation, DRV perpetration, GBV victimisation and GBV perpetration, each in the short term and in the long term. In each QCA, an observation was defined by a comparison between an intervention and a control; thus, where trials entered into pairwise intervention-control meta-analyses were multiarm in design, each intervention-control comparison constituted a separate observation. As a result, the number of observations in each QCA model was higher than the number of trials in the corresponding pairwise meta-analysis.

### **Initial calibration and conditions**

Our analysis began by calibrating the outcome using pooled intervention-comparison effect sizes generated for the NMAs. As a starting point, we used a threshold of OR = 0.90 to distinguish 'most effective' interventions (outcome value of 1) from 'not most effective' (outcome value of 0) interventions. This distinction is important because trials with large sample sizes may generate statistically significant effects on an outcome, but with an effect size that may be too small to be a useful, efficient investment of resources by schools. Because we defined our outcome variable as 'most effective' rather than 'effective', interventions classed as 'not most effective' may have statistically meaningful effects on key

outcomes but may not be considered best bets for implementers. In addition, an OR of 0.90 appeared to be a consistent breakpoint in the distribution of intervention-control comparisons included in each analysis and is close to the pooled effect size for most overall pairwise meta-analysis models where a significant result was not in evidence. Once we assigned individual intervention-control comparisons to 'most effective' or 'not most effective', we sense checked the resulting classifications on the basis of the overall pattern of (disaggregated) results from each trial.

We identified a core set of four conditions to enter into each model, based on the previous syntheses. These conditions were:

- country context of interventions, whether low-income and middle-income or high-income, which was found to be a plausible moderator for most outcomes;
- percentage of girl children in the trial sample, which was found to be a plausible moderator for victimisation outcomes, but which was included for perpetrator outcomes for consistency (and because a configurational analysis might nuance the null findings from metaregressions on perpetration outcomes);
- single-component interventions, given the unexpected effectiveness of these interventions in several pairwise meta-analyses;
- structural social domain of implementation, given the theorised salience of these components in activating higher-level change mechanisms, and numerically indicative results from the component metaregression.

Because percentage of girl children was a continuous variable, it required calibration into a binary condition. We calibrated samples with more than 60% girl children as 'girls plus' (condition value of 1) on the basis that a sample with more than 60% girl children would be unlikely to arise by chance alone and would also include interventions delivered to girls exclusively. While we included single-component interventions as a condition, we did not include curriculum or multicomponent interventions as conditions as these types were not clearly linked with effectiveness in meta-analyses.

Upon calibration of outcome and conditions, we did not proceed with models for long-term GBV victimisation and perpetration. The number of observations available in each model ( $n = 12$  and  $n = 10$ , respectively) compared to the number of conditions, and the number of interventions that could be reasonably classed as most effective, was too small for any QCA to be meaningful.

### ***Dating and relationship violence victimisation***

#### **Short-term analyses**

An initial model with the four conditions described above generated three contradictory configurations (i.e. combinations of conditions linked with both most effective and not most effective interventions). Upon scrutinising the contradictory configurations, we added two conditions.

First, we created a condition reflecting interventions that were most effective at reducing short-term DRV perpetration and that included a focus on more than just perpetration. Thus, interventions that did not report short-term DRV perpetration were labelled as 0. In addition, one intervention (Fernandez-Gonzalez and Sánchez-Álvarez 2020<sup>169</sup>) was considered most effective in short-term DRV perpetration analyses but focused on perpetration rather than victimisation (and thus could not be expected to provide a strong theoretical test of impact on victimisation); this intervention was labelled as 0. We viewed this condition as being an important theoretical test of whether reductions in perpetration are causal for reductions in victimisation, given the focus on school-based intervention and the context within which DRV often occurs in adolescents.

Second, we created a condition that reflected whether interventions included single-gender components. We viewed this as being an important theoretical test given the contrast between analyses

showing gendered moderation of victimisation outcomes and the numerically worse effectiveness of interventions including single-gender components. When these two conditions were included in the model, we were able to remove conditions for country context and for social structural domain of implementation as these did not improve the model.

This generated a consistent analysis with full coverage of most effective cases, reflected in [Table 24](#), which was taken to Boolean minimisation. The reduced configurations demonstrate several salient points. The first is the central importance of ‘co-reducing’ DRV perpetration; that is, a key way that interventions can reduce victimisation is by targeting perpetration in the context of a focus on more than just perpetration. As evidenced by reduced results including configuration configurations abC (not high percentage of girls in the intervention sample AND not single-component intervention AND effective in reducing short-term DRV perpetration) and aCd (not high percentage of girls in the intervention sample AND effective in reducing short-term DRV perpetration AND no single-gender components); this condition was central to effectiveness in a way that was invariant to high-level intervention features. But what are also interesting are the reduced configurations demonstrating pathways to most effectiveness where effectiveness in reducing DRV perpetration was not in evidence; specifically reduced configurations aBcD and Abcd. Configuration aBcD suggests that where perpetration reductions are not in evidence (c), DRV victimisation reductions can still be achieved in the short-term via single-component interventions (B) with single-gender components (D) where samples do not have high levels of girl children (a); equally, as noted in configuration Abcd, where samples include a high level of girl children (A), interventions that go beyond single-component approaches (b) without effectiveness in perpetration (c) or single-gender components (d) can be most effective. This reflects the potential importance of gender-relevant programming.

### Long-term analyses

We attempted the initial model with four conditions for long-term DRV victimisation but generated a similar set of contradictory conditions.

**TABLE 24** Qualitative comparative analysis for short-term DRV victimisation

Set	Outcome	Interventions
abcD	Not most effective	Fourth R (Cissner 2014) <sup>81</sup>
abCd	Most effective	Expect Respect (Roberts 2009), <sup>95</sup> JOVEN (Gonzalez-Guarda 2015), <sup>193</sup> Shifting Boundaries building (Taylor 2011) <sup>100</sup>
abcd	Not most effective	Dat-e (Sanchez-Jimenez 2018), <sup>49</sup> Shifting Boundaries classroom, Shifting Boundaries building and classroom (Taylor 2011), <sup>100</sup> Safe Dates (Foshee 1998) <sup>36</sup>
aBcd	Not most effective	Interaction-based curriculum, law and justice-based curriculum (Taylor 2008), <sup>248</sup> Ending Violence (Jaycox 2006), <sup>185</sup> Incremental theory of personality (Fernandez-Gonzalez 2020) <sup>169</sup>
aBcD	Most effective	Scientist-Practitioner Programme (Muck 2018) <sup>94</sup>
aBCd	Most effective	Practitioner Programme (Muck 2018), <sup>94</sup> Katie Brown Educational Program (Joppa 2016), <sup>186</sup> Teen Choices (Levesque 2016) <sup>182</sup>
Abcd	Most effective	SHARP (Miller 2015), <sup>187</sup> PREPARE (Mathews 2016) <sup>90</sup>
ABcD	Not most effective	Twilight, True Love and You (Lynch 2014) <sup>88</sup>
abCD	Most effective	Alcohol and dating violence, single-sex intervention (Filho 2017) <sup>84</sup>
<b>Reduced pathway</b>	<b>aBcD OR Abcd OR abC OR aCd</b>	

A, high percentage of girls in the trial sample; B, single-component intervention; C, perpetration and victimisation interventions that were effective in reducing short-term DRV perpetration; D, single-gender components; an upper-case letter indicates the presence of a condition, while a lower-case letter indicates its absence.

We then implemented a full model with six conditions as undertaken for short-term DRV victimisation. We noted that the single-gender components condition did not contribute to model fit and thus dropped it from the analysis; similarly, we dropped the condition for country context and for social structural domain of implementation. The remaining model still contained a contradictory configuration, so we scrutinised interventions for differences and noticed that a condition that appeared to resolve this contradiction related to the presence of multiple components with active parental involvement, for example, JOVEN,<sup>193</sup> which included multiple rounds and activities of parental involvement. The final model included four conditions: percentage of girl children in the trial sample, single-component intervention type, a condition indicating whether interventions were most effective at reducing long-term DRV perpetration with a focus more than on just perpetration, and a condition for active parental involvement.

As noted in [Table 25](#), this generated a consistent analysis with full coverage of most effective cases. The three reduced configurations again reflect the central importance of concomitant reductions in perpetration to victimisation (reduced configuration aCd), but also reflect the importance of high numbers of girls in the trial sample in the absence of reductions in perpetration, and the potential value of single-component interventions in HIC contexts. The absence of a high level of parental involvement in intervention components was part of all reduced configurations.

### Dating and relationship violence perpetration

#### Short-term analyses

An initial analysis for short-term DRV perpetration with the four core conditions generated contradictory configurations. We were unable to test or develop, based on our syntheses, a satisfactory explanation for the contradictory configurations in evidence, despite returning to this analysis with insights as to candidate conditions from other analyses. This included considering conditions relating to a range of intervention components and levels of these components, as well as other intervention characteristics relating to type, focus and domain of implementation.

**TABLE 25** Qualitative comparative analysis for long-term DRV victimisation

Set	Outcome	Interventions
abCd	Most effective	It's Your Game (Peskin 2014), <sup>189</sup> Me and You (Peskin 2019), <sup>188</sup> Fourth R (Cissner 2014), <sup>81</sup> Safe Dates (Foshee 1998) <sup>36</sup>
abcd	Not most effective	Green Dot (Coker 2017), <sup>82</sup> Project Respect (Meiksin 2020), <sup>35</sup> Skhokho schools (Jewkes 2019), <sup>155</sup> Program H and M (Bando 2019), <sup>184</sup> Second Step (Espelage 2013) <sup>192</sup>
abcD	Not most effective	Skhokho families and schools (Jewkes 2019) <sup>155</sup>
aBcd	Most effective	Incremental theory of personality (Fernandez-Gonzalez 2020) <sup>169</sup>
aBCd	Most effective	<i>Teen Choices</i> (Levesque 2016) <sup>211</sup>
Abcd	Most effective	PREPARE (Mathews 2016) <sup>90</sup>
abCD	Not most effective	JOVEN (Gonzalez-Guarda 2015) <sup>193</sup>
<b>Reduced pathway</b>	<b>Abcd OR aCd OR aBd</b>	

A, high percentage of girls in the trial sample; B, single-component intervention; C, perpetration and victimisation interventions that were effective in reducing long-term DRV perpetration; D, high levels of active parental involvement components; an upper-case letter indicates the presence of a condition, while a lower-case letter indicates its absence.



### Long-term analyses

An initial analysis for long-term DRV perpetration with the four core conditions generated contradictory configurations. In considering these contradictory configurations, we first noted that the condition for a high percentage of girls in the trial sample was irrelevant and thus excluded it from the analysis. We also noted that two additional conditions were useful to distinguish most effective and not most effective interventions:

- the presence of components facilitating interpersonal relationships between students as a domain of implementation (focused on developing relationships between students and interpersonal aspects of student behaviour; for example dialogue and interaction-based components)
- interventions including a high level of guided practice.

We calibrated the first condition using the intervention typology coding and we calibrated the second condition by considering the number, diversity and salience of guided practice components in intervention descriptions. In practice, interventions classed as having two or more specific guided practice components were classed as having high levels of guided practice. In addition, we reconsidered the calibration of Green Dot (Coker *et al.* 2017)<sup>82</sup> as not most effective in this analysis. The pooled OR for student-level outcomes in Green Dot was 0.93, but because of the uniquely longitudinal nature of this analysis, the pattern of outcomes suggesting that results only 'bedded in' in years 3 and 4 of the evaluation and the presence of student-level tests (used to derive OR = 0.93) and school-level tests (which showed consistent and significant time trends between intervention and control schools), we agreed that Green Dot was a boundary case for most effectiveness for DRV perpetration in a way that was not in evidence for DRV victimisation. As a result, we recoded Green Dot as most effective for this outcome. Finally, we noted that country context did not contribute to the model and removed it as a condition.

This generated a consistent analysis with full coverage of most effective cases, reflected in [Table 26](#), which was taken to Boolean minimisation. The three reduced configurations are notable in the first instance for the relative unimportance of social structural components, as the presence of these components did not enter into any pathways to most effectiveness. Second, while single-component interventions alone appear to be important to effectiveness (reduced configuration Abcd), the remaining two pathways illustrate the potential synergistic effects of interactions between student interpersonal

**TABLE 26** Qualitative comparative analysis for long-term DRV perpetration

Set	Outcome	Interventions
abcd	Not most effective	Second Step (Espelage 2013) <sup>192</sup>
abCd	Most effective	CBIM (Miller 2012, <sup>77</sup> Miller 2020 <sup>180</sup> ), JOVEN (Gonzalez-Guarda 2015), <sup>193</sup> It's Your Game (Peskin 2014) <sup>189</sup>
abCD	Most effective	Fourth R (Cissner 2014), <sup>81</sup> Me and You (Peskin 2019) <sup>188</sup>
aBCD	Most effective	Fourth R (Wolfe 2009), <sup>190</sup> Green Dot (Coker 2017), <sup>82</sup> Safe Dates (Foshee 1998) <sup>36</sup>
Abcd	Most effective	Incremental theory of personality (Fernandez-Gonzalez 2020), <sup>169</sup> Teen Choices (Levesque 2016) <sup>182</sup>
aBCd	Not most effective	Skhokho schools, Skhokho families and schools (Jewkes 2019), <sup>155</sup> Program H and M (Bando 2019), <sup>184</sup> Project Respect (Meiksin 2020) <sup>35</sup>
aBcD	Not most effective	PREPARE (Mathews 2016) <sup>90</sup>
<b>Reduced pathway</b>	<b>Abcd OR abC OR aCD</b>	

A, single-component intervention; B, social structural domain of implementation; C, implementation of student interpersonal components; D, high level of guided practice; an upper-case letter indicates the presence of a condition, while a lower-case letter indicates its absence.

components and high levels of guided practice in shaping pathways to effectiveness (reduced configurations abC and aCD).

### **Gender-based violence victimisation: short-term**

An initial analysis for short-term GBV victimisation with the four core conditions generated contradictory configurations. Based on our experience of the DRV victimisation analyses, we generated a similar condition for most effectiveness in short-term GBV perpetration. This did not, however, resolve all contradictory configurations.

In scrutinising the remaining contradictory configurations, we noticed that a condition that distinguished most effective interventions from not most effective interventions was the presence of narrative components. Based on the consistent numerical patterns seen in component meta-regressions, we believed that this component could plausibly contribute to understanding pathways to effectiveness. When these two conditions were included in the model, we were able to remove conditions for country context and for social structural domain of implementation as these did not improve the model.

The resultant set of conditions generated a fully consistent model with complete coverage of all most effective interventions. This analysis (Table 27), when taken to Boolean minimisation, reflected the importance of reducing GBV perpetration, but also the importance of avoiding narrative components. A potentially surprising finding is that most effective interventions that were also effective at reducing GBV perpetration were not single-component interventions (reduced configuration abCd), and single-component interventions that were most effective did not need to reduce GBV perpetration to be most effective at reducing GBV victimisation (Bcd). Another pathway (reduced configuration Abc) indicated that interventions that were most effective at reducing GBV victimisation were not single-component interventions and were implemented in samples with high percentages of girl children.

**TABLE 27** Qualitative comparative analysis for short-term GBV victimisation

Set	Outcome	Interventions
abcd	Not most effective	Shifting Boundaries classroom (Taylor 2011), <sup>100</sup> Shifting Boundaries building (Taylor 2011) <sup>100</sup>
abCD	Not most effective	Benzies and Batchies (de Lijster 2016) <sup>168</sup>
aBcd	Most effective	Practitioner Programme, Scientist-Practitioner Programme (Muck 2018) <sup>94</sup>
aBcD	Not most effective	Interaction-based curriculum, law and justice-based curriculum (Taylor 2008) <sup>248</sup>
ABcd	Most effective	My Voice, My Choice (Rowe 2015) <sup>96</sup>
abCd	Most effective	Let Us Protect Our Future (Jemmott 2018), <sup>176</sup> Fourth R (Cissner 2014), <sup>81</sup> BITB-HSC (Edwards 2019), <sup>197</sup> Shifting Boundaries building and classroom (Taylor 2011) <sup>100</sup>
abcD	Not most effective	GEMS India (ICRW 2017), <sup>85</sup> Second Step (Espelage 2013) <sup>192</sup>
AbcD	Most effective	IMPower/50 : 50 (Baiocchi 2017), <sup>160</sup> PREPARE (Mathews 2016) <sup>90</sup>
Abcd	Most effective	IMPower (Decker 2018) <sup>174</sup>
<b>Reduced pathway</b>	<b>abCd OR Bcd OR Abc</b>	

A, high percentage of girls in the trial sample; B, single-component intervention; C, perpetration and victimisation interventions that were effective in reducing short-term DRV victimisation; D, presence of narrative components; an upper-case letter indicates the presence of a condition, while a lower-case letter indicates its absence.

### Gender-based violence perpetration: short-term

An initial analysis for short-term GBV victimisation with the four core conditions generated contradictory configurations. Based on our learning from models for DRV perpetration, we included conditions for high levels of guided practice and for the presence of student interpersonal components. This generated a completely consistent model; however, conditions for high percentage of girl children in the trial sample, country context and single-component interventions were no longer necessary to explain patterns of most effectiveness.

Thus, the final model included conditions for social structural domain of implementation, student interpersonal components, and a high level of guided practice. As seen in [Table 28](#), pathways to effectiveness required either the presence of social structural components with student interpersonal components (reduced configuration ABc) or the presence of a high level of guided practice (reduced configuration aC).

## Network meta-analysis

We undertook two types of NMAs: type-based NMAs, in which we considered the comparative effectiveness of intervention types identified in six (single-component, curriculum, multicomponent, multilevel), and component-based NMAs, in which we considered and ‘disentangled’ the effectiveness of different intervention component types. In our protocol, we described analyses linked to intervention type as our ‘primary’ analyses while our component-based NMAs were exploratory. Ranking of interventions and the results of sensitivity analyses are reported in [Report Supplementary Material 1](#).

### Feasibility assessment

We undertook type-based NMAs for all outcomes. We ultimately undertook component-based NMAs using the domains of implementation (intrapersonal, interpersonal, staff, structural social and structural environmental) as our components to balance the number of trials and the number of components entered into the analyses. We did not explore interactions between components as the data structure

**TABLE 28** Qualitative comparative analysis for short-term GBV perpetration

Set	Outcome	Interventions
abc	Not most effective	Law and justice-based curriculum (Taylor 2008), <sup>248</sup> Practitioner Programme, Scientist-Practitioner Programme (Muck 2018), <sup>94</sup> Second Step (Espelage 2013) <sup>192</sup>
abC	Most effective	Let Us Protect Our Future (Jemmott 2018) <sup>176</sup>
aBc	Not most effective	Sexual Harassment and Teens (Durand 1997), <sup>83</sup> CBIM (Miller 2020), <sup>180</sup> Interaction-based curriculum (Taylor 2008), <sup>248</sup> Shifting Boundaries classroom (Taylor 2011) <sup>100</sup>
aBC	Most effective	Benzies and Batchies (de Lijster 2016), <sup>168</sup> Fourth R (Cissner 2014), <sup>81</sup> BITB-HSC (Edwards 2019) <sup>197</sup>
Abc	Not most effective	Shifting Boundaries building (Taylor 2011) <sup>100</sup>
ABc	Most effective	Shifting Boundaries building and classroom (Taylor 2011) <sup>100</sup>
ABC	Not most effective	GEMS India (ICRW 2017) <sup>85</sup>
<b>Reduced pathway</b>	<b>ABc OR aC</b>	

A, social structural domain of implementation; B, implementation of student interpersonal components; C, high level of guided practice; an upper-case letter indicates the presence of a condition, while a lower-case letter indicates its absence.

(in which components tended to ‘stack’) did not support reliable estimation of components. Finally, we only undertook component NMAs for DRV outcomes as the data structures for GBV outcomes were not rich enough to support component NMAs.

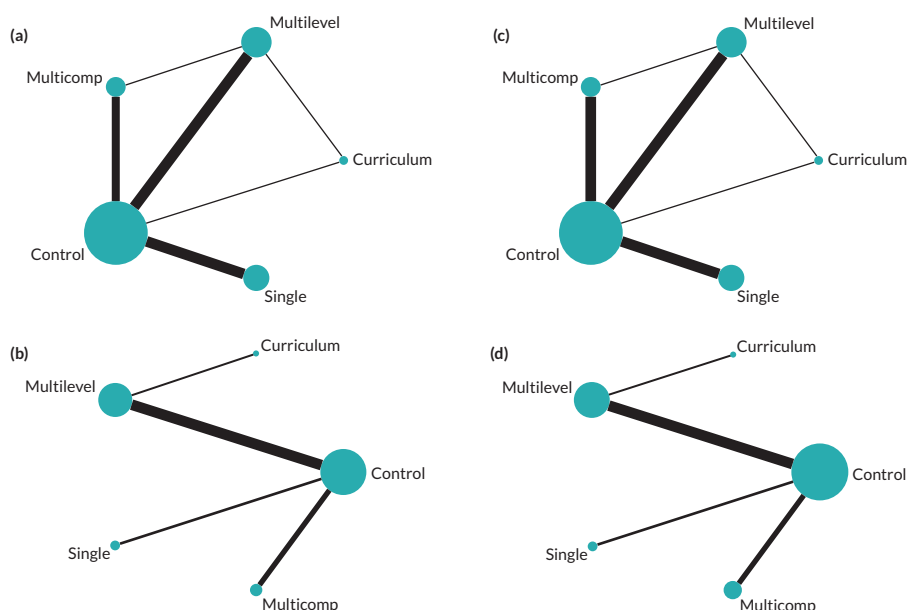
We proposed to assess potential intransitivity (i.e. the coherence of the network over multiple comparisons and trials) by effect modification using three key factors: age, LMIC context and percentage of girl children in the trial sample. In the course of our feasibility assessment, we did not note any meaningful or systematic variation by comparison in the network on the basis of age. However, we did note on the basis of prior metaregressions presented in [Chapters 6 and 7](#) that while LMIC context was a potential effect modifier for all analyses, percentage of girl children was primarily an effect modifier for victimisation outcomes only. Thus, we undertook network metaregressions for all analyses with the relevant effect modifiers, using a common regression coefficient across the network.

#### **Dating and relationship violence victimisation: short-term**

The network diagram for this NMA is displayed in [Figure 16](#). Based on results from this NMA ([Table 29](#)), there was no signal of any intervention type being significantly better than either control or any other intervention type. An overall design-by-treatment interaction test did not find significant evidence of inconsistency in the network ( $\chi^2 = 0.29$ ,  $df = 3$ ,  $p = 0.96$ ). Rankings of intervention effectiveness suggested that single-component and multilevel approaches were ranked highest, with SUCRAs of 0.8 and 0.7, respectively. Sensitivity analysis by either LMIC context or percentage of girl children did not change any conclusions.

#### **Dating and relationship violence victimisation: long-term**

The network diagram for this NMA is displayed in [Figure 16](#) and indicates that there is no inconsistency to be assessed in this network. Based on results from this NMA (see [Table 29](#)), both single-component interventions [OR = 0.60, 95% CI (0.41 to 0.86)] and multilevel interventions [OR = 0.83, 95% CI (0.69 to 1.00)] were significantly better than control. In addition, single-component interventions were more effective than multicomponent interventions [OR = 0.64, 95% CI (0.41 to 0.99)], though all other pairwise comparisons were not significant. Rankings of intervention effectiveness suggested that single-component interventions were most effective (SUCRA 1.0), followed by multilevel interventions (SUCRA



**FIGURE 16** Network diagrams for DRV outcomes. (a) Represents short-term DRV victimisation; (b) represents long-term DRV victimisation; (c) represents short-term DRV perpetration; and (d) represents long-term DRV perpetration.

TABLE 29 Network meta-analysis results for DRV victimisation and perpetration outcomes

DRV victimisation: short-term					DRV perpetration: short-term						
	Control	Single	Curriculum	Multicomp	Multilevel		Control	Single	Curriculum	Multicomp	Multilevel
Control		0.88 (0.75 to 1.03)	0.97 (0.70 to 1.34)	1.01 (0.84 to 1.22)	0.89 (0.76 to 1.05)	Control		0.81 (0.65 to 1.02)	0.90 (0.68 to 1.19)	0.99 (0.79 to 1.24)	0.87 (0.70 to 1.08)
Single	1.14 (0.97 to 1.33)		1.10 (0.77 to 1.57)	1.15 (0.90 to 1.47)	1.02 (0.81 to 1.27)	Single	1.23 (0.98 to 1.55)		1.11 (0.78 to 1.58)	1.22 (0.89 to 1.68)	1.07 (0.79 to 1.45)
Curriculum	1.03 (0.75 to 1.43)	0.91 (0.64 to 1.30)		1.04 (0.72 to 1.51)	0.92 (0.69 to 1.24)	Curriculum	1.11 (0.84 to 1.47)	0.90 (0.63 to 1.28)		1.10 (0.78 to 1.56)	0.96 (0.72 to 1.29)
Multicomp	0.99 (0.82 to 1.20)	0.87 (0.68 to 1.11)	0.96 (0.66 to 1.38)		0.88 (0.70 to 1.11)	Multicomp	1.01 (0.81 to 1.26)	0.82 (0.60 to 1.12)	0.91 (0.64 to 1.29)		0.87 (0.66 to 1.17)
Multilevel	1.12 (0.95 to 1.32)	0.98 (0.79 to 1.23)	1.08 (0.81 to 1.45)	1.13 (0.90 to 1.42)		Multilevel	1.15 (0.93 to 1.44)	0.94 (0.69 to 1.27)	1.04 (0.78 to 1.39)	1.14 (0.86 to 1.52)	
Wald ( $\chi^2$ to df to p)	0.29 to 3 to 0.96					Wald test ( $\chi^2$ to df to p)	0.16 to 3 to 0.98				

continued

TABLE 29 Network meta-analysis results for DRV victimisation and perpetration outcomes (continued)

DRV victimisation: long-term						DRV perpetration: long-term					
	Control	Single	Curriculum	Multicomp	Multilevel		Control	Single	Curriculum	Multicomp	Multilevel
Control		0.60 (0.41 to 0.86)	0.92 (0.61 to 1.40)	0.94 (0.73 to 1.20)	0.83 (0.69 to 1.00)	Control		0.57 (0.40 to 0.83)	0.97 (0.63 to 1.50)	0.83 (0.65 to 1.06)	0.86 (0.69 to 1.06)
Single	1.67 (1.16 to 2.42)		1.54 (0.88 to 2.69)	1.57 (1.01 to 2.45)	1.39 (0.92 to 2.10)	Single	1.74 (1.20 to 2.52)		1.69 (0.94 to 3.02)	1.44 (0.90 to 2.29)	1.49 (0.96 to 2.32)
Curriculum	1.09 (0.71 to 1.65)	0.65 (0.37 to 1.13)		1.02 (0.63 to 1.66)	0.9 (0.62 to 1.31)	Curriculum	1.03 (0.67 to 1.59)	0.59 (0.33 to 1.06)		0.85 (0.53 to 1.37)	0.88 (0.61 to 1.29)
Multicomp	1.06 (0.83 to 1.36)	0.64 (0.41 to 0.99)	0.98 (0.60 to 1.59)		0.88 (0.65 to 1.21)	Multicomp	1.21 (0.95 to 1.55)	0.70 (0.44 to 1.11)	1.17 (0.73 to 1.89)		1.04 (0.77 to 1.39)
Multilevel	1.20 (1.00 to 1.46)	0.72 (0.48 to 1.09)	1.11 (0.76 to 1.61)	1.13 (0.83 to 1.54)		Multilevel	1.17 (0.94 to 1.45)	0.67 (0.43 to 1.04)	1.13 (0.78 to 1.65)	0.96 (0.72 to 1.29)	

**Note**

Effects are expressed as ORs with 95% CIs, in which the column-forming intervention is compared to the row-forming intervention.

0.6) and curriculum and multicomponent interventions tied (SUCRA 0.4). Sensitivity analysis by either LMIC context or percentage of girl children did not change any conclusions.

#### Dating and relationship violence perpetration: short-term

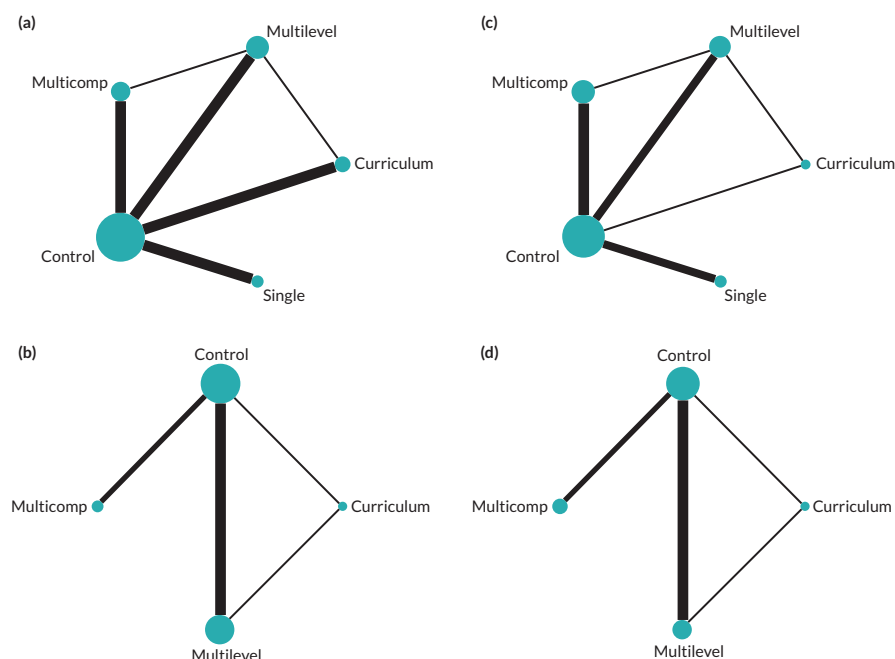
The network diagram for this NMA is displayed in [Figure 16](#). Based on results from this NMA (see [Table 29](#)), there was no signal of any intervention type being significantly better than either control or any other intervention type. An overall design-by-treatment interaction test did not find significant evidence of inconsistency in the network ( $\chi^2 = 0.16$ ,  $df = 3$ ,  $p = 0.98$ ). Rankings of intervention effectiveness suggested that single-component and multilevel approaches were ranked highest, with SUCRAs of 0.8 and 0.7, respectively. Sensitivity analysis by LMIC context did not change any conclusions.

#### Dating and relationship violence perpetration: long-term

The network diagram for this NMA is displayed in [Figure 16](#); it indicates that there is no inconsistency to be assessed in this network. Based on results from this NMA (see [Table 29](#)), only single-component interventions [OR = 0.57, 95% CI (0.40 to 0.83)] were significantly better than control, though no pairwise comparisons between active interventions in the network were statistically significant. Rankings of intervention effectiveness reflected this, suggesting that single-component interventions were dominant (SUCRA 1) with multicomponent interventions ranked second (SUCRA 0.6). Sensitivity analysis by LMIC context did change conclusions, suggesting that in addition to single-component interventions [OR = 0.57, 95% CI (0.40 to 0.82)], multilevel interventions [OR = 0.73, 95% CI (0.55 to 0.96)] were effective as compared to control in high-income contexts. In addition, curriculum interventions had a numerically improved effectiveness in this sensitivity analysis [OR = 0.82, 95% CI (0.52 to 1.30)].

#### Gender-based violence victimisation: short-term

The network diagram for this NMA is displayed in [Figure 17](#). Based on results from this NMA ([Table 30](#)), curriculum interventions were statistically better than control for this outcome [OR = 0.72, 95% CI (0.54 to 0.95)], but no other comparisons against control or any other intervention type were significant.



**FIGURE 17** Network diagrams for GBV outcomes. (a) Represents short-term DRV victimisation; (b) represents long-term DRV victimisation; (c) represents short-term DRV perpetration; and (d) represents long-term DRV perpetration.

TABLE 30 Network meta-analysis results for GBV victimisation and perpetration outcomes

GBV victimisation: short-term						GBV perpetration: short-term					
	Control	Single	Curriculum	Multicomp	Multilevel		Control	Single	Curriculum	Multicomp	Multilevel
Control		0.87 (0.66 to 1.14)	0.72 (0.54 to 0.95)	0.95 (0.77 to 1.16)	0.90 (0.75 to 1.09)	Control		1.00 (0.90 to 1.11)	0.88 (0.70 to 1.10)	0.95 (0.85 to 1.06)	0.89 (0.78 to 1.02)
Single	1.15 (0.88 to 1.52)		0.83 (0.57 to 1.21)	1.09 (0.78 to 1.53)	1.04 (0.74 to 1.46)	Single	1.00 (0.90 to 1.11)		0.88 (0.69 to 1.11)	0.95 (0.81 to 1.11)	0.89 (0.75 to 1.07)
Curriculum	1.39 (1.05 to 1.85)	1.21 (0.83 to 1.77)		1.32 (0.94 to 1.85)	1.26 (0.94 to 1.69)	Curriculum	1.14 (0.91 to 1.43)	1.14 (0.90 to 1.44)		1.08 (0.83 to 1.41)	1.02 (0.83 to 1.25)
Multicomp	1.06 (0.86 to 1.29)	0.92 (0.65 to 1.29)	0.76 (0.54 to 1.06)		0.95 (0.74 to 1.23)	Multicomp	1.05 (0.94 to 1.18)	1.05 (0.90 to 1.23)	0.93 (0.71 to 1.20)		0.94 (0.81 to 1.09)
Multilevel	1.11 (0.91 to 1.34)	0.96 (0.69 to 1.35)	0.79 (0.59 to 1.07)	1.05 (0.81 to 1.35)		Multilevel	1.12 (0.98 to 1.28)	1.12 (0.94 to 1.33)	0.98 (0.80 to 1.21)	1.06 (0.92 to 1.23)	
Wald test ( $\chi^2$ to df to p)	7.24 to 3 to 0.06					Wald test ( $\chi^2$ to df to p)	7.95 to 3 to 0.047				



GBV victimisation: long-term					GBV perpetration: long-term						
	Control	Single	Curriculum	Multicomp	Multilevel		Control	Single	Curriculum	Multicomp	Multilevel
Control		0.93 (0.66 to 1.32)	0.89 (0.75 to 1.07)	0.95 (0.77 to 1.18)		Control		0.82 (0.54 to 1.26)	0.89 (0.70 to 1.14)	0.95 (0.72 to 1.25)	
Curriculum	1.07 (0.76 to 1.52)		0.96 (0.67 to 1.37)	1.02 (0.69 to 1.53)		Curriculum	1.21 (0.79 to 1.86)		1.08 (0.66 to 1.76)	1.15 (0.79 to 1.68)	
Multicomp	1.12 (0.94 to 1.34)	1.05 (0.73 to 1.50)		1.07 (0.80 to 1.44)		Multicomp	1.12 (0.88 to 1.43)	0.93 (0.57 to 1.51)		1.07 (0.74 to 1.54)	
Multilevel	1.05 (0.85 to 1.30)	0.98 (0.66 to 1.46)	0.93 (0.70 to 1.26)			Multilevel	1.05 (0.80 to 1.38)	0.87 (0.60 to 1.26)	0.94 (0.65 to 1.35)		
Wald test ( $\chi^2$ to df to p)	5.39 to 1 to 0.02					Wald test ( $\chi^2$ to df to p)	8.12 to 1 to 0.004				

**Note**

Effects are expressed as ORs with 95% CIs, in which the column-forming intervention is compared to the row-forming intervention.

An overall design-by-treatment interaction test did not find significant evidence of inconsistency in the network ( $\chi^2 = 0.16$ ,  $df = 3$ ,  $p = 0.06$ ), but this was marginally non-significant and thus results from this NMA should be treated with caution. As expected, rankings from this NMA suggested that curriculum interventions were most effective (SUCRA 0.9), followed by single-component interventions (SUCRA 0.6). Sensitivity analysis by LMIC context did not change interpretation of point estimates of effectiveness, though all estimates now had CIs that crossed the line of no effect. A similar effect was seen when including percentage of girl children in the trial sample; furthermore, in this analysis, any differences between intervention types were minimal.

#### **Gender-based violence victimisation: long-term**

The network diagram for this NMA is displayed in [Figure 17](#); no single-component interventions entered into this analysis. Based on results from this NMA (see [Table 30](#)), there was no signal of any intervention type being significantly better than either control or any other intervention type. However, there was strong evidence of inconsistency in the network ( $\chi^2 = 5.39$ ,  $df = 1$ ,  $p = 0.02$ ), driven primarily by the multilevel-curriculum comparison represented by the Dating Matters trial.<sup>275</sup> Because the inconsistency model is statistically preferred to a consistency model, rankings are not presented as under inconsistency there is no one set of rankings generated. Sensitivity analysis by either LMIC context or percentage of girl children did not change any conclusions and thus did not present any plausible explanation for inconsistency in this analysis.

#### **Gender-based violence perpetration: short-term**

The network diagram for this NMA is displayed in [Figure 17](#). Based on results from this NMA (see [Table 30](#)), there was no signal of any intervention type being significantly better than either control or any other intervention type. However, there was some evidence of inconsistency in the network ( $\chi^2 = 7.95$ ,  $df = 3$ ,  $p = 0.047$ ), driven primarily and in the first instance by the multilevel-curriculum comparison represented by the Dating Matters trial.<sup>275</sup> Because the inconsistency model is statistically preferred to a consistency model, rankings are not presented as under inconsistency there is no one set of rankings generated. However, stratification by LMIC context in a sensitivity analysis (see [Report Supplementary Material 1](#)) did change the pattern of results, suggesting significant improvements against control for both curriculum [OR = 0.85, 95% CI (0.74 to 0.97)] and multilevel (OR = 0.85, 95% CI 0.75 to 0.97) interventions. It is thus likely that imbalance by LMIC context would resolve the inconsistency in this analysis.

#### **Gender-based violence perpetration: long-term**

The network diagram for this NMA is displayed in [Figure 17](#); no single-component interventions entered into this analysis. Based on results from this NMA (see [Table 30](#)), there was no signal of any intervention type being significantly better than either control or any other intervention type. However, there was strong evidence of inconsistency in the network ( $\chi^2 = 8.12$ ,  $df = 1$ ,  $p = 0.004$ ), driven primarily by the multilevel-curriculum comparison represented by the Dating Matters trial.<sup>275</sup> Because the inconsistency model is statistically preferred to a consistency model, rankings are not presented as under inconsistency there is no one set of rankings generated. Sensitivity analysis by LMIC context did not change any conclusions (see [Report Supplementary Material 1](#)) and thus did not present any plausible explanation for inconsistency in this analysis.

#### **Component network meta-analyses**

We undertook exploratory component NMAs for DRV outcomes. These are presented in [Table 31](#). In both short-term analyses, we could not separate social and environmental structural components as these were coterminous.

#### **Dating and relationship violence victimisation: short-term**

Analysis of short-term DRV victimisation outcomes did suggest that a component-based model generated a decrement in fit as compared to a model that treated component groups as distinct

**TABLE 31** Component network meta-analysis results for DRV outcomes

Domain	OR (95% CI)	Domain	OR (95% CI)
<b>DRV victimisation: short-term</b>		<b>DRV perpetration: short-term</b>	
Intrapersonal	0.92 (0.83 to 1.02)	Intrapersonal	0.91 (0.82 to 1.02)
Interpersonal	1.04 (0.92 to 1.18)	Interpersonal	1.04 (0.9 to 1.20)
Staff	1.10 (0.92 to 1.30)	Staff	1.08 (0.89 to 1.31)
Structural	0.87 (0.74 to 1.02)	Structural	0.89 (0.75 to 1.07)
Goodness of fit ( $\chi^2$ , df, p)	0.31, 3, 0.96	Goodness of fit ( $\chi^2$ , df, p)	1.73, 6, 0.62
<b>DRV victimisation: long-term</b>		<b>DRV perpetration: long-term</b>	
Intrapersonal	0.66 (0.49 to 0.89)	Intrapersonal	0.76 (0.52 to 1.10)
Interpersonal	1.02 (0.73 to 1.42)	Interpersonal	0.86 (0.57 to 1.29)
Staff	1.30 (0.87 to 1.94)	Staff	1.03 (0.64 to 1.66)
Structural environmental	0.90 (0.53 to 1.52)	Structural environmental	1.00 (0.54 to 1.85)
Structural social	0.99 (0.62 to 1.58)	Structural social	1.15 (0.68 to 1.93)
Goodness of fit ( $\chi^2$ , df, p)	14.09, 4, 0.007	Goodness of fit ( $\chi^2$ , df, p)	38.05, 4, <0.0001

interventions ( $\chi^2 = 0.31$ , df = 3,  $p = 0.96$ ). None of the four component classes tested generated a significant association with the outcome, though both intrapersonal and structural components were numerically associated with improved effectiveness.

#### Dating and relationship violence victimisation: long-term

Analysis of long-term DRV victimisation outcomes suggested that a component-based model did not fit the data as well as a component group-based model ( $\chi^2 = 14.09$ , df = 4,  $p = 0.007$ ). While the presence of an intrapersonal component class was associated with greater effectiveness [OR = 0.66, 95% CI (0.49 to 0.89)], none of the other component classes tested were significantly associated with improved outcomes.

#### Dating and relationship violence perpetration: short-term

Analysis of short-term DRV perpetration outcomes did suggest that a component-based model generated a decrement in fit as compared to a model that treated component groups as distinct interventions ( $\chi^2 = 1.73$ , df = 6,  $p = 0.62$ ). None of the four component classes tested generated a significant association with the outcome, though as with short-term DRV victimisation, both intrapersonal and structural components were numerically linked with improved effectiveness.

#### Dating and relationship violence perpetration: long-term

Analysis of long-term DRV victimisation outcomes suggested that a component-based model did not fit the data as well as a component group-based model ( $\chi^2 = 38.05$ , df = 4,  $p < 0.0001$ ). None of the component classes tested were significantly associated with improved outcomes, and only intrapersonal and interpersonal components were numerically linked with improved effectiveness.



## Chapter 9 Discussion

In this chapter, we summarise our findings and provide a realist account that integrates our syntheses, summarise the impact of our stakeholder consultation, relate our findings to previous research and describe the strengths and limitations of our work.

### Summary and realist reflection on empirical findings (RQ6)

#### *Programme components*

Most programmes to prevent DRV or GBV addressed both perpetration and victimisation. Fewer programmes addressed bystander intervention skills.

Most programmes aimed to equip potential perpetrators and victims with the capabilities and motivations needed to contribute to preventing DRV/GBV (either primary prevention of any such behaviour or secondary prevention of continuation/escalation of such behaviour). For example, many programmes aimed to provide students with the knowledge, attitudes or skills needed to prevent perpetration or victimisation. Student components included:

- guided practising of skills
- group discussions
- individual reflection; visual/image- or narrative-based learning
- student competitions in class.

These components could be guided by adults or peers, or be self-paced. Most components were delivered to mixed rather than single-sex groups. In addition, many programmes aimed to develop staff capabilities and motivations to address DRV or GBV. These could prepare staff to deliver learning for students, or target staff directly as change agents within schools. Staff components offered training or lesson plans/scripts. A few programmes also included components addressing parent/family involvement (which could range from sending out materials to involvement in intervention activities).

Some programmes aimed to modify school social structures or physical environments to prevent DRV/GBV. Components addressing school social structures included visits from community organisations, changes to school policies, participative customisation of activities and school clubs. Components addressing the physical environment included visual materials in shared space and monitoring the safety of school spaces.

Overall, programmes could be categorised discretely as: single-component programmes (usually short and focused on a novel technology or activity) which generally focused on a single change mechanism; curricular programmes integrated into the wider school curriculum, which generally focused on one or more change mechanisms at the student level; multicomponent programmes involving multiple modes of intervention and focusing on multiple change mechanisms operating at the student or staff level; and multilevel programmes, which were complex interventions involving multiple modes focused on change mechanisms within schools at multiple levels including the individual, classroom and school structural/environment context.

#### *Programme theory*

Drawing on programme theories of change, we theorised a complex set of mechanisms which involved school staff drawing on programme resources to transform school organisation to promote students' school belonging and engagement with pro-social behaviours. This was theorised to occur via programmes aiming to achieve 'weakened classification' and 'reframing'. 'Weakened classification' is defined as involving the strengthening of relationships between and among staff and students.

It also involved strengthening relationships between the classroom and wider school and between professional roles (which were in turn theorised to help ensure consistent responses to violence and harassment, ensure staff supervision of school spaces, and provide students with a role in school-level decision-making). Weakened classification was also theorised to involve strengthening positive relationships between schools and communities, in order to provide better parent/family support and involvement, and allow better student access to community support services. Conversely, some programmes, particularly in high-crime areas, could be understood in terms of a mechanism increasing some boundaries between the school and community in order to reject pro-violence community norms. 'Reframing' is defined as aiming to increase student involvement in decisions at the level of the classroom (e.g. interactive, experiential learning) and the school (e.g. contribution to policies and decisions).

'Weakened classification' and 'reframing' were in turn theorised to engender increased students' sense of school connection and sense of safety in the school building. In turn, this could encourage increased learning of pro-social skills, and increased pro-social peer norms and interactions, associated with potential reductions in DRV and sexual harassment.

Programmes varied in the extent to which they addressed all or some aspects of the above mechanism. Multilevel programmes aimed to trigger such mechanisms at multiple levels of the school system including the individual, classroom and school. Classroom-level interventions aimed instead to trigger mechanism at the level of staff-student relationships, reframing learning to be more interactive and engendering increasing student commitment to learning.

Other programmes appeared to focus on other mechanisms which did not involve wholesale transformations to school organisation or improvements in student commitment to school. Instead, these mechanisms aimed to promote specific individual or group capabilities or motivations for preventing or terminating DRV/GBV among students and staff. These could involve a mechanism promoting students' positive development and behaviours (e.g. conflict management within relationships). Another mechanism could instead merely involve the prevention of negative behaviours (e.g. by identifying DRV or GBV as unacceptable behaviours).

Based on these review findings about theorised mechanisms, it was hypothesised that multilevel programmes triggering the above weakening of classification and reframing mechanisms across multiple levels, within the school and in terms of the school's relationship with families and the community, would achieve larger effects and be more sustainable (because they might effect an enduring transformation in school organisation and social structures) than the other mechanisms. Based on these review findings about theorised mechanisms. It was also hypothesised that programmes triggering mechanisms promoting individual- and group-level capabilities and motivations rather than school organisational changes would achieve smaller effects and be less sustainable (because these involved fewer mechanisms of prevention and did not aim to transform school organisation or social structures).

### ***Influences on successful implementation***

At the school level, implementation was facilitated by school resources and infrastructure, such as space, supplies and organisation (which increased ease of implementation and reduced the necessity of programme modification); availability of space or supplies (which increased the fidelity of implementation); school organisation and leadership capacity (which enabled better fidelity of implementation); time constraints such as the structure of the school day and competing priorities such as academic performance (which could hamper implementation and hence reduce fidelity and dose); and positive school perceptions of the importance of addressing DRV/GBV and comprehension of the importance of the topic to their student body (which could facilitate adoption of interventions, and promote staff commitment to delivery). These factors could cross-compensate so that, for example, where there was strong staff commitment to DRV/GBV prevention, this could compensate for time or resource limitations.

At the student level, implementation was facilitated by intervention interactivity (which could increase engagement and acceptability); good student attendance across population subgroups (which could facilitate reach); the development of positive relationships among students (which could increase student engagement over time); programme fit to the student population, for example in terms of culture, age, academic background, sexual orientation and disability (which could increase fidelity, engagement and acceptability); and student maturity and positive behaviours (which could facilitate delivery and engagement).

A number of facilitator-level factors enabled implementation. One of these was facilitator content knowledge (which could promote fidelity and acceptability). There was more variation in knowledge among school-based than external facilitators. Other factors were external support (which could increase feasibility) and the availability of school-based facilitation (which could increase sustainability by embedding interventions in school curricula).

Programme-level factors included the ease of delivery (which could promote fidelity); and the ease of intervention modification to the particularities of the setting (which was associated with better delivery). These last findings might suggest that simpler (single-component) programmes might work in a greater diversity of schools than more complex programmes.

Based on these findings from the review of process evaluations, it was hypothesised that multilevel programmes triggering the above classification and framing mechanisms across multiple levels would only work in school contexts with high organisational capacity and high resourcing. Based on these findings about process, it was further hypothesised that programmes triggering mechanisms involving individual- and group-level capabilities and motivations would work in all school contexts including in more resource-poor contexts. Also, it was further hypothesised that such programmes would be more effective when they were locally modifiable and interactive, and involved support from external agencies. It may be that programmes are best delivered when they align with school organisational readiness, defined as a receptive school climate, staff buy-in or strong school leadership. This was associated with better implementation.

### ***Programme effects on outcomes***

There were overall programme effects on long-term but not short-term DRV perpetration and victimisation. A possible reason for this is that many of the programmes involved changes to school organisation, social structures and physical environments, and that such changes required time to achieve and to result in benefits to students. Even where simpler programmes addressed individual-level mechanisms involving changes in capabilities and motivations, these may have required time to generate reductions in DRV because of the time required for students to enter new relationships in which these capabilities and motivations might translate into reduced violence. The long-term nature of the impacts might also be explained by developmental windows, so that behaviours manifest as students enter particular psychosocial developmental periods, for example coinciding with the initiation of dating behaviours.

There were no overall programme effects on GBV victimisation or perpetration. This might be because, whereas DRV tends to be a behaviour occurring in the private context of dating and relationships, GBV is a more public activity. Consequently, while DRV might be more amenable to change via partners learning new capabilities and motivations in relation to gender, relationships and violence, GBV may tend to be somewhat more influenced by social norms about how students behave in front of their peers, which may be more challenging to modify. However, despite the lack of overall effects on GBV, there was some evidence in HICs only of long-term effects on GBV victimisation and perpetration. This might have occurred because, although challenging to address for the reasons stated above, GBV can be reduced by longer-term transformations in higher-capacity school systems when supported perhaps by broader social and normative trends.

Programme components generally did not readily explain effectiveness. However, there was some evidence that single-component programmes were more effective for long-term DRV victimisation and perpetration. This is a surprising finding given our expectation that programmes aiming to weaken classification and framing via modifying school organisation would be more sustainable. A possible explanation for this finding is that simpler programmes involving single components might, as we hypothesised, be more feasible to deliver in a greater breadth of schools and so, even if they are somewhat less effective and sustainable when delivered, their effects over the long term may succeed in being built up because they are more often well delivered in more schools. There was no evidence for programme effects on bystander outcomes, adding to the picture that programmes are not effective in disrupting mechanisms that generate socially normative, publicly visible violence.

### ***Evidence of programme effects on intermediate outcomes***

Intermediate outcomes are those proximal outcomes that the intervention is aimed to address in order to achieve impacts on DRV or GBV. There was some evidence that, in the short but not the long term, programmes reduced DRV violence acceptance, DRV knowledge, attitudes to intervening and attitudes to personal help-seeking but not individual self-efficacy or social norms. This might be interpreted as further evidence that current DRV programmes are most likely to achieve their effects via changes in individual capabilities and motivations, and not via social norms. However, it might also be that social norms are challenging to measure so that any effects on them were imperfectly assessed. There was some evidence that, in the short and long term, programmes reduced GBV violence acceptance and GBV knowledge but not attitudes to intervening, attitudes to personal help-seeking or social norms. There was also evidence for short-term effects on individual self-efficacy. This might be interpreted as evidence that, if programmes are to achieve impacts on GBV, this is unlikely to occur via impacts on social norms (at least with regard to interventions currently evaluated).

### ***Mediation***

#### **Factors mediating dating and relationship violence outcomes**

Mediation analyses examine whether programme effects on intermediate outcomes might explain programme effects on distal outcomes. There was some evidence that reductions in violence acceptance mediated programme effects on DRV victimisation. There was also some evidence that reduced violence acceptance mediated programme effects on DRV perpetration. There was inconsistent evidence that increased knowledge mediated programme effects on DRV victimisation and perpetration. Single studies found evidence that gender stereotyping attitudes mediated programme effects on DRV victimisation but that conflict-management skills, belief in need for help or school belonging did not do so. Single studies found evidence from some analyses that gender stereotyping and awareness of the need for help mediated programme effects on some DRV perpetration outcomes but that conflict-management skills did not do so. There was no evidence that bystander actions or school belonging mediated programme effects on DRV perpetration. This adds to the picture that these programmes might reduce DRV primarily via impacts on individual capabilities and motivations than group norms, and that the most important individual capabilities and motivations centre on knowledge, and attitudes towards violence and possibly gender. The evidence suggests that conflict-management skills and sense of school belonging were less implicated in the impacts of these programmes. However, it may be that some of these null findings in relation to mediation reflect measurement problems.

#### **Factors mediating gender-based violence outcomes**

There was evidence from one study that a programme reduced GBV victimisation outcomes by improving school belonging. A single study of mediation by school belonging found evidence for this occurring for some but not all GBV perpetration outcomes. Single studies examined mediation of GBV perpetration, finding stronger evidence for mediation by violence acceptance than for mediation by bystander actions. A single study found evidence that delinquency behaviours mediated programme effects on GBV perpetration. This evidence might suggest that, where GBV is prevented, this is most likely to occur via mechanisms involving transformations to the school social environment, though



evidence from the analysis of programme outcomes on GBV suggest that such mechanisms might rarely be triggered sufficiently to achieve significant reductions in GBV.

### **Factors mediating bystander outcomes**

A single study examined whether self-efficacy mediated bystander outcomes, finding evidence that this was the case for observed but not self-reported bystander behaviour.

### **Moderation**

#### **Factors moderating dating and relationship violence victimisation**

Our syntheses suggested that programme effects on DRV victimisation were not moderated by gender, prior experience of DRV victimisation, dating history, age, ethnicity, acculturation or sexual orientation. There was some evidence from metaregression that the proportion of the study sample who were female was associated with an increase in effectiveness for preventing DRV victimisation. This might be interpreted as evidence that school programmes are more likely to be effective in settings where a critical mass of female students might have encouraged greater overall student engagement with intervention activities and hence the de-normalisation of violence. However, this effect did not emerge in the findings from the synthesis of process evaluations (e.g. that increases in the proportion of students who were female was associated with improved programme delivery or reach) and so should be treated with caution.

#### **Factors moderating dating and relationship violence perpetration**

There was evidence from multiple studies for gender moderating programme effects on DRV perpetration with greater benefits for boys, particularly for emotional and physical DRV perpetration. There was weaker evidence that programme effects were greater for those with prior experience of DRV perpetration. There was no evidence that dating history, age, ethnicity, acculturation or sexual orientation moderated programme effects on DRV perpetration. There was no evidence from meta-regression that the proportion of the study sample who were female was associated with a difference in effectiveness for DRV perpetration. The finding that, for some DRV perpetration outcomes, effects were larger for boys suggests that these programmes were not gender-neutral. The programmes may have been interpreted by students as programmes aiming to reduce male perpetration of DRV and, also informed by the above findings on mediation, might have achieved these effects via changes in male attitudes to violence. There was weak evidence that such mechanisms might have been slightly stronger among those who had previously engaged in perpetration.

#### **Factors moderating gender-based violence victimisation**

There was some evidence that programmes had greater effects reducing GBV victimisation among boys than girls, driven primarily by longer-term evidence. Two studies examined prior experience of GBV victimisation finding little evidence of moderation. There was also little evidence of moderation by age, ethnicity, sexual orientation or poverty status. There was some evidence from metaregression that the proportion of the study sample who were female was associated with an increase in effectiveness for reducing GBV victimisation. The finding of a larger effect for male than female GBV victimisation is surprising and hard to interpret. The finding that effects on GBV perpetration were larger in schools with more female students aligns with the evidence above on DRV victimisation and might be explained by greater student engagement.

#### **Factors moderating gender-based violence perpetration**

There was some, albeit patchy, evidence that programmes were more effective for girls than for boys, and some evidence that programmes were more effective for those with prior experience of GBV perpetration. There was little or no evidence for moderation by age, ethnicity, sexual orientation or poverty status. There was no evidence from metaregression that the proportion of the study sample who were female was associated with a difference in effectiveness for GBV perpetration. The finding that programmes may have been more effective in reducing GBV perpetration among girls more than boys is unexpected and not easily interpreted. It might possibly be that programmes encouraging girls

not to perpetrate GBV were more novel and therefore impactful than similar messages concerning male perpetration. However, this is speculative and this finding adds to the overall picture that GBV programmes do not work as theorised.

### ***Explaining heterogeneity in effectiveness***

As noted previously, intervention components did not generally explain effectiveness via meta-regression. There are a number of reasons why this could be the case, including differential implementation across trials and non-differential misclassification due to the difficulty of identifying intervention components. However, a likely reason why is that single-component meta-regressions do not reflect the complexity of intervention functioning, especially in a non-psychotherapeutic context. A similar finding was in evidence from our component NMAs for DRV outcomes, which did not provide clear signals as to the differential contribution to effectiveness of component classes. However, QCAs were substantially able to differentiate between most effective interventions and not most effective interventions, with the exception of short-term DRV perpetration.

An important finding from our QCA is that a central causal condition for reduction of victimisation is reduction of perpetration, across short-term and long-term DRV victimisation and short-term GBV perpetration. However, a number of other pathways to the reduction of victimisation exist, generally characterised by the inclusion of single-gender components or a critical mass of girls. A critical mass of girls was especially important where interventions went beyond single components. There was some evidence relating to 'proscribed components', or components whose absence was part of a causal pathway to most effectiveness. For long-term DRV victimisation, the absence of parental involvement was central to achieving effectiveness; similarly, for short-term GBV victimisation, the absence of narrative components was an important part of causal pathways. It is possible that the absence of these components reduces opportunities to minimise GBV and DRV, or to receive conflicting messages about their importance.

However, if reductions in victimisation are principally achieved through reductions in perpetration, how are reductions in perpetration achieved? A key limitation of our QCA was our inability to develop a satisfactory model for short-term DRV perpetration. However, for both long-term DRV perpetration and short-term GBV perpetration, interventions that were most effective incorporated a range of opportunities for guided practice of skills and attitudes, and interpersonal components focusing on student relationships. Importantly, the implementation of social structural components was central to effectiveness for short-term GBV perpetration, but not for DRV perpetration.

Another way in which we explored heterogeneity in effectiveness was via NMAs by intervention type. Our findings largely confirmed the findings of our pairwise meta-analyses; specifically, that there was limited evidence that any one intervention type was significantly better than any other, but also that single-component interventions appeared unusually effective as compared to control. However, NMAs for GBV outcomes were generally inconsistent, suggesting that additional evidence is needed to understand the comparative effectiveness of intervention types.

### ***Realist integration***

Drawing on syntheses of theories of change and process evaluations, we hypothesised that multilevel programmes triggering mechanisms that involved the weakening of classification and reframing would achieve larger effects and be more sustainable than programmes triggering individual- and group-level capabilities and motivations. We hypothesised that the multilevel programmes would only work in schools with high organisational capacity, whereas simpler programmes would work in all schools (see [Figure 1](#)).

Our subsequent analyses did not support these findings. We found that the programmes were effective in reducing DRV perpetration and victimisation (in the long but not the short term) but not GBV victimisation and perpetration. There was some evidence that the programmes we reviewed worked not

by complex mechanisms involving transformation of school organisation and social structure, but via increasing student capabilities and motivations concerning the unacceptability of violence. This might be termed the 'basic safety' mechanism which aimed to destabilise or interrupt violent behaviours via communicating the unacceptability of violence. This would be in contrast to more complex mechanisms involving, for example, increased school commitment and the development of social norms supporting a range of positive behaviours that actively displace violent behaviours. Support for this comes from evidence that bystander programmes were not effective and that students' school belonging and social norms were not mediators of programme effects.

There was evidence that this simpler basic safety mechanism may have involved reductions in DRV perpetration among males and those with previous experience of perpetration. That simpler interventions if delivered well may be effective in reducing DRV would be a significant finding, especially for low-income countries and other resource-poor settings. Another possibility is that interventions did impact on student school belonging and social norms, and that these mediated impacts on DRV, but that this was not detected because of limitations in how school belonging and social norms are measured. Further research on the reliability and validity of these measures would be useful. A further possibility is that the complex mechanisms are capable of generating reductions in both DRV and GBV but that the particular interventions that have been evaluated to date are not capable of triggering these mechanisms.

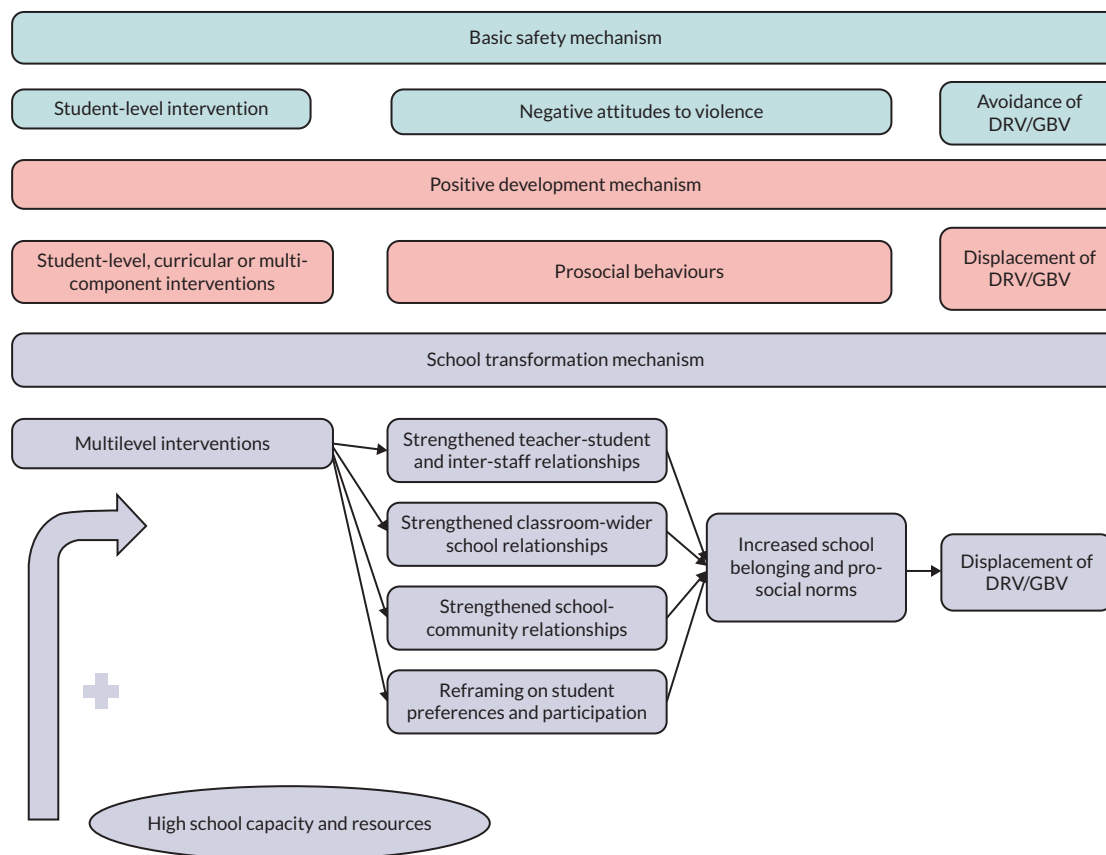
We theorise that individual-level basic safety mechanisms are more likely to effect changes in DRV rather than GBV perpetration because of the more private, dyadic nature of DRV (meaning it is amenable to reduction via changes in partners' capabilities and motivations, and it does not require changes in broader social norms). However, the more public nature of GBV means that it might be influenced by social norms which programmes appear not to successfully address. DRV programmes are likely to work via mechanisms involving modifying boys' knowledge of, and attitudes to, DRV and may generate reductions in DRV across school contexts. There is some evidence from single studies that attitudes to gender might also mediate effects on DRV victimisation and perpetration. Our findings suggest that even these individual-level mechanisms are likely to require time, perhaps because of the time that students take to become involved in relationships in which violent behaviours might have occurred in the absence of the theorised mechanisms. The long-term nature of these benefits might have arisen because even impacts occurring via relatively simple mechanisms require time to manifest (e.g. because of the time needed for young people to form new relationships) or because changes were sensitive to development windows. The long-term nature of these impacts might create challenges persuading schools to deliver such interventions. A further issue is that, in terms of sector, benefits arise for the health sector but costs are incurred by the education sector. All this suggests a joined approach is required, where public health funding supports programme delivery in schools. Future evaluations providing evidence of cost effectiveness would also be useful in persuading take-up in the education sector.

We found that the programmes could be effective in preventing GBV but that this was only likely in high-income settings. It is possible that transformation of school organisations and social environments is needed to reduce GBV and that this requires a context of existing high school organisational capacity. This is especially supported by the findings from the QCA that social structural components are important parts of causal recipes in reducing short-term GBV perpetration. Such work is more challenging in low- and middle-income settings because of lower school capacity to effect institutional and social transformation, and higher levels of inequality and GBV. Schools would be less ready to implement such complex programmes. It may also be that transformations in GBV-related social norms are more feasible in HICs because of the greater malleability of norms concerning gender and violence in contexts where traditional gender norms are increasingly being eroded. Another possibility is that impacts on GBV are harder to achieve in schools in some low-income countries because of the involvement of school staff in GBV in some contexts.

Implementation evidence suggested that simpler interventions were easier to implement and receive. This might explain why there was some evidence that simple, single-component programmes may have achieved larger effects via the basic safety mechanism described above. There is some support for this from our synthesis of process evaluations, which suggested that school capacity and resourcing were key limiting factors, and that programmes ease of modification and delivery facilitated better implementation. There was also evidence that simpler programmes of this sort were easier to deliver in a greater diversity of school contexts. It may therefore be more important to deliver simpler programmes achieving basic levels of school safety with fidelity than attempt more complex, multicomponent or multilevel programmes. It may be particularly important to prioritise such programmes in resource-poor settings. Some schools may be better placed to implement more complex multicomponent or multilevel programmes in terms of a receptive school climate, staff buy-in or strong school leadership. It may be that a one-size-fits-all approach is not appropriate and programmes need to align with schools' baseline readiness. Interventions involving initial need and capacity assessment have not been trialled in prevention of DRV/GBV.

In terms of gender, we found some evidence that programmes achieved larger effects in schools with more female students, which might be explained by the greater aggregate student engagement with programmes in schools populated by girls. Our findings that the effects of GBV programmes on victimisation were greater for boys and on perpetration were greater for girls is counterintuitive (we may speculate that this reflected the greater novelty of these messages for girls compared to boys) and suggests that the mechanisms of such programmes are not as currently theorised.

Our findings are summarised below, and in [Figures 18](#) and [19](#).



**FIGURE 18** Hypotheses arising from syntheses.

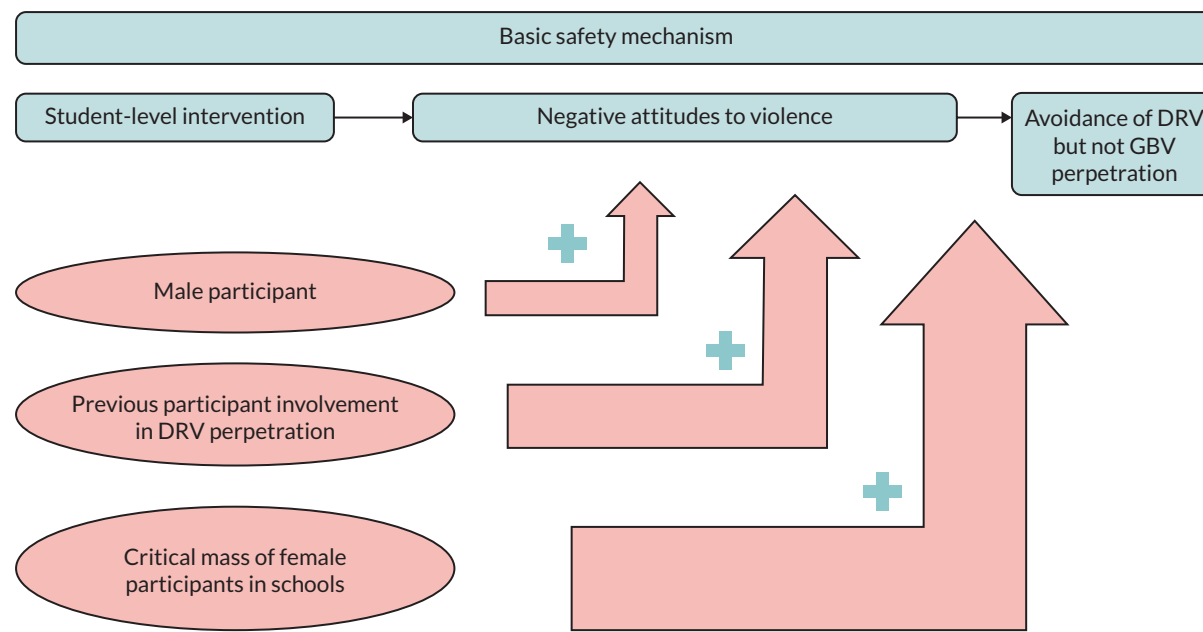


FIGURE 19 Realist integration of findings.

- **It may be less important to do something complex via multilevel interventions than to do something simpler to a higher standard, via well-implemented single-component interventions.** This is, in many respects, the central finding of our syntheses.
- **School resourcing and basic safety are important.** Evidence from our syntheses suggested that under-resourced schools and schools in areas of high deprivation may have a greater focus on ensuring the 'basic safety' of students by suppressing harmful behaviours rather than encouraging positive behaviours.
- **Destabilising harmful practices is foundational, whereas establishing prosocial behaviours will take more time and investment.** Relatedly, interventions may focus on reducing harmful behaviours by making them less desirable but may also seek to replace these with prosocial behaviours. It may be that schools focus on foundational mechanisms before engaging with broader prosocial mechanisms.
- **The basic safety mechanism might be more able to generate reductions in DRV than GBV.** DRV might be more influenced by individual-level capabilities and motivations amenable to change via simple basic safety interventions, whereas GBV might require normative change achieved through more complex programmes aiming to displace violence via more positive norms and behaviours.
- **School preparation and intervention choice provide important context.** School preparation, whether defined as a receptive school climate, staff buy-in or strong school leadership, was linked with smoothing the path to implementation and unlocking a wider range of mechanisms beyond strictly student-directed mechanisms.
- **Female critical mass may matter.** Programmes might be more effective when targeting populations with a critical mass of young women, perhaps reflecting greater aggregate engagement.
- **In high-income but not low-income countries, it may be that interventions can reduce GBV** perhaps by schools having more capacity to achieve transformations, in potentially more conducive societal contexts.

## Consultation with stakeholders

We sought feedback and advice throughout the project from a group working in practice settings or interfacing with education and violence prevention policy. We convened three online meetings with the group to share our review findings and obtain their input on our analytical focus and interpretation. We fed this consultation into our investigator discussions and analysis.

Our *first meeting* of the advisory group was held in October 2020. We focused the discussion on two substantive issues that went on to frame our approach for categorising and analysing intervention studies in the review: what DRV and GBV education look like in England right now, and what priorities for DRV and GBV education would be most salient for this review.

### ***What does gender-based violence/dating and relationship violence education look like in England right now?***

There was consensus that provision in England (at the time, 2020) is patchy, with schools adopting different approaches to violence prevention education and offering varied support to students. The content and quality of delivery are often reliant on the skills and knowledge of individual teachers designing curricula and/or delivering to students and there was agreement that school-wide culture change is necessary if intervention is to be effective. A key factor determining divergent provision in England is the nature of schools' administration and governance; Local Authority-controlled schools differ to independent schools and academies in the nature and scope of what they can/must deliver to students. This sits in contrast to the context in Wales, where the VAWDASV Act (2015) requires schools to have a violence against women and girls (VAWG) policy. This is supported by a toolkit of educational resources to inform curriculum reform and the Welsh Government encourages a WSA to violence prevention.<sup>278</sup>

The advisory group also felt that the current concepts informing curriculum and intervention design were siloed and labelling. In particular, GBV curricula tend to adopt binary or gendered perspectives, which are unhelpful for lesbian, gay, bisexual, transgender, queer (LGBTQ) and non-binary students. Connections between topics (like DRV and GBV) may not be clear and some issues, like honour-based violence or child sexual abuse, are simply absent. Moreover, distinctions between DRV and GBV were often unclear, due in part to overlap between behaviours. It was acknowledged that the online training modules provided by the Department for Education were helpful tools for improving the knowledge base of staff in educational settings; however advisory group members agreed that this training was being used in a piecemeal, ad hoc way rather than leading to comprehensive understanding of the issues. Finally, there was a sense that COVID-19 had changed the context for schools particularly in relation to RE; practice advisors commented that the demand for information and support had increased, both from young people navigating difficult relationships and those trying to support them. This is especially alarming given anecdotal evidence that RSE and personal, social, health and economic education (PSHE) declined in UK schools during the pandemic, due to prioritising other lessons.

### ***What are the priorities and opportunities in gender-based violence/dating and relationship violence education for this research to focus on?***

While school curricula devoted to RE, RSE and health education are mandatory in England, there is a need and an opportunity to improve the consistency and effectiveness of messaging by schools. Schools and teachers require more training to build knowledge and confidence, and, critically, time to think through topics in the curriculum and their own starting place (values, beliefs and attitudes) with these issues. The advisory group felt there was also an opportunity for specialist RSE provision, tailored for subgroups of young people with particular inclusion needs. This may also require a spectrum of interventions, for example combining small group support and curricula-style input.

The advisory group signalled four priority areas for intervention research to consider:

- School leadership: Head teachers and other school leaders (governors, etc.) are critical to ensuring that time and resources are available to teachers and classroom assistants to access training and develop skills/competency (not just knowledge). They are also the lever for adopting WSAs and embedding culture change regards violence acceptance and social norms.
- Subject experts: In addition to their community-based service provision, many specialist, third-sector organisations are involved in delivering school-based interventions but funding models to recognise

this are not well developed. There are opportunities for better health and education partnership working, and integrated care pathways to identify and support vulnerable young people.

- Student voice: Identifying methods/processes for intervention content to be driven by, and tailored to, students' needs is a priority. As a rule, content should be 'experience near' – that is, related to young people's personal experience and understanding rather than established knowledge or expert opinion – and not driven by age and/or development. More could also be done to link DRV/GBV topics to young people's rights.
- Context: A more nuanced understanding of context is critical to improving research about what works and how it works for young people. DRV/GBV interventions are operating in wider RSE and PSHE frameworks as well as wider school and community contexts and culture. A focus on shared mechanisms – that is, pathways to effectiveness across different forms of violence – may be most helpful.

Our *second meeting* of the advisory group was held in June 2021. Members of the research team presented on included studies in the systematic review, the theoretical synthesis of intervention programme theories, and intervention typology options, as well as data from a synthesis of process evaluations. The advisory group were happy that the review inclusion results were clear, and that theory synthesis and component typologies were relevant and relatable. The group reflected on the breadth and complexity of the implementation data gathered from process evaluations; they also supported the finding that the ease of intervention modification (and low demand) is related to implementation success.

In the context of these results, we explored two key questions: what are the gaps in the evidence base, and how can we make the most of the findings of this review?

### ***What are the gaps in the evidence base and what are the other ways in which we should consider/categorise the studies and the interventions in the review?***

The group noted an absence of studies focusing on LGBTQ young people. Only two studies focus on homophobic violence and examined the relative benefits of programmes for LGBTQ students. There was a suggestion that impacts by student-level factors, such as minority ethnic group or age, may be particularly important to disaggregate. For example, do results from US studies of middle school-aged children mean we should be working with Year 5 and 6 primary school students in the UK? Advisory members also highlighted the need to consider the way in which social norms, school culture and school policies/responses (to violent/abusive behaviour) explain the differential effectiveness of interventions. The concept of 'context' – signalled as a priority area in the first advisory group meeting – was raised again. In particular, the group questioned whether it was possible to distinguish baseline differences in schools (context) where interventions had been more (or less) effective. This might include differences in whether teachers felt empowered to respond to situations of DRV/GBV, or differences in the staff group/senior leadership in their openness to learning and change.

### ***How can we make the most of the alignment of this review with existing agendas and approaches?***

Advisory group members felt that a lack of oversight of RSE curricula and programmes would continue to result in diverse provision, of varied quality, and that interventions needed to align with central (PSHE), evidence-based guidance. Many organisations working with young people and families affected by DRV/GBV are adopting trauma-informed approaches to their provision; the group questioned whether any interventions in the review were framed in similar ways. Advisory members recommended that school-level policies – on DRV/GBV and wider issues – be regarded as part of the intervention in this space, that is a multilevel approach. Approaches for dealing with other forms of oppression, such as racism, set a precedent for responses to GBV/DRV and the group felt it was important to highlight interactions between GBV, DRV, hate crime, and safeguarding requirements in schools. Approaches that support assessments of how safe young people feel in school, via pastoral provision or Mental Health Support Teams were all aligned with a strong violence prevention agenda. In addition, noting

how interventions work with established pathways for referral to specialist support and/or community VAWG partnerships were also ways to ensure alignment with wider agendas.

Finally, consideration of school exclusion policies may be important for enhancing impact of DRV/GBV interventions. Where perpetrators are excluded from school settings, this may deny them access to learning or exposure to prosocial norms. Similarly, policies that require parental permission/consent for involvement in extended RSE programmes may limit reach and impact.

Our *final meeting* with policy and practice advisors was held in March 2022. We presented findings from the meta-analysis of effectiveness data, moderator and mediator analyses. We used the meeting to consider one substantive question with the policy and practice group.

### ***What do our findings mean for intervention functioning, and for current provision in the UK?***

The lack of evidence for short-term effectiveness of interventions for DRV or GBV was surprising for the group, particularly considering long-term impacts on victimisation in girls and for DRV and GBV interventions in high-income countries. There was a question as to whether this might indicate a potential sleeper effect of universal prevention curricula with younger students – with impacts only showing later in their lives when they engaged in dating or intimate partner relationships. There was agreement that the positive signals on some of the short-term mechanisms/mediators, such as violence acceptance and attitudes to intervening and help-seeking, were helpful for schools to consider as they designed their RSE programmes. The group agreed that findings support the need for better-quality and long-term outcome studies, given the likely link between these in the current review.

The group also discussed the potential value (or not) of a WSA to DRV/GBV prevention. While programme theories advocated comprehensive, multilevel interventions as recommended practice, the effectiveness data do not necessarily support this as standard. Advisors agreed, however, that a recommendation for well-delivered ‘simple’ interventions nuanced by the importance of context would be welcomed by schools. This also raised implications for school/organisational ‘readiness’ to implement. A receptive school climate with staff buy-in, strong leadership and connections to families are likely pre-conditions for intervention impact. Advisors reflected that the schools represented in intervention trials were less likely to reflect those most in need of support, where the safeguarding load is high and where staff and leadership are regularly firefighting. That said, high-profile incidents can motivate change and be a catalyst for intervention readiness.

There were no clear patterns in our findings for age or sexual and gender identity, which was surprising to stakeholders but may reflect a lack of precision in the measurement or disaggregation of categories. For example, wide age ranges (10–18 years) may mask important developmental differences. The advisory group felt that there were implications for current prevention approaches, in particular universal prevention, with the weight of evidence targeting secondary and/or tertiary consequences of DRV/GBV. Most educational curricula are designed for universal delivery and these data may necessitate greater diversity in provision, for example a menu or suite of programmes tailored for different levels/needs.

Advisors suggested that the importance of the wider culture of the school and broader community contexts on GBV and DRV were critical to emphasise. Schools and providers should reflect on the common roots underlying issues, for example, addressing the gender stereotypes to target DRV/GBV may also help with challenging homophobia and bullying. Teachers and intervention facilitators need to make these strands work together; this may be both efficient, in curriculum time/resources, and impactful by consolidating learning and skills.

Finally, advisors highlighted the importance of strong messaging around bystander interventions. These approaches are gaining popularity in schools and many settings are looking for programmes to develop



their provision in this space. But the evidence base is weak, particularly in young adolescents, and may be unsuitable for recommendation as part of RSE curricula in schools.

## Consultations with young people

We also undertook a parallel set of consultations with young people. These consultations are ongoing as part of our dissemination work. In particular, we met with the Association for Young People's Health specialist youth advisory group twice, with a group of experts by experience twice, with the ALPHA group at DECIPHer once, and with a specialist girls' and young women's project in the South East. Their comments shaped our interpretations in several ways.

### *Immanence of gender-based violence and presence of dating and relationship violence*

An important line of enquiry sparked by our consultations with young people related to a central difference between GBV and DRV; a difference that emerged as important in accounting for our pattern of findings. Specifically, young people observed that GBV is omnipresent in a way that DRV is not. This reflects findings from policy documents, but it also provides an important explanation for the differential value of structural components. For example, young people described experiences of feeling targeted by older men when in school uniforms; peers making comments about clothes and bodies; frequent experiences of gendered physical harassment but also gendered comments about appearance, including from teachers; and, importantly, homophobic violence and harassment, such as rumours about LGBTQ young people looking at people naked. Young people also reflected on how GBV is not strictly male to female, but also perpetrated, for example, between girls on the basis of appearance, or between boys on the basis of pubertal development. DRV was described as more hidden, though visible, for example, in inappropriate relationships between students and older adults, or in male students' inappropriate comments about their sexual histories with partners.

### *Salience and appropriateness of dating and relationship violence and gender-based violence education*

A common theme across all consultations was the manifest awkwardness of DRV and GBV teaching in schools, driven primarily by student behaviour (e.g. 'immature' comments) and teacher discomfort; in some cases, this manifested as a perceived lack of safety in classrooms. This chimed with our synthesis of process evaluations and the importance of appropriate implementation. Staff preparation to deliver this content was also flagged as a major weakness, with many young people reporting that staff were either unprepared or unwilling to support disclosures or acknowledge the trauma of DRV and GBV. An additional point young people raised was the remarkable 'straightness' of school lessons on DRV and GBV, commenting that existing programming erases LGBTQ young people in their descriptions of DRV and GBV.

### *The value of proactive intervention*

Finally, young people commented on the perception that interventions were tokenistic and responsive, for example to a spate of reported incidents of DRV or GBV. For example, all-school assemblies were noted as a common but disappointing forum for DRV and GBV teaching, noting that in many cases, teachers indicated their lack of interest or a view that this information was not important. All-school assemblies also did not appear to be responsive to the needs of survivors who might be in the audience. This chimes with our general finding that included studies did not appear to be aware of the specific needs of those with a prior history of DRV or GBV, either as victims or survivors or as perpetrators. In some cases, especially where schools appeared to be prepared to support students, young people noted that targeted, high-quality intervention could be an effective way to diffuse DRV and GBV information.

### *Reflections on consultation*

Because this was a systematic review in the public domain, ethical approval was not necessary in the main, except for access to additional data sources. In addition, our work with stakeholders was undertaken to

the highest standards of involvement and engagement; that is, we worked with them from a learning, as opposed to a data generation, perspective; and we ensured that engagement with young people, especially when these young people may have been vulnerable (e.g. engagement with a specialist young people's project), was overseen by AYPH's specialist youth workers separate from the research team.

Findings from our consultation was instrumental in shaping our interpretation of findings, as was documented above. In the future, we expect that engagement with young people and other stakeholders will inform ongoing research with a view towards valuing their perspectives in developing relevant interventions.

## Implications of the findings

### *In light of prior research*

The study findings indicate that well-implemented single-component interventions may be more effective than undertaking complex, multicomponent interventions and in some cases multilevel interventions. While perhaps initially counterintuitive to existing evidence relating to the value of complex interventions, it is in line with developments of the understanding of complexity in intervention research.<sup>279</sup> An intervention may be considered complex, for example, because of the number of components or socio-ecological levels at which the intervention is targeted.<sup>280</sup> However, increasingly it is recognised that complexity also arises through the interaction between the intervention (however 'complex') and the context in which the intervention is delivered.<sup>47,279,281</sup>

In line with Hawe *et al.* (2009),<sup>252</sup> and the adoption of a 'systems approach', the concept of interventions can be interpreted as 'events within systems' which aim to disrupt the functioning of complex systems through changing relationships, displacing entrenched practice and redistribute and transform resources. This approach is of value when interpreting the current study findings that implementing a simple intervention well is more effective than more complex multicomponent, or in some cases multilevel interventions. It helps contextualise these findings, whereby interventions interact with the context in which they are delivered, and that more simple interventions may be more likely well received and acceptable within school environments, receiving greater staff buy-in and delivered with the intended fidelity, especially in under-resourced schools with low staff capacity.

The benefit of employing a systems perspective when interpreting the current findings may lie in the extent to which the approach expands on understanding interventions that minimally disrupt the system, towards more disruptive system changes operating at multiple system levels.<sup>47</sup> The ability to fully disrupt the school system at multiple levels depends on the ability for the intervention to be implemented well, or at all, how receptive the system is to the intervention and whether appropriate intervention components are sufficiently activated (i.e. the system being disrupted in the first place). This is of significant importance when interpreting the current study findings, and especially those relating to GBV. The findings suggest that these require normative, and social structural change via more complex mechanisms which may not be sufficiently triggered by current interventions. In addition, activation may be sought by using interventions which are being delivered in environments which are less receptive and capable of implementing multi component or multilevel interventions. Interventions which require whole-system level change, for example those underpinned by the WHO Health Promoting Schools framework often show promise, but experience implementation shortcomings, often because the length of time required to implement these changes is often underestimated.<sup>282</sup>

Existing systematic reviews of interventions for young people have focused on DRV without considering intervention impacts jointly with GBV.<sup>5-8</sup> It is understood that interventions focusing on DRV may impact GBV, and those focusing on GBV may impact DRV, as both phenomena are potentially underpinned by common mechanisms and structural features that lead to high rates of both within schools and among school populations. It is hypothesised that DRV and GBV have a common basis in exploiting gender

and sexual inequalities, and in antisocial norms.<sup>283</sup> Indeed, QCA findings suggest that the optimum way to reduce both DRV and GBV victimisation is to reduce perpetration. However, the current findings also indicate that programmes were effective in reducing DRV perpetration and victimisation (in the long, but not short term), but were not effective in reducing GBV victimisation and perpetration. The mechanisms by which the interventions were effective did not relate to those described as complex, or relating to school organisation and social structure (i.e. higher levels in the socio-ecological models), but instead related to 'basic safety' and individual-level capabilities and motivations about unacceptability of violence. Therefore, the current study findings suggest that the basic safety mechanism might be more able to generate reductions in DRV than GBV, and so despite their theorised common aetiological mechanisms, the mechanisms through which the interventions are effective are different for DRV compared to GBV. This is a challenge to schools and interventionists alike, who may wish to develop and implement a common programme to address multiple linked outcomes.

The success of interventions in reducing DRV and to a lesser degree GBV via different mechanisms is of key importance in the theorisation of interventions addressing these forms of violence. When designing interventions and theorising their role in interrupting the system, it is important to address whether the interventions are designed to address the underlying mechanisms behind the behaviour (i.e. the root causes), or the manifestations of the behaviours themselves. The study findings have identified differences in what might be effective for DRV as compared to GBV. Consequently, by designing and theorising interventions intended to address the range of different, violence-specific mechanisms that need to be activated, along with incorporating an understanding the context and actors within the system,<sup>284</sup> is likely to contribute towards successfully implemented interventions and overall intervention effectiveness.

The current research findings indicate that best way to reduce both DRV and GBV victimisation is through reducing perpetration of these behaviours. It is noted that the reduction in perpetration is best achieved through components which target students' interpersonal relationships with each other. While this may relate to improving positive relationships and/or reducing negative relationships, we found that key components related to interpersonal skills and especially those that provided a space to practice these skills. In line with existing evidence on RSE, it is known that young people want and need comprehensive, positive, inclusive and skills-based learning to enhance their ability to develop healthy and positive relationships throughout their lives, and that they want this to be delivered through interactive, participatory, learner-centred approaches.<sup>285-291</sup> Of importance for discussion is the relevance of gender in relation to the current findings; while there was an absence of evidence for a moderating effect of gender studies reporting process evaluation outcomes, analyses of outcome evaluations showed that programmes may see larger effects in schools with more female students. While perhaps counterintuitive, especially given that overall the strongest evidence suggests that to reduce victimisation was found to be through perpetration, a reduction in perpetration may not be the only way to reduce victimisation. This finding, while of importance to the effectiveness of interventions, continues to perpetuate the inequalities whereby girls are disproportionately affected and expected to take responsibility for issues relating to, for example, sexual and reproductive health, relationships and sexual harassment. The sexual double standard and lack of international male engagement with a focus on addressing gender inequalities within relationship and sex education is significant gaps as identified by UNESCO and the WHO among others.<sup>285,292-295</sup> More work is needed in RSE and interventions addressing DRV and GBV to engage with boys and young men across the life course, including from primary school. This work should promote positive relationships for all, and especially challenge the gender inequalities that underlie not just the experiences of these behaviours, but also the interventions designed to target mechanisms/behaviours. More work is also needed to establish how best to engage young men to address these issues.

Along with substantive contributions to the literature, the research also provides some theoretical insights. The research incorporated the use of Markham and Averyard's (2003)<sup>228</sup> theory of human functioning and school organisation to promote student health to support the structuring of concepts

for the review. While supporting the understanding of the deeper structural influences on young people's behaviour,<sup>46</sup> the theory is in line with the WHO definition of health-promoting schools<sup>296</sup> and with 'ecological systems' approaches.<sup>252</sup> Each of these approaches emphasises the importance of the school environment, which can be described as a complex social system whereby their functioning is shaped by interactions among a diverse range of changing actors including staff, pupils and parents.<sup>229</sup> The interactions occur between settings within the school such as classes or parent-teacher meetings.<sup>252</sup> The schools themselves are nested within educational supra-systems and are situated within a broader social context where the students and staff interact with other settings such as family and community. While the focus of the interventions in this review are exclusively school-based, they must be considered in light of the wider ecological context through which schools operate. For example, the process of 'lowering boundaries', as outlined by Markham and Aveyard (2003)<sup>228</sup> with schools enhancing student commitment by promoting cultural congruence between schools and the wider community, may pose a particular problem for issues relating to DRV and GBV in some contexts. Weakening framing between school and community may not be desirable in contexts where school safety is challenged by the surrounding community, and where school is a 'safe haven' for students. Therefore, reconsideration of the theory and implementation syntheses from the perspective of local context is of particular salience, where intervention selection and implementation should be interpreted as increasing commitment to school, peers and prosocial behaviours, rather than weakening framing and boundaries. There is therefore real value in the application of the Markham and Aveyard (2003)<sup>228</sup> theory to DRV and GBV, but subject to ensuring the wider appreciation of schools and their broader social contexts and the roles these contexts have in the implementation of interventions.

### *In light of policy and practice*

The UK Government's 2021 Tackling Violence Against Women and Girls Strategy,<sup>297</sup> similar to the Welsh Government's Violence against Women and Girls, Domestic Abuse and Sexual Violence (Wales) Act,<sup>298</sup> demonstrates commitment to tackling and eradicating all forms of violence against women experiencing, and at risk of domestic and sexual abuse and all forms of GBV. In line with these aims, UK RSE has undergone significant reform in recent years, with RSE becoming a statutory subject in all English secondary schools, and RE statutory for primary schools in 2020. Similarly, in Wales significant reform of the whole educational system is underway and parents will no longer have the right to withdraw their children from RSE from September 2022. However, despite the recognition of the importance of the issues of DRV and GBV highlighted in this review that are encompassed within the curriculum, Government definitions for domestic violence and abuse continue to relate to only those aged 16 or older.<sup>299</sup> This, therefore, has implications for the recognition, reporting, recording and understanding of the extent of both DRV and GBV among young people, but importantly poses significant challenges not only for victims who want and need support, but also for perpetrators seeking help to change.

Governmental agencies responsible for inspections of schools in England conducted a review of sexual abuse in schools and colleges (Office for Standard in Education, Children's Services and Skills; Ofsted<sup>32</sup>) and a review of peer-on-peer sexual harassment among secondary school pupils in Wales.<sup>300</sup> These reports illuminated the gendered nature of the deficits in current RSE and school-based settings. In England, 60% of girls in schools experienced contact forms of sexual harassment compared to 25% of males. In relation to non-contact forms of sexual harassment, such as unwanted or inappropriate sexual comments, the report found that 90% of young women in schools experienced this compared to 75% of males. In Wales, many female pupils (86%) report personal experience of peer-on-peer harassment or seeing others experience this. This compares with a lower proportion of male pupils (74%). Twice as many girls (61%) report having personal experience of peer-on-peer sexual harassment than boys (29%). The gendered distribution of experience of sexual harassment and violence within schools is well characterised both in the UK and internationally.<sup>4</sup> An important addition to this from the findings of this study relates to the potential for intervention programmes tackling DRV and GBV victimisation and perpetration to equally engage with boys as well as girls, or at least not be gender blind in relation to addressing gender inequalities of these topics. The current research findings suggest programmes may be more effective when targeting populations with a majority of young women. While this contributes

to the already disproportionate burden young women experience in relation to DRV and GBV, it raises greater questions for practice, as for research, about how best to engage young men in programmes relating to DRV and GBV so as to promote equal accountability and responsibility.

Both reports recommend comprehensive RSE as a fundamental component to addressing harassment and violence in schools and each specifically recommends improved teacher training and professional learning in RSE and associated topics such as harassment and violence. As identified in the exploration of intervention implementation, the long-term nature of implementing and evaluating DRV and GBV interventions poses significant challenges for schools (and researchers). Given that school readiness and engagement play a crucial role in the implementation of interventions, complex or otherwise, greater support and capacity are needed for schools to engage in the prevention of DRV and GBV, whether through RSE curriculum, wider WSAs or through the participation of rigorous, well-designed research. The establishment of dedicated health and well-being roles, either for existing staff, or newly created health and well-being roles for staff who are specifically trained in areas of health and well-being, including those such as harassment, violence etc., would provide greater scope and capacity to engage in preventive work, and to implement this work successfully. This is especially important given our review's findings that staff-directed components were not valuable in explaining intervention effectiveness.

### **Where further research is needed**

The next stage of research is to utilise the current review findings to update knowledge about intervention functioning, consider the different mechanisms operating behind effective DRV and GBV interventions and incorporate this into the design and development of an intervention to be optimised, feasibility tested and piloted.<sup>279</sup>

To avoid the development of interventions that will not be implementable or effective in crowded and rapidly changing systems, interdisciplinary working with those involved in the education system (and thus with understanding of schools' systems and structures) will be vital.<sup>47</sup> This includes not just policy and practice professionals, but school senior leadership teams, teaching staff and students, as well as young people's wider networks (e.g. family, religious groups).

Future evaluations should consider providing evidence relating to the cost effectiveness of interventions. This is an important element which may play a key role in persuading up-take within the education sector, especially where staff engagement and capacity to implement interventions is low. As cost savings associated with interventions may also be borne by sectors outside of education (e.g. benefits to the economy through increased work productivity, and reduced costs for the legal system), high-quality evidence of the cost effectiveness of interventions for DRV and GBV may support funding for such interventions within school settings where budgets are frequently constrained.

The focus of this review was solely on school-based interventions, given that schools play a key role in health promotion,<sup>282</sup> especially in relation to sexual health, relationships and well-being, primarily via RSE. Schools, however, represent only one of the key bounded ecological systems which play a role in young people's lives,<sup>47</sup> via which interventions may take place to address DRV, GBV and other health behaviours. Schools are part of a broader network of agents who make up a variety of systems and structures whose interactions influence behavioural practices and health. Given the complex interplay between students, teachers and school systems and structures, and their wider engagement with the family and community, there is scope to expand the exploration of effective interventions and components in different settings, including primary schools, as equal uncertainty is present in these settings about 'what works', including for groups where health inequalities are present but interventions are poorly understood (e.g. sexual minority groups).

As highlighted by Oringanje *et al.* (2016),<sup>301</sup> and as evidenced in the literature, there is significant variation in the way some behavioural mediators and outcomes are measured and reported in RSE

intervention studies, and this is also seen among existing reviews of GBV.<sup>5,6,41</sup> In relation to outcome measures, this poses challenges for providing estimates of ICCs, and thus the required sample size for cluster RCTs in this area, as well as comparisons of effectiveness. The ever-changing nature of DRV and GBV also mean that existing measures may no longer be fit for purpose. The challenges of measurement are also demonstrated through mediated impacts on DRV. For example, in this research it is possible that interventions may have impacted on student school belonging and social norms, and that these mediated intervention impacts on DRV, but that measurements of school belonging and social norms were not effectively captured. More research is needed on the reliability and validity of both outcome measures of DRV and GBV, as well as potential important mediators.

## Strengths and limitations

Our analysis had several strengths. First, our search was uniquely exhaustive, drew on a range of supplementary methods and was not limited by language or type of publication (i.e. we included all relevant research published to June 2021 regardless of language of publication or format of report). As a result, we included a number of trials that previous meta-analyses have not included and drew on a rich body of evidence for our syntheses. Second, our consideration of a range of evidence sources and forms, in a correspondingly broad range of syntheses, meant that we were able to draw inferences and address nuances in the body of evidence that any one synthesis alone would have been unable to explore. Third, despite a range of prior reviews on interventions for DRV and GBV, this systematic review is the first to include a comprehensive analysis of theories of change, components or implementation; the first to relate a range of intervention components to intervention effectiveness; the first to jointly consider DRV and GBV as distinct but related constructs; the first to use QCA or NMA to explore intervention effectiveness; and the first to synthesise equity impacts and mediational evidence from interventions.

However, our analysis has several limitations. First, we cannot exclude the possibility that studies were missed, either due to the limits of database searching or due to publication bias. Several studies could not be retrieved despite our best efforts, though these were all likely implementation evidence, and indeed our analyses of publication bias suggested some risk of selective publication and reporting. In any systematic review of complex interventions, there are ambiguities in what is included and excluded. For example, we excluded interventions relating to the prevention of child sexual abuse or general social-emotional learning, even where the downstream prevention of DRV and GBV was researchers' implicit goal. Thus, there may be interventions evaluated in, for example primary school children that could be effective at reducing DRV and GBV, but that were not included in our review. Second, our analyses of theories of change, components and implementation evidence (and, to a lesser extent, our QCA) were inductive. While this is not strictly a limitation, we cannot rule out that a different set of investigators would have come to different conclusions. To mitigate against this, we drew on extensive auditing, second coding and reflection amongst the investigators. Third, our meta-analyses required extensive data transformation to prepare estimates for analysis, including because a number of trials did not properly account for clustering. While our approaches (such as imputing ICC = 0.05) are on balance likely to be conservative, it is possible that a different set of meta-analysts may have organised outcome constructs or data transformations differently. Finally, we did not include every form of GBV defined by legal standards in the UK; for example, we did not include honour-based violence, forced marriage or female genital mutilation. This was because these forms of violence are not peer-to-peer or perpetrated in school settings, even though schools should be alert to these risks.

## Equity, diversity and inclusion

Throughout this work, we have sought to centre equity and diversity. Our investigator team drew on a range of identities and perspectives, including by gender, sexual orientation, age, ethnicity, and

professional preparation and discipline. In addition, through our consultations with stakeholders, we sought diverse groups of young people and of policy stakeholders to inform our work.

Because DRV and GBV are fundamentally gendered threats to public health, health inequities were central to our analysis. This was both in the fact of the types of violence we are seeking to prevent and in every aspect of our analysis plan. For example, we used an explicitly gendered lens in understanding intervention implementation and accounted for the importance of gender-relevant programming in our QCA. In addition, our analysis of equity, which is the first of its kind in this area, drew to light how intervention effectiveness may be conditional on gender.

## Conclusion

Dating and relationship violence and GBV remain threats to population health and the health of adolescents and young people. School-based interventions show promise, though evidence is uneven and is more conclusive for DRV than for GBV. It is possible, if not likely, that differential mechanisms will need to be activated for DRV than for GBV, and existing theories of change do not adequately account for differences between DRV and GBV in terms of intervention functioning. Schools may need to focus on interventions that can be implemented with intention and to a high standard rather than challenging, multilevel strategies, and should carefully consider intervention fit.





# Acknowledgements

Lindsay Starbuck (Association for Young People's Health) delivered the consultations with young people. Helen Coelho and Laura Trigg (University of Exeter) provided early help with screening and quality assessment.

We acknowledge the members of our policy and practice advisory group for their feedback and guidance throughout the project. The members were:

- Kristian Tomblin (Devon County Council)
- Jonathan Baggaley (PSHE Association)
- Lucy Emmerson (Sex Education Forum)
- Joan Roberts (School Health Research Network)
- Paul Storey (MESMAC)
- Catherine Bewley, Lisa Harrington and Pheeb Jameson (Galop)
- Nicola Douglas and Miranda Pio (Standing Together Against Domestic Abuse)
- Chrissy Stower (Splitz).

Finally, we acknowledge with sincere gratitude the young people who contributed to consultations in a range of fora.

Ethics approval was granted by the University of Exeter College of Medicine and Health Research Ethics Committee (ID 488499) on 18 August 2021 for analysis of additional data sets.

## Contributions of authors

**GJ Melendez-Torres** (<https://orcid.org/0000-0002-9823-4790>) (Professor of Clinical and Social Epidemiology) is the Principal Investigator and directed the review, overseeing all stages and components. He contributed to screening, data extraction, quality assessment and synthesis of intervention theories and components, implementation and economic evidence and the realist integration of findings. He led pairwise meta-analyses, synthesis of mediation and moderation evidence, and metaregressions, qualitative comparative analyses and network meta-analyses.

**Noreen Orr** (<https://orcid.org/0000-0002-5822-0247>) (Research Fellow) contributed to screening, data extraction and quality assessment, led the synthesis of intervention theories, and contributed to syntheses of intervention components and implementation evidence and the realist integration of findings.

**Caroline Farmer** (<https://orcid.org/0000-0003-3553-8970>) (Research Fellow) contributed to screening, and led data extraction, quality assessment and narrative synthesis of intervention effectiveness. She supervised synthesis of economic evidence and contributed to meta-analyses, network meta-analyses and the realist integration of findings.

**Naomi Shaw** (<https://orcid.org/0000-0001-7387-1809>) (Information Specialist) designed and led searches, led screening and contributed to all syntheses.

## ACKNOWLEDGEMENTS

**Annah Chollet** (<https://orcid.org/0000-0003-0329-2750>) (Postgraduate student) led synthesis of implementation evidence and contributed to synthesis of intervention theories and the realist integration of findings.

**Andrew J Rizzo** (<https://orcid.org/0000-0001-8578-9562>) (Postdoctoral Fellow) led synthesis of intervention components and contributed to synthesis of intervention theories, metaregressions and network meta-analyses, and the realist integration of findings.

**Fraizer Kiff** (<https://orcid.org/0000-0002-6008-3907>) (Graduate Research Assistant) contributed to data extraction and quality assessment and led synthesis of economic evidence.

**Emma Rigby** (<https://orcid.org/0000-0003-1781-9859>) (Chief Executive) co-led the consultation strategy and contributed to the realist integration of findings.

**Ann Hagell** (<https://orcid.org/0000-0002-3690-1674>) (Research Lead) co-led the consultation strategy and contributed to the realist integration of findings.

**Sidnei R Priolo Filho** (<https://orcid.org/0000-0003-1320-9674>) (Associate Professor of Psychology) contributed to the synthesis of implementation evidence and the realist integration of findings.

**Bruce Taylor** (<https://orcid.org/0000-0002-8115-1438>) (Senior Fellow) contributed to the synthesis of moderation evidence and the realist integration of findings.

**Honor Young** (<https://orcid.org/0000-0003-0664-4002>) (Senior Lecturer in Social Sciences and Health) contributed to all syntheses.

**Chris Bonell** (<https://orcid.org/0000-0002-6253-6498>) (Professor of Public Health Sociology) contributed to all syntheses and led the realist integration of findings.

**Vashti Berry** (<https://orcid.org/0000-0001-6438-3731>) (Associate Professor in Prevention Science) co-led the consultation strategy and contributed to all syntheses.

All authors contributed to the review methods and all contributed to, read and approved the final manuscript.

## Data-sharing statement

All data underpinning this report are available on request from the corresponding author.

## Ethics approval

The research involved no human participants and drew solely on evidence in the public realm, so ethics approval was not required.

## Department of Health and Social Care disclaimer

This publication presents independent research commissioned by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, MRC, CCF, NETSCC, the Public Health Research programme or the Department of Health and Social Care. The funder had no role in study design and had no role in the collection, management, analysis or interpretation of data. The funder had no role in writing manuscripts based on study findings or in the decision to submit findings for publication.

## References

1. Jewkes R, Flood M, Lang J. From work with men and boys to changes of social norms and reduction of inequities in gender relations: a conceptual shift in prevention of violence against women and girls. *Lancet* 2015;**385**:1580–9.
2. Exner-Cortens D, Eckenrode J, Rothman E. Longitudinal associations between teen dating violence victimization and adverse health outcomes. *Pediatrics* 2013;**131**:71–8.
3. Exner-Cortens D, Eckenrode J, Bunge J, Rothman E. Revictimization after adolescent dating violence in a matched, national sample of youth. *J Adolesc Health* 2017;**60**:176–83.
4. Taquette SR, Monteiro DLM. Causes and consequences of adolescent dating violence: a systematic review. *J Inj Violence Res* 2019;**11**:137–47.
5. De La Rue L, Polanin JR, Espelage DL, Pigott TD. A meta-analysis of school-based interventions aimed to prevent or reduce violence in teen dating relationships. *Rev Educ Res* 2017;**87**:7–34.
6. Kettrey HH, Marx RA, Tanner-Smith EE. Effects of bystander programs on the prevention of sexual assault among adolescents and college students. *Campbell Syst Rev* 2019;**15**:e1013.
7. Stanley N, Ellis J, Farrelly N, Hollinghurst S, Downe S. Preventing domestic abuse for children and young people: a review of school-based interventions. *Child Youth Serv Rev* 2015;**59**:120–31.
8. Fellmeth GLT, Heffernan C, Nurse J, Habibula S, Sethi D. Educational and skills-based interventions for preventing relationship and dating violence in adolescents and young adults. *Cochrane Database Syst Rev* 2013;**2013**:CD004534.
9. Earnest AA, Brady SS. Dating violence victimization among high school students in Minnesota: associations with family violence, unsafe schools, and resources for support. *J Interpers Violence* 2014;**31**:383–406.
10. Barter C, Stanley N. Inter-personal violence and abuse in adolescent intimate relationships: mental health impact and implications for practice. *Int Rev Psychiatry* 2016;**28**:485–503.
11. Young H, Long SJ, Melendez-Torres GJ, Kim HS, Hewitt G, Murphy S, *et al*. Dating and relationship violence victimization and perpetration among 11–16 year olds in Wales: a cross-sectional analysis of the School Health Research Network (SHRN) survey. *J Public Health* 2021;**43**:111–22.
12. Pulerwitz J, Blum R, Cislighi B, Costenbader E, Harper C, Heise L, *et al*. Proposing a conceptual framework to address social norms that influence adolescent sexual and reproductive health. *J Adolesc Health* 2019;**64**:S7–9.
13. Reyes HLM, Foshee VA, Niolon PH, Reidy DE, Hall JE. Gender role attitudes and male adolescent dating violence perpetration: normative beliefs as moderators. *J Youth Adolesc* 2016;**45**:350–60.
14. Théorêt V, Hébert M, Fernet M, Blais M. Gender-specific patterns of teen dating violence in heterosexual relationships and their associations with attachment insecurities and emotion dysregulation. *J Youth Adolesc* 2021;**50**:246–59.
15. Hébert M, Blais M, Lavoie F. Prevalence of teen dating victimization among a representative sample of high school students in Quebec. *Int J Clin Health Psychol* 2017;**17**:225–33.
16. Marcantonio TL, Weese J, Willis M. Rates of forced sexual experiences among high school students from 2001 to 2019. *J Interpers Violence* 2022;**37**:NP21045–69.

17. Macdowall W, Gibson LJ, Tanton C, Mercer CH, Lewis R, Clifton S, *et al.* Lifetime prevalence, associated factors, and circumstances of non-volitional sex in women and men in Britain: findings from the third National Survey of Sexual Attitudes and Lifestyles (Natsal-3). *Lancet* 2013;**382**:1845–55.
18. Shorey RC, Cohen JR, Lu Y, Fite PJ, Stuart GL, Temple JR. Age of onset for physical and sexual teen dating violence perpetration: a longitudinal investigation. *Prev Med* 2017;**105**:275–9.
19. Foshee VA, Gottfredson NC, Reyes HLM, Chen MS, David-Ferdon C, Latzman NE, *et al.* Developmental outcomes of using physical violence against dates and peers. *J Adolesc Health* 2016;**58**:665–71.
20. Miller E, Culyba AJ, Paglisotti T, Massof M, Gao Q, Ports KA, *et al.* Male adolescents' gender attitudes and violence: implications for youth violence prevention. *Am J Prev Med* 2020;**58**:396–406.
21. Girlguiding. *Research Briefing: It Happens all the Time. Girls' and Young Women's Experiences of Sexual Harassment*. 2021. URL: [www.girlguiding.org.uk/globalassets/docs-and-resources/research-and-campaigns/girlguiding-research-briefing\\_girls-experiences-of-sexual-harassment\\_june2021.pdf](http://www.girlguiding.org.uk/globalassets/docs-and-resources/research-and-campaigns/girlguiding-research-briefing_girls-experiences-of-sexual-harassment_june2021.pdf) (accessed April 2023).
22. Mueller AS, James W, Abrutyn S, Levin ML. Suicide ideation and bullying among US adolescents: examining the intersections of sexual orientation, gender, and race/ethnicity. *Am J Public Health* 2015;**105**:980–5.
23. Johns MM, Lowry R, Rasherry CN, Dunville R, Robin L, Pampati S, *et al.* Violence victimization, substance use, and suicide risk among sexual minority high school students – United States, 2015–2017. *MMWR Morb Mortal Wkly Rep* 2018;**67**:1211–5.
24. Martin-Storey A, Pollitt AM, Baams L. Profiles and predictors of dating violence among sexual and gender minority adolescents. *J Adolesc Health* 2021;**68**:1155–61.
25. Norris AL, Orchowski LM. Peer victimization of sexual minority and transgender youth: a cross-sectional study of high school students. *Psychol Violence* 2020;**10**:201–11.
26. Shorey RC, Fite PJ, Choi H, Cohen JR, Stuart GL, Temple JR. Dating violence and substance use as longitudinal predictors of adolescents' risky sexual behavior. *Prev Sci* 2015;**16**:853–61.
27. Vivolo-Kantor AM, Olsen EOM, Bacon S. Associations of teen dating violence victimization with school violence and bullying among US high school students. *J Sch Health* 2016;**86**:620–7.
28. Loxton D, Dolja-Gore X, Anderson AE, Townsend N. Intimate partner violence adversely impacts health over 16 years and across generations: a longitudinal cohort study. *PLOS ONE* 2017;**12**:e0178138.
29. Costa BM, Kaestle CE, Walker A, Curtis A, Day A, Toumbourou JW, *et al.* Longitudinal predictors of domestic violence perpetration and victimization: a systematic review. *Aggress Violent Behav* 2015;**24**:261–72.
30. Reidy DE, Kearns MC, Houry D, Valle LA, Holland KM, Marshall KJ. Dating violence and injury among youth exposed to violence. *Pediatrics* 2016;**137**:e20152627.
31. Espelage DL, Merrin GJ, Hatchel T. Peer victimization and dating violence among LGBTQ youth: the impact of school violence and crime on mental health outcomes. *Youth Violence Juv Justice* 2016;**16**:156–73.
32. Ofsted. *Review of Sexual Abuse in Schools and Colleges*. 2021. URL: [www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges](http://www.gov.uk/government/publications/review-of-sexual-abuse-in-schools-and-colleges/review-of-sexual-abuse-in-schools-and-colleges) (accessed April 2023).

33. Herbert A, Heron J, Barter C, Szilassy E, Barnes M, Howe L, *et al.* Risk factors for intimate partner violence and abuse among adolescents and young adults: findings from a UK population-based cohort [version 3; peer review: 2 approved]. *Wellcome Open Res* 2021;**5**:176.
34. Young H, Turney C, White J, Lewis R, Bonell C. Formative mixed-method multicase study research to inform the development of a safer sex and healthy relationships intervention in further education (FE) settings: the SaFE Project. *BMJ Open* 2019;**9**:e024692.
35. Meiksin R, Crichton J, Dodd M, Morgan GS, Williams P, Willmott M, *et al.* A school intervention for 13- to 15-year-olds to prevent dating and relationship violence: the Project Respect pilot cluster RCT. *Public Health Res* 2020;**8**:1–337.
36. Foshee VA, Bauman KE, Arriaga XB, Helms RW, Koch GG, Linder GF. An evaluation of Safe Dates, an adolescent dating violence prevention program. *Am J Public Health* 1998;**88**:45–50.
37. Espelage DL, Low S, Polanin JR, Brown EC. Clinical trial of Second Step© middle-school program: impact on aggression & victimization. *J Appl Dev Psychol* 2015;**37**:52–63.
38. Sargent KS, Jouriles EN, Rosenfield D, McDonald R. A high school-based evaluation of TakeCARE, a video bystander program to prevent adolescent relationship violence. *J Youth Adolesc* 2017;**46**:633–43.
39. Bonell C, Jamal F, Harden A, Wells H, Parry W, Fletcher A. Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. *Public Health Res* 2013;**1**.
40. Taylor BG, Stein ND, Mumford EA, Woods D. Shifting Boundaries: an experimental evaluation of a dating violence prevention program in middle schools. *Prev Sci* 2013;**14**:64–76.
41. De Koker P, Mathews C, Zuch M, Bastien S, Mason-Jones AJ. A systematic review of interventions for preventing adolescent intimate partner violence. *J Adolesc Health* 2014;**54**:3–13.
42. Piolanti A, Foran HM. Efficacy of interventions to prevent physical and sexual dating violence among adolescents: a systematic review and meta-analysis. *JAMA Pediatr* 2022;**176**:142–9.
43. Lundgren R, Amin A. Addressing intimate partner violence and sexual violence among adolescents: emerging evidence of effectiveness. *J Adolesc Health* 2015;**56**:S42–50.
44. Tancred T, Melendez-Torres GJ, Papparini S, Fletcher A, Stansfield C, Thomas J, *et al.* Interventions integrating health and academic education in schools to prevent substance misuse and violence: a systematic review. *Public Health Res* 2019;**7**:17.
45. Melendez-Torres GJ, Tancred T, Fletcher A, Thomas J, Campbell R, Bonell C. Does integrated academic and health education prevent substance use? Systematic review and meta-analyses. *Child Care Health Dev* 2018;**44**:516–30.
46. Moore GF, Evans RE. What theory, for whom and in which context? Reflections on the application of theory in the development and evaluation of complex population health interventions. *SSM Popul Health* 2017;**3**:132–5.
47. Moore GF, Evans RE, Hawkins J, Littlecott H, Melendez-Torres GJ, Bonell C, *et al.* From complex social interventions to interventions in complex social systems: future directions and unresolved questions for intervention development and evaluation. *Evaluation* 2019;**25**:23–45.
48. Moore GF, Littlecott HJ, Turley R, Waters E, Murphy S. Socioeconomic gradients in the effects of universal school-based health behaviour interventions: a systematic review of intervention studies. *BMC Public Health* 2015;**15**:907.
49. Sanchez-Jimenez V, Munoz-Fernandez N, Ortega-Rivera J. Efficacy evaluation of ‘Dat-e Adolescence’: a dating violence prevention program in Spain. *PLOS ONE* 2018;**13**:e0205802.

50. Department for Education. *Relationships Education, RSE and Health Education: School Practice in Early Adopter Schools*. 2021. URL: [www.gov.uk/government/publications/rshe-school-practice-in-early-adopter-schools](http://www.gov.uk/government/publications/rshe-school-practice-in-early-adopter-schools) (accessed April 2023).
51. Department for Education. *Relationships Education, Relationships and Sex Education (RSE) and Health Education: Statutory Guidance for Governing Bodies, Proprietors, Head Teachers, Principals, Senior Leadership Teams, Teachers*. London, UK: Crown Stationery Office; 2019. URL: [www.gov.uk/government/publications/relationships-education-relationships-and-sex-education-rse-and-health-education](http://www.gov.uk/government/publications/relationships-education-relationships-and-sex-education-rse-and-health-education) (accessed April 2023).
52. CRD. *Systematic Reviews. CRD's Guidance for Undertaking Reviews in Health Care*. York, UK: Centre for Reviews and Dissemination, University of York; 2009. URL: [www.york.ac.uk/media/crd/Systematic\\_Reviews.pdf](http://www.york.ac.uk/media/crd/Systematic_Reviews.pdf) (accessed April 2023).
53. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.3 (updated February 2022). Cochrane; 2022. URL: [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook) (accessed April 2023).
54. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71.
55. Welch V, Petticrew M, Tugwell P, Moher D, O'Neill J, Waters E, White H. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. *PLOS Med* 2012;**9**:e1001333.
56. Bonell C, Allen E, Warren E, McGowan J, Bevilacqua L, Jamal F, et al. Effects of the Learning Together intervention on bullying and aggression in English secondary schools (INCLUSIVE): a cluster randomised controlled trial. *Lancet* 2018;**392**:2452–64.
57. Booth A, Harris J, Croot E, Springett J, Campbell F, Wilkins E. Towards a methodology for cluster searching to provide conceptual and contextual 'richness' for systematic reviews of complex interventions: case study (CLUSTER). *BMC Med Res Methodol* 2013;**13**:118.
58. Meiksin R, Melendez-Torres GJ, Miners A, Falconer J, Witzel TC, Weatherburn P, et al. E-health interventions targeting STIs, sexual risk, substance use and mental health among men who have sex with men: four systematic reviews. *Public Health Res* 2022;**10**:4.
59. Jamal F, Fletcher A, Harden A, Wells H, Thomas J, Bonell C. The school environment and student health: a systematic review and meta-ethnography of qualitative research. *BMC Public Health* 2013;**13**:798.
60. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ* 2019;**366**:l4898.
61. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford: Oxford University Press; 2015.
62. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment. *Pharmacoeconomics* 2006;**24**:355–71.
63. Sutcliffe K, Thomas J, Stokes G, Hinds K, Bangpan M. Intervention Component Analysis (ICA): a pragmatic approach for identifying the critical features of complex interventions. *Syst Rev* 2015;**4**:140.
64. Bonell C, Wells H, Harden A, Jamal F, Fletcher A, Thomas J. Theories of how the school environment impacts on student health: systematic review. *Health Place* 2013;**24**:242–9.
65. Noblit GW, Hare RD. *Meta-ethnography: Synthesizing Qualitative Studies*. Thousand Oaks, CA: SAGE Publications Ltd; 1988.

66. Malpass A, Carel H, Ridd M, Shaw A, Kessler D, Sharp D, *et al.* Transforming the perceptual situation: a meta-ethnography of qualitative work reporting patients' experiences of mindfulness-based approaches. *Mindfulness* 2012;**3**:60–75.
67. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;**8**:45.
68. Tanner-Smith EE, Tipton E. Robust variance estimation with dependent effect sizes: practical considerations including a software tutorial in Stata and SPSS. *Res Synth Methods* 2013;**5**:13–30.
69. Bonell C, Dickson K, Hinds K, Melendez-Torres GJ, Stansfield C, Fletcher A, *et al.* *The Effects of Positive Youth Development Interventions on Substance Use, Violence and Inequalities: Systematic Review of Theories of Change, Processes and Outcomes.* *Public Health Res* 2016;**4**. <https://doi.org/10.3310/phr04050>
70. O'Neill J, Tabish H, Welch V, Petticrew M, Pottie K, Clarke M, *et al.* Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol* 2014;**67**:56–64.
71. Skeen S, Laurenzi CA, Gordon SL, du Toit S, Tomlinson M, Dua T, *et al.* Adolescent mental health program components and behavior risk reduction: a meta-analysis. *Pediatrics* 2019;**144**:e20183488.
72. Melendez-Torres GJ, Sutcliffe K, Burchett HED, Rees R, Thomas J. Developing and testing intervention theory by incorporating a views synthesis into a qualitative comparative analysis of intervention effectiveness. *Res Synth Methods* 2019;**10**:389–97.
73. Rücker G, Petropoulou M, Schwarzer G. Network meta-analysis of multicomponent interventions. *Biom J* 2020;**62**:808–21.
74. Melendez-Torres GJ, Bonell C, Thomas J. Emergent approaches to the meta-analysis of multiple heterogeneous complex interventions. *BMC Med Res Methodol* 2015;**15**:47.
75. Pawson R, Tilley N. *Realistic Evaluation.* London, UK: SAGE Publications Ltd; 1997.
76. Carroll C, Booth A, Leaviss J, Rick J. 'Best fit' framework synthesis: refining the method. *BMC Med Res Methodol* 2013;**13**:37.
77. Miller E, Tancredi DJ, McCauley HL, Decker MR, Virata MC, Anderson HA, *et al.* 'Coaching Boys into Men': a cluster-randomized controlled trial of a dating violence prevention program. *J Adolesc Health* 2012;**51**:431–8.
78. Achyut P, Bhatla N, Khandekar S, Maitra S, Verma RK. *Building Support for Gender Equality Among Young Adolescents in School: Findings from Mumbai, India.* New Delhi, India: International Center for Research on Women; 2011.
79. Arnab S, Brown K, Clarke S, Judd B, Baxter J, King R, *et al.*; *European Conference on Games Based Learning. Serious game for Relationships and Sex Education (RSE): promoting discourse on pressure and coercion in adolescent relationships.* Proceedings of the European Conference on Games Based Learning, 2012, vol. 1. pp. 611–4, 4–5 October 2012, Cork, Ireland.
80. Cascardi M, Avery-Leaf S. Case study of a school-based universal dating violence prevention program. *SAGE Open* 2014;**4**:1–9.
81. Cissner AB, Ayoub LH. *Building Healthy Teen Relationships: An Evaluation of the Fourth R Curriculum with Middle School Students in the Bronx.* U.S. Washington, DC: Center for Court Innovation; 2014.
82. Coker AL, Bush HM, Cook-Craig PG, DeGue SA, Clear ER, Brancato CJ, *et al.* RCT testing bystander effectiveness to reduce violence. *Am J Prev Med* 2017;**52**:566–78.

83. Durand AJ. *Sexual Harassment and Junior High School Students: The Effect of a Program Designed to Address Peer-to-Peer Sexual Harassment*. PhD thesis. Fort Collins, CO: Colorado State University; 1997.
84. Filho SRP. *Avaliação de uma intervenção para prevenção da violência no namoro*. PhD thesis. São Carlos, Brazil: Universidade Federal de São Carlos; 2017.
85. International Center for Research on Women (ICRW). *Changing Course. Implementation and Evaluation of the Gender Equity Movement in Schools (GEMS) Program in Specific Sites - Vietnam, India and Bangladesh*. New Delhi, India: International Center for Research on Women; 2017.
86. Jaycox LH, McCaffrey DF, Weidmer Ocampo B, Marshall GN, Collins RL, Hickman LJ, et al. *Curbing Teen Dating Violence: Evidence from a School Prevention Program*. RB-9194-CDC. Santa Monica, CA: RAND Corporation; 2006.
87. Lee PS, Chen HT, Hong ZR. Effects of a bystander intervention curriculum for preventing sexual harassment. *J Res Edu Sci* 2018;**63**:1–35.
88. Lynch A. *Twilight, True Love and You: A Bibliotherapy Approach to Preventing Dating Abuse in Adolescent Girls*. PhD thesis. Guildford, UK: University of Surrey; 2014.
89. Macgowan MJ. An evaluation of a dating violence prevention program for middle school students. *Violence Vict* 1997;**12**:223–35.
90. Mathews C, Eggers SM, Townsend L, Aaro LE, de Vries PJ, Mason-Jones AJ, et al. Effects of PREPARE, a multi-component, school-based HIV and intimate partner violence (IPV) prevention programme on adolescent sexual risk behaviour and IPV: cluster randomised controlled trial. *AIDS Behav* 2016;**20**:1821–40.
91. Meiksin R, Campbell R, Crichton J, Morgan GS, Williams P, Willmott M, et al. Implementing a whole-school relationships and sex education intervention to prevent dating and relationship violence: evidence from a pilot trial in English secondary schools. *Sex Educ* 2020;**20**:658–74.
92. Merrell R. *The Impact of a Drama Intervention Program on the Response of the Bystander to Bullying Situations*. Doctoral thesis. Rochester, NY: University of Rochester; 2004.
93. Miller E, Chopel A, Jones KA, Dick RN, McCauley HL, Jetton J, et al. *Integrating Prevention and Intervention: A School Health Center Program to Promote Healthy Relationships*. Washington, DC: National Institute of Justice; 2015.
94. Muck C, Schiller EM, Zimmermann M, Kartner J. Preventing sexual violence in adolescence: comparison of a scientist-practitioner program and a practitioner program using a cluster-randomized design. *J Interpers Violence* 2018;**36**:NP1913–1940NP.
95. Roberts KEC. *An Evaluation of the Expect Respect: Preventing Teen Dating Violence High School Program*. PhD thesis. Athens, OH: Ohio University; 2009.
96. Rowe LS, Jouriles EN, McDonald R. Reducing sexual victimization among adolescent girls: a randomized controlled pilot trial of My Voice, My Choice. *Behav Ther* 2015;**46**:315–27.
97. Scull T, Malik C, Morrison A, Keefe E. Promoting sexual health in high school: a feasibility study of a web-based media literacy education program. *J Health Commun* 2021;**26**:147–60.
98. Scull TM, Kupersmidt JB, Malik CV, Morgan-Lopez AA. Using media literacy education for adolescent sexual health promotion in middle school: randomized control trial of Media Aware. *J Health Commun* 2018;**23**:1051–63.
99. Silverman AB. *Evaluating the Efficacy of the Rhode Island Teen Dating Violence Prevention Program (TDVPP): A Process and Outcome Approach to Determining the Success of Both Primary and Secondary Prevention Projects*. PhD thesis. South Kingston, RI: University of Rhode Island; 2000.



100. Taylor B, Stein ND, Woods D, Mumford E. *Shifting Boundaries: Final Report on an Experimental Evaluation of Youth Dating Violence Prevention Program in New York City Middle Schools*. Report submitted to the National Institute of Justice. Washington, DC: Police Executive Research Forum; 2011.
101. Yoder J, Williford A, Ortega L, Espelage DL, LoMurray S, Ruiz D, Kennedy N. Qualitative process evaluation of rural schools: uptake of change processes and contextual factors influencing implementation within a primary prevention program for youth. *Prev Sci* 2020;**21**:1093–103.
102. Farrelly N. *Violence Prevention for Children: An Evaluation of a Primary School Based Programme*. PhD thesis. Preston, UK: University of Central Lancashire; 2020.
103. Keddie A, Ollis D. Context matters: the take up of Respectful Relationships Education in two primary schools. *Aust Educ Res* 2020;**48**:15.
104. de Albuquerque KK. Gender dialogues in education: considerations about The Project Law Maria Da Penha goes to schools. *Estudos Feministas* 2020;**28**:e60485.
105. Hale B, Fox C, Gadd D. *Research report. Evaluation of three European schools-based domestic violence prevention education programmes. Relationships without Fear (England), La Máscara del Amor (Spain), Filles et Garçons, en route pour l'Égalité*. Manchester, UK: READAPT Project; 2012.
106. Henderson S, Reid-Howie Associates Ltd. *Evaluation of the Zero Tolerance 'Respect' Pilot Project*. Edinburgh: Scottish Executive Central Research Unit; 2002.
107. UNICEF. *Lessons from Serbia. Addressing Gender-Based Violence Through the School without Violence Programme*. Belgrade, Serbia: UNICEF Serbia Country Office; 2016.
108. Weingarten C, Rabago J, Reynolds J, Gates K, Yanagida E, Baker C. Examining the utility of a train-the-trainer model for dissemination of sexual violence prevention in schools. *Child Abuse Negl* 2018;**80**:70–9.
109. AVA, Comic Relief. *Prevention on the Schools' Agenda. Summary Report of Key Learning. Working with Secondary Schools in England to Prevent Violence Against Women and Girls*. London, UK: Comic Relief; 2013.
110. Bell J, Stanley N. Learning about domestic violence: young people's responses to a Healthy Relationships programme. *Sex Educ* 2006;**6**:237–50.
111. Bosch Jorba C. *Prevenió de les relacions abusives en la parella avaluació de l'eficàcia d'un programa educatiu per a adolescents*. PhD thesis. Barcelona, Spain: Universitat Autònoma de Barcelona; 2012.
112. Bruno L, Joelsson T, Franzen AG, Gottzen L. Heroes and others: tensions and challenges in implementing Mentors in Violence Prevention in Swedish schools. *J Gend Based Viol* 2020;**4**:141–55.
113. Casas Tello MM. *La prevención de la violencia en la pareja entre adolescentes a través del taller: la máscara del amor*. PhD thesis. Valencia, Spain: Universidad de Valencia; 2013.
114. CRG Research Ltd. *An independent evaluation of the TRUST Education Project*. Final report. Contract no. CON03/0809. 2016.
115. DMSS Research & Consultancy. *Tender's Healthy Relationship Education in Schools Funded by Comic Relief*. Final evaluation. Summary report. North Dalton, UK: DMSS Research & Consultancy; 2012.
116. Genovés VG, Tello MC. The prevention of dating violence in youth: the «Mask of Love» workshop. *Rev Educ* 2009;**349**:335–60.

117. Jalušič V. Spolno nadlegovanje v šolah: lekcije, potencial in omejitve pristopa » aktivnih opazovalcev «. *Šolsko Polje* 2019;**30**:89–118, 83–84, 91.
118. Kempes MM, van Pelt L, Beerthuisen MGJ, Boom J, Brugman D. *Programma-integriteit en effecten van Stay in Love+. Een preventieprogramma voor 12-15 jarige VMBO scholieren dat partnergeweld beoogte voorkomen*. The Hague, The Netherlands: Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC), Ministerie van Justitie; 2010.
119. Mateos Inchaurreondo A, Urrea Monclus A, Juan LO, Cojocarú D. A pilot study on the implementation of the GENER@T Program. *Rev Cercet Interv Soc* 2020;**68**:32–52.
120. Matos M, Machado C, Caridade S, Silva MJ. Prevention of the dating violence: intervention with youngsters in school context. *Psicologia: Teoria e Prática* 2006;**8**:55–75.
121. Maxwell C, Chase E, Warwick I, Aggleton P, Wharf H. *Freedom to Achieve: Preventing Violence, Promoting Equality: A Whole School Approach*. London, UK: Womankind Worldwide; 2010.
122. McElwee JD, Fox CL. Young people's perceptions of the 'Love Hurts' programme: is theatre an effective means of addressing teenage relationship abuse? *Br Educ Res J* 2020;**46**:1026–43.
123. McGinn M. *Care: A Communication and Relationship Education Primary Prevention Social Cognitive Model Aimed at Developing the Self-efficacy of Adolescents to Verbally Respond to Sexual Coercion*. PhD thesis. Dublin: University of Dublin, Trinity College; 2017.
124. Namy S, Heilman B, Stich S, Crownover J, Leka B, Edmeades J. Changing what it means to 'become a man': participants' reflections on a school-based programme to redefine masculinity in the Balkans. *Cult Health Sex* 2015;**17**:206–22.
125. Perez-Marco A, Soares P, Davo-Blanes MC, Vives-Cases C. Identifying types of dating violence and protective factors among adolescents in Spain: a qualitative analysis of Lights4Violence materials. *Int J Environ Res Public Health* 2020;**17**:2443.
126. Racionero-Plaza S, Ugalde L, Merodio G, Gutierrez-Fernandez N. 'Architects of their own brain'. Social impact of an intervention study for the prevention of gender-based violence in adolescence. *Front Psychol* 2020;**10**:3070.
127. Rogers M, Rumley T, Lovatt G. The change up project: using social norming theory with young people to address domestic abuse and promote healthy relationships. *J Fam Violence* 2018;**34**:507–19.
128. Sanders-McDonagh E, Rogers SA, Horvath M, Selwood S. *Evaluation Report of the Tender Healthy Relationships Project*. National Partnership Programme. London, UK: Tender; Middlesex University; 2015.
129. Sorbring E, Bolin A, Ryding J. A game-based intervention – a technical tool for social workers to combat adolescent dating-violence. *Adv Soc Work* 2015;**16**:125–39.
130. Walton K. *Domestic Violence Prevention in London Schools. Peer Educator Pilot Project Evaluation Report*. London, UK: Domestic Violence Responses; 2007.
131. Williams DJ, Neville FG. Qualitative evaluation of the Mentors in Violence Prevention pilot in Scottish high schools. *Psychol Violence* 2017;**7**:213–23.
132. Cahill K, Romei K. *Preventing Gender-based Violence in Schools in East and Southern Africa: From Consultation to Capacity Building*. Melbourne: Youth Research Centre, University of Melbourne; 2019.
133. Chipeta E. *Paving the Way for Change: How Protect Our Youth Clubs are Helping Students Challenge Norms and Confront School-Related GBV. A Case Study from Malawi*. Boston, MA: Bantwana Initiative of World Education, Inc.; 2019. URL: [www.ungei.org/publication/paving-way-change](http://www.ungei.org/publication/paving-way-change) (accessed April 2023).

134. Cockcroft A, Marokoane N, Kgakole L, Mhati P, Tswetla N, Sebilo I, Andersson N. Acceptability and challenges of introducing an educational audio-drama about gender violence and HIV prevention into schools in Botswana: an implementation review. *AIDS Care* 2019;**31**:1397–402.
135. Cooper A, Moolman B, Matandela M. *Outcomes Evaluation: SKILLZ Street Plus: A Grassroot Soccer Initiative*. Human Sciences Research Council; 2017. URL: <http://repository.hsrc.ac.za/handle/20.500.11910/11279> (accessed April 2023).
136. Maphosa B. *Reducing Gender-Based Violence: An Action Research Among Zimbabwean Youth*. PhD thesis. Durban, South Africa: Durban University of Technology; 2017.
137. Merrill KG, Merrill JC, Hershow RB, Barkley C, Rakosa B, DeCelles J, Harrison A. Linking at-risk South African girls to sexual violence and reproductive health services: a mixed-methods assessment of a soccer-based HIV prevention program and pilot SMS campaign. *Eval Program Plann* 2018;**70**:12–24.
138. Mgalla Z, Schapink D, Ties Boerma J. Protecting school girls against sexual exploitation: a guardian programme in Mwanza, Tanzania. *Reprod Health Matters* 1998;**6**:19–30.
139. Nicholson F, Mukaro T. *The Effectiveness of a Model for Addressing School-Related Gender-Based Violence (SRGBV) in South Africa: An Evaluation of the 'Zero Tolerance School Alliance' Limpopo Province, South Africa*. TVEP; 2018. URL: [www.ungei.org/sites/default/files/The-Effectiveness-of-a-Model-for-Addressing-School-Related-Gender-Based-Violence-%28SRGBV%29-in-South-Africa-An-Evaluation-of-the-Zero-Tolerance-School-Alliance-2018-eng.pdf](http://www.ungei.org/sites/default/files/The-Effectiveness-of-a-Model-for-Addressing-School-Related-Gender-Based-Violence-%28SRGBV%29-in-South-Africa-An-Evaluation-of-the-Zero-Tolerance-School-Alliance-2018-eng.pdf) (accessed April 2023).
140. Sarr KG. *Evaluation Finale. Fonds de solidarité prioritaire (FSP): FSP 2015–25. Appui à la lutte contre les violences de genre en milieu scolaire (VGMS)*. UNICEF; 2019. URL: [www.ungei.org/fr/publication/evaluation-finale-fonds-de-solidarite-prioritaire-2015-25](http://www.ungei.org/fr/publication/evaluation-finale-fonds-de-solidarite-prioritaire-2015-25) (accessed April 2023).
141. Schwandt HM, Underwood C. Engaging school personnel in making schools safe for girls in Botswana, Malawi, and Mozambique. *Int J Educ Dev* 2016;**46**:53–8.
142. dos Santos KB, Murta SG. Peers as agents of dating violence prevention: feasibility analysis of an intervention. *Trends Psychol* 2019;**27**:631–46.
143. García Escobar MB. *Promotores educadores: Una estrategia preventiva para enfrentar las situaciones de violencia familiar y sexual en niños, niñas y adolescentes en Villa El Salvador y San Martín de Porres–Lima, Perú (Período 2012–2014)*. Lima, Peru: Pontificia Universidad Católica del Perú; 2020.
144. Gomez Gonzalez ML. *Relaciones Románticas Constructivas 'RRC': Evaluación de un programa psicoeducativo para el desarrollo de relaciones de pareja constructivas en la adolescencia*. PhD thesis. Bogota, Colombia: Universidad de Los Andes; 2014.
145. Murta SG, Moore RA, Miranda AAV, Cangussú EDA. Effects of a dating violence prevention program. *Psico-USF* 2016;**21**:381–93.
146. Guedes de Oliveira RN, Serpa da Fonseca RM. A abordagem da violência nas relações de intimidade de adolescentes por meio de um jogo on line. *CIAIQ2016* 2016;**2**:875–84.
147. Jordan J, Mossman E. 'Back off buddy, this is my body, not yours': empowering girls through self-defense. *Violence Against Women* 2018;**24**:1591–613.
148. Kearney S, Gleeson C, Leung L, Ollis D, Joyce A. *Respectful Relationships Education in Schools: The Beginnings of Change*. Final evaluation report. Melbourne, VIC: Our Watch; 2016.
149. Loudon W. *Review of Appropriateness and Efficacy of the Safe Schools Coalition Australia Program Resources*. Canberra, ACT: Department of Education, Australian Government; 2016.

150. Mills M. The human relationships education curriculum and gender and violence programs. *Change Transform Educ* 1998;**1**:68–81.
151. Ollis D. A 'respectful relationships' approach: could it be the answer to preventing gender-based violence? *Redress* 2011;**20**:19–26.
152. Ollis D. The power of feminist pedagogy in Australia: vagina shorts and the primary prevention of violence against women. *Gend Educ* 2017;**29**:461–75.
153. Shevlin A, Gill PR. Parental attitudes to the Australian anti-bullying safe schools program: a critical discourse analysis. *Soc Psychol Educ* 2020;**23**:891–915.
154. Struthers K, Tilbury C, Parmenter N. *Young People as Agents of Change in Preventing Violence Against Women*. Research report. Sydney, ACT: Australia's National Research Organisation for Women's Safety; 2019.
155. Jewkes R, Gevers A, Chirwa E, Mahlangu P, Shamu S, Shai N, Lombard C. RCT evaluation of Skhokho: a holistic school intervention to prevent gender-based violence among South African Grade 8s. *PLOS ONE* 2019;**14**:e0223562.
156. Taylor BG, Mumford EA, Liu W, Stein ND. The effects of different saturation levels of the Shifting Boundaries intervention on preventing adolescent relationship abuse and sexual harassment. *J Exp Criminol* 2017;**13**:79–100.
157. Taylor BG, Stein N, Burden FF. Exploring gender differences in dating violence/harassment prevention programming in middle schools: results from a randomized experiment. *J Exp Criminol* 2010;**6**:419–45.
158. Sabella RA. *The Effectiveness of a Developmental Guidance Unit and Self-instruction Module About Sexual Harassment Among Seventh-grade Students*. PhD thesis. Gainesville, FL: University of Florida; 1995.
159. Rosenbluth B, Whitaker DJ, Sanchez E, Valle LA. The Expect Respect project: preventing bullying and sexual harassment in US elementary schools. In: Smith PK, Pepler D, Rigby K, editors. *Bullying in Schools: How Successful Can Interventions Be?* New York, NY: Cambridge University Press; 2004. pp. 211–33.
160. Baiocchi M, Omondi B, Langat N, Boothroyd DB, Sinclair J, Pavia L, et al. A behavior-based intervention that prevents sexual assault: the results of a matched-pairs, cluster-randomized study in Nairobi, Kenya. *Prev Sci* 2017;**18**:818–27.
161. Boduszek D, Debowska A, Jones AD, Ma M, Smith D, Willmott D, et al. Prosocial video game as an intimate partner violence prevention tool among youth: a randomised controlled trial. *Comput Human Behav* 2019;**93**:260–6.
162. Miller JR. *A Social Learning Perspective Toward the Prevention of Dating Violence: An Evaluation of a Group Counseling Model*. PhD thesis. Doral, FL: Miami Institute of Psychology, Caribbean Center for Advanced Studies; 1998.
163. dos Santos KB, Murta SG, Vinha LGD, de Deus JS. Efficacy of a bystander intervention for preventing dating violence in Brazilian adolescents: short-term evaluation. *Psicologia: Reflexão e Crítica* 2019;**32**:14.
164. Munoz-Rivas MJ, Redondo-Rodriguez N, Ronzon-Tirado RC. Dating violence prevention: evaluation of the program PREVIO. *Rev Psicol Clin con Niños Adolesc* 2019;**6**:18–23.
165. Gage AJ, Honoré JG, Deleon J. *Pilot Test of a Dating Violence-prevention Curriculum Among High School Students: Emerging Evidence of Effectiveness in a Low-income Country*. New Orleans: Tulane University; 2016.

166. Niolon PH, Vivolo-Kantor AM, Tracy AJ, Latzman NE, Little TD, DeGue S, *et al.* An RCT of Dating Matters: effects on teen dating violence and relationship behaviors. *Am J Prev Med* 2019;**57**:13–23.
167. Chamberland A, Cantin-Drouin M, Damant D. Assessment of the impact of Saisir: a dating violence prevention program. *Can Soc Work Rev/Revue Canadienne de Service Social* 2014;**31**:125–39.
168. de Lijster GP, Felten H, Kok G, Kocken PL. Effects of an interactive school-based program for preventing adolescent sexual harassment: a cluster-randomized controlled evaluation study. *J Youth Adolesc* 2016;**45**:874–86.
169. Fernández-González LC E, Sánchez-Álvarez N. Efficacy of a brief intervention based on an incremental theory of personality in the prevention of adolescent dating violence: a randomized controlled trial. *Psychosoc Interv* 2020;**29**:9–18.
170. Yom YH, Lee KE. Effects of a CD-ROM educational program on sexual knowledge and attitude. *Comput Inform Nurs* 2005;**23**:214–9.
171. Ekhtiari SY, Shojaeizadeh D, Rahimi Foroushani A, Ghofranipour F, Ahmadi B. The effect of an intervention based on the PRECEDE-PROCEED model on preventive behaviors of domestic violence among Iranian high school girls. *Iran Red Crescent Med J* 2013;**15**:21–8.
172. Garmaroudi G, Sarlak B, Sadeghi R. يیتشادهب داقت عا یوگلا ساسارب یشزوم آهلخادم ریثأت یسررب  
J انارعت رمش هطسوت مود هرود رتخد نازوم آشناد رد یگنخ تنوشخ زا مدننکیری گشیپ ی اهراتفسر رب  
*Knowl Health Basic Med Sci* 2016;**11**:1395.
173. Jung HY, Min HS. Development & evaluation of web-based dating violence prevention program for middle school students. *J Korean Acad Soc Nurs Educ* 2013;**19**:627–39.
174. Decker MR, Wood SN, Ndinda E, Yenokyan G, Sinclair J, Maksud N, *et al.* Sexual violence among adolescent girls and young women in Malawi: a cluster-randomized controlled implementation trial of empowerment self-defense training. *BMC Public Health* 2018;**18**:1341.
175. Devries KM, Knight L, Allen E, Parkes J, Kyegombe N, Naker D. Does the Good Schools Toolkit reduce physical, sexual and emotional violence, and injuries, in girls and boys equally? A cluster-randomised controlled trial. *Prev Sci* 2017;**18**:839–53.
176. Jemmott JB, 3rd, O’Leary A, Jemmott LS, Ngwane ZP, Teitelman AM, Makiwane MB, Bellamy SL. Effect of a behavioral intervention on perpetrating and experiencing forced sex among South African adolescents: a secondary analysis of a cluster randomized trial. *JAMA Netw Open* 2018;**1**:e181213.
177. Cameron CA, Byers ES, Miller SA, McKay SL, St Pierre M, Glenn S. *Dating Violence Prevention in New Brunswick. Reported Prepared for Status of Women Canada.* Fredericton, NB: University of New Brunswick; 2007.
178. Jaime MC, Stocking M, Freire K, Perkinson L, Ciaravino S, Miller E. Using a domestic and sexual violence prevention advocate to implement a dating violence prevention program with athletes. *Health Educ Res* 2016;**31**:679–96.
179. Tutty LM. *Healthy Relationships: Preventing Teen Dating Violence. An Evaluation of the Teen Violence Prevention Program.* Toronto, ON: Canadian Women’s Foundation; 2011.
180. Miller E, Jones KA, Ripper L, Paglisotti T, Mulbah P, Abebe KZ. An athletic coach-delivered middle school gender violence prevention program: a cluster randomized clinical trial. *JAMA Pediatr* 2020;**174**:241–9.
181. Jouriles EN, McDonald R, Rosenfield D, Sargent KS. Increasing bystander behavior to prevent adolescent relationship violence: a randomized controlled trial. *J Consult Clin Psychol* 2019;**87**:3–15.

182. Levesque DA, Johnson JL, Welch CA, Prochaska JM, Paiva AL. Teen dating violence prevention: cluster-randomized trial of teen choices, an online, stage-based program for healthy, nonviolent relationships. *Psychol Violence* 2016;**6**:421–32.
183. Walther DJ. Wife abuse prevention: effects of information on attitudes of high school boys. *J Prim Prev* 1986;**7**:84–90.
184. Bando R, Hidalgo N, Land A. Education with a social focus on gender attitudes: experimental evidence from secondary education in El Salvador. *J Econ Race Policy* 2019;**2**:225–39.
185. Jaycox LH, McCaffrey D, Eiseman B, Aronoff J, Shelley GA, Collins RL, Marshall GN. Impact of a school-based dating violence prevention program among Latino teens: randomized controlled effectiveness trial. *J Adolesc Health* 2006;**39**:694–704.
186. Joppa MC, Rizzo CJ, Nieves AV, Brown LK. Pilot investigation of the Katie Brown Educational Program: a school-community partnership. *J Sch Health* 2016;**86**:288–97.
187. Miller E, Goldstein S, McCauley HL, Jones KA, Dick RN, Jetton J, et al. A school health center intervention for abusive adolescent relationships: a cluster RCT. *Pediatrics* 2015;**135**:76–85.
188. Peskin MF, Markham CM, Shegog R, Baumler ER, Addy RC, Temple JR, et al. Adolescent dating violence prevention program for early adolescents: the Me & You randomized controlled trial, 2014–2015. *Am J Public Health* 2019;**109**:1419–28.
189. Peskin MF, Markham CM, Shegog R, Baumler ER, Addy RC, Tortolero SR. Effects of the it's your game ... keep it real program on dating violence in ethnic-minority middle school youths: a group randomized trial. *Am J Public Health* 2014;**104**:1471–7.
190. Wolfe DA, Crooks C, Jaffe P, Chiodo D, Hughes R, Ellis W, et al. A school-based program to prevent adolescent dating violence: a cluster randomized trial. *Arch Pediatr Adolesc Med* 2009;**163**:692–9.
191. Pacifici C, Stoolmiller M, Nelson C. Evaluating a prevention program for teenagers on sexual coercion: a differential effectiveness approach. *J Consult Clin Psychol* 2001;**69**:552–9.
192. Espelage DL, Low S, Polanin JR, Brown EC. The impact of a middle school program to reduce aggression, victimization, and sexual violence. *J Adolesc Health* 2013;**53**:180–6.
193. Gonzalez-Guarda RM, Guerra JE, Cummings AA, Pino K, Becerra MM. Examining the preliminary efficacy of a dating violence prevention program for Hispanic adolescents. *J Sch Nurs* 2015;**31**:411–21.
194. Kershner RE. *The Efficacy of Adolescent Rape Prevention Education*. PhD thesis. Morgantown, WV: West Virginia University; 1995.
195. Malo-Juvera VS. The effect of literary instruction on adolescents' rape myth acceptance. *Res Teach English* 2014;**48**:407–27.
196. Avery-Leaf S, Cascardi M, O'Leary KD, Cano A. Efficacy of a dating violence prevention program on attitudes justifying aggression. *J Adolesc Health* 1997;**21**:11–7.
197. Edwards KM, Banyard VL, Sessarego SN, Waterman EA, Mitchell KJ, Chang H. Evaluation of a bystander-focused interpersonal violence prevention program with high school students. *Prev Sci* 2019;**20**:488–98.
198. Fay KE, Medway FJ. An acquaintance rape education program for students transitioning to high school. *Sex Educ* 2006;**6**:223–36.
199. Coyle KK, Anderson P, Franks HM, Walker JD, Glassman JR. You-me-us: results of a cluster randomized trial of a healthy relationships approach to sexual risk reduction. *J Prim Prev* 2019;**40**:607–29.

200. Taylor B, Stein N, Burden F. The effects of gender violence/harassment prevention programming in middle schools: a randomized experimental evaluation. *Violence Vict* 2010;**25**:202–23.
201. Raible CA, Dick R, Gilkerson F, Mattern CS, James L, Miller E. School nurse-delivered adolescent relationship abuse prevention. *J Sch Health* 2017;**87**:524–30.
202. Watson LA. 'Wow ... They Care, Right?' *Making Schools Safe(r) for Lesbian, Gay, Bisexual, Transgender and Questioning Youth*. DEd thesis. East Bay, CA: California State University; 2012.
203. Diegel RP. *Participation in a Dating Violence Prevention Psychoeducational Support Group for Adolescent Females: A Phenomenological Inquiry*. PhD thesis. Cincinnati, OH: Union Institute; 1999.
204. Edwards KM, Waterman EA, Lee KDM, Himlin L, Parm K, Banyard VL. Feasibility and acceptability of a high school relationship abuse and sexual assault bystander prevention program: school personnel and student perspectives. *J Interpers Violence* 2021;**36**:NP7070–85.
205. Makleff S, Garduno J, Zavala RI, Barindelli F, Valades J, Billowitz M, et al. Preventing intimate partner violence among young people – a qualitative study examining the role of comprehensive sexuality education. *Sex Res Social Policy* 2019;**17**:314–25.
206. Wernick LJ, Dessel AB, Kulick A, Graham LF. LGBTQ youth creating change: developing allies against bullying through performance and dialogue. *Child Youth Serv Rev* 2013;**35**:1576–86.
207. Singh AA. The use of Popular Opinion Leader (POL) groups and the reduction of 'gay bullying' in middle school: a case study inquiry of group leader experiences. *J Spec Group Work* 2013;**38**:184–206.
208. Edwards KM, Lee KDM, Waterman EA. Implementation science analysis of a school-delivered bystander-focused violence prevention curriculum. *Sch Soc Work J* 2019;**44**:1–23.
209. Gardner SP. *Final Summary Report for the Evaluation of the Connections: Dating and Emotions Curriculum*. Brookings, SD: South Dakota State University; 2005.
210. Grimm AJ. *Go for the Gold: Relationship Education Program Evaluation*. MSc thesis. Dayton, OH: University of Dayton; 2011.
211. Levesque DA, Johnson JL, Prochaska JM. Teen choices, an online stage-based program for healthy, nonviolent relationships: development and feasibility trial. *J Sch Health* 2017;**16**:376–85.
212. Turner CJ. *Using Peers to Influence Dating Relationships: An Evaluation of a Dating Violence Prevention Program*. Master's thesis. Newark, DE: University of Delaware; 2006.
213. Elias-Lambert N, Boyas JF, Black BM, Schoech RJ. Preventing substance abuse and relationship violence: proof-of-concept evaluation of a social, multi-user, tablet-based game. *Child Youth Serv Rev* 2015;**53**:201–10.
214. Fawson PR, Broce R, Bonner B, Wright R. Adolescents' experiences: programming implications for in-school dating violence prevention programs. *Sch Soc Work J* 2016;**41**:1–16.
215. Guillot-Wright SP, Lu Y, Torres ED, Le VD, Hall HR, Temple JR. Design and feasibility of a school-based text message campaign to promote healthy relationships. *Sch Ment Health* 2018;**10**:428–36.
216. Harrington LC. *'It Doesn't Take Much': Dissemination of Violence Prevention Messaging by High School Athletes on Girls' Sports Teams*. Master's thesis. Seattle, WA: University of Washington; 2019.
217. Elias-Lambert N, Black B, Sharma Y. Middle school youth: satisfaction with and responses to a dating violence and sexual assault prevention program. *J Sch Violence* 2010;**9**:136–53.

218. Ball B, Kerig PK, Rosenbluth B. 'Like a family but better because you can actually trust each other': the Expect Respect dating violence prevention program for at-risk youth. *Health Promot Pract* 2009;**10**:45S–58S.
219. Belknap RA, Haglund K, Felzer H, Pruszynski J, Schneider J. A theater intervention to prevent teen dating violence for Mexican-American middle school students. *J Adolesc Health* 2013;**53**:62–7.
220. Enriquez M, Kelly PJ, Cheng AL, Hunter J, Mendez E. An intervention to address interpersonal violence among low-income midwestern Hispanic-American teens. *J Immigr Minor Health* 2012;**14**:292–9.
221. Eisman AB, Hicks M, Kernsmith PD, Rupp L, Smith-Darden JP, Zimmerman MA. Adapting an evidence-based positive youth development intervention to prevent sexual and teen dating violence. *Transl Behav Med* 2019;**20**:20.
222. Ouellett ML. *A Multicultural Organization Development Examination of School-Based Change Strategies to Address the Needs of Gay Youth*. PhD thesis. Amherst, MA: University of Massachusetts Amherst; 1998.
223. Ball B, Holland KM, Marshall KJ, Lippy C, Jain S, Souders K, Westby RP. Implementing a targeted teen dating abuse prevention program: challenges and successes experienced by Expect Respect facilitators. *J Adolesc Health* 2015;**56**:S40–6.
224. Kervin D, Obinna J. Youth action strategies in the primary prevention of teen dating violence. *J Fam Soc Work* 2010;**13**:362–74.
225. Rizzo AJ, Orr N, Shaw N, Farmer C, Chollet A, Young H, Berry V, Rigby E, Hagell A, Bonell C, Melendez-Torres GJ. Exploring the activities and target audiences of school-based violence prevention programs: systematic review and intervention component analysis. *Trauma Violence Abuse* 2023;**24**:3593–614.
226. Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychology research. *Qual Res Psychol* 2015;**12**:202–22.
227. Schütz A, Brodersen A, Gurswitsch A, Natanson MA, Schütz I. *Collected Papers*. The Hague, The Netherlands: Martinus Nijhoff; 1962.
228. Markham WA, Aveyard P. A new theory of health promoting schools based on human functioning, school organisation and pedagogic practice. *Soc Sci Med* 2003;**56**:1209–20.
229. Keshavarz N, Nutbeam D, Rowling L, Khavarpour F. Schools as social complex adaptive systems: a new way to understand the challenges of introducing the health promoting schools concept. *Soc Sci Med* 2010;**70**:1467–74.
230. Edwards KM, Sessarego SN, Mitchell KJ, Chang H, Waterman EA, Banyard VL. Preventing teen relationship abuse and sexual assault through bystander training: intervention outcomes for school personnel. *Am J Community Psychol* 2020;**65**:160–72.
231. Achyut P, Bhatla N, Verma H, Uttamacharya SG, Bhattacharya S, Verma R. *Towards Gender Equality. The GEMS Journey Thus Far. An Evaluation Report of the Gender Equity Movement in Schools (GEMS) Program in Jharkhand*. New Delhi, India: International Center for Research on Women; 2016.
232. Knowles EA. *Experiences of Gay and Lesbian Educators Who Work in Massachusetts Schools Participating in the Safe Schools Program*. PhD thesis. Amherst, MA: University of Massachusetts; 1997.
233. Madsen C. *Adolescents' Critical Consciousness of Sexism and Violence in Relationships: An Evaluation of the Toronto Board of Education's 'Parallel Retreats' Program*. Master's thesis. Toronto, Canada: University of Toronto; 1994.



234. Ekhtiari YS, Shojaeizadeh D, Foroushani AR, Ghofranipour F, Ahmadi B. Effect of an intervention on attitudes towards domestic violence among Iranian girls. *J Pak Med Assoc* 2014;**64**:987–92.
235. Busch-Armendariz NB, Kalergis K, Little A, Woo H. *An Evaluation of the Texas Team's Teen Dating Violence Awareness and Prevention Toolkit*. Austin, TX: Texas Dating Violence Prevention Team, University of Texas at Austin; 2008.
236. Genereux AS. *Teen Dating Violence Prevention to Healthy Relationship Education: Case Study of a Montanan Teen Dating Violence Prevention Program*. PhD thesis. Madison, WI: University of Wisconsin; 2020.
237. Whitaker DJ, Rosenbluth B, Valle LA, Sanchez E. Expect Respect: a school-based intervention to promote awareness and effective responses to bullying and sex harassment. In Espelage DL, Swearer SM, editors. *Bullying in American Schools: A Social-ecological Perspective on Prevention and Intervention*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2003. pp. 327–50.
238. Coker AL, Bush HM, Brancato CJ, Clear ER, Recktenwald EA. Bystander program effectiveness to reduce violence acceptance: RCT in high schools. *J Fam Violence* 2019;**34**:153–64.
239. Espelage DL, Low S, Van Ryzin MJ, Polanin JR. Clinical trial of Second Step Middle School Program: impact on bullying, cyberbullying, homophobic teasing, and sexual harassment perpetration. *School Psych Rev* 2015;**44**:464–79.
240. Foshee VA, Bauman KE, Greene WF, Koch GG, Linder GF, MacDougall JE. The Safe Dates program: 1-year follow-up results. *Am J Public Health* 2000;**90**:1619–22.
241. Foshee VA, Bauman KE, Ennett ST, Suchindran C, Benefield T, Linder GF. Assessing the effects of the dating violence prevention program 'Safe Dates' using random coefficient regression modeling. *Prev Sci* 2005;**6**:245–58.
242. Foshee VA, Bauman KE, Ennett ST, Linder GF, Benefield T, Suchindran C. Assessing the long-term effects of the Safe Dates program and a booster in preventing and reducing adolescent dating violence victimization and perpetration. *Am J Public Health* 2004;**94**:619–24.
243. Foshee VA. Involving schools and communities in preventing adolescent dating abuse. In Arriaga XB, Oskamp S, editors. *Addressing Community Problems: Psychological Research and Interventions*. Thousand Oaks, CA: SAGE Publications Ltd; 1998. pp. 104–29.
244. Gale JY. *Youthink in Action: Transforming Homophobic School Culture Through Participatory Action Research and Theatre of the Oppressed*. Master's thesis. Los Angeles, CA: University of Southern California; 2011.
245. Jaime MCD, McCauley HL, Tancredi DJ, Decker MR, Silverman JG, O'Connor B, Miller E. Implementing a coach-delivered dating violence prevention program with high school athletes. *Prev Sci* 2018;**19**:1113–22.
246. Astor RA, Meyer HA, Pitner RO. Elementary and middle school students' perceptions of violence-prone school subcontexts. *Elem Sch J* 2001;**101**:511–28.
247. Munoz-Fernandez N, Ortega-Rivera J, Nocentini A, Menesini E, Sanchez-Jimenez V. The efficacy of the 'Dat-E Adolescence' prevention program in the reduction of dating violence and bullying. *Int J Environ Res Public Health* 2019;**16**:31.
248. Taylor B, Stein N, Mack AR, Horwood TJ, Burden F. *Experimental Evaluation of Gender Violence/Harassment Prevention Programs in Middle Schools*. Fairfax, VA: ICF International; 2008.
249. Cramer EP, Ross AI, McLeod DA. The impact on peer facilitators of facilitating a school-based healthy relationship program for teens. *Sch Soc Work J* 2015;**40**:23–41.

250. Brunk T. *The Use of Drama to Explore Violence Within Relationships: A Study of Young Men's Experiences*. MA thesis. Waterloo, ON: Wilfrid Laurier University; 1993.
251. Ryding J. *A Game-based Intervention for Adolescent Dating Violence: A Focus Group Investigation of 'Green Acres High'*. Master's thesis. Karlskrona, Sweden: Blekinge Tekniska Högskola; 2013.
252. Hawe P, Shiell A, Riley T. Theorising interventions as events in systems. *Am J Community Psychol* 2009;**43**:267–76.
253. Hawe P. Lessons from complex interventions to improve health. *Annu Rev Public Health* 2015;**36**:307–23.
254. Hawe P. Minimal, negligible and negligent interventions. *Soc Sci Med* 2015;**138**:265–8.
255. Coker AL, Bush HM, Clear ER, Brancato CJ, McCauley HL. Bystander program effectiveness to reduce violence and violence acceptance within sexual minority male and female high school students using a cluster RCT. *Prev Sci* 2020;**21**:434–44.
256. Lapointe A, Dunlop C, Crooks C. Feasibility and fit of a mental health promotion program for LGBTQ+ YOUTH. *J Youth Dev* 2018;**13**:100–17.
257. Bragg S, Ponsford R, Meiksin R, Bonell C, Emmerson L. Dilemmas of school-based relationships and sexuality education for and about consent. *Sex Educ* 2020;**1**:15.
258. Davidov DM, Hill K, Bush HM, Coker AL. The green light for Green Dot: a qualitative study of factors influencing adoption of an efficacious violence prevention program in high school settings. *Violence Against Women* 2019;**26**:1077801219886377.
259. Garces-Foley A. *School Superintendents' Opinions of LGBT-Inclusive Instruction: A Policy Case Study of California's Fair Education Act Implementation*. PhD thesis. Fairfax, VA: George Mason University; 2017.
260. Payne EC, Smith MJ. Refusing relevance: school administrator resistance to offering professional development addressing LGBTQ issues in schools. *Educ Adm Q* 2018;**54**:183–215.
261. Fonn E. *A Process Evaluation of the PREPARE Intervention to Promote Healthy Sexual Practices Among Adolescents in Cape Town, South Africa*. Master's thesis. As, Norway: NMBU; 2017.
262. Hertel LA. *Considering Gender in Intimate Partner Violence Prevention for Youth*. PhD thesis. Minneapolis, MN: Walden University; 2020.
263. Gibbs DA, Krieger KE, Cutbush SL, Clinton-Sherrod AM, Miller S. Implementer-initiated adaptation of evidence-based interventions: kids remember the blue wig. *Health Educ Res* 2016;**31**:405–15.
264. Cheney B. *A Case Study of the Development and Implementation of a Modified Integrated Thematic Instructional Model for Sexual Harassment Prevention*. PhD thesis. Manhattan, KS: Kansas State University; 1998.
265. Makleff S, Garduño J, Zavala RI, Valades J, Barindelli F, Cruz M, Marston C. Evaluating complex interventions using qualitative longitudinal research: a case study of understanding pathways to violence prevention. *Qual Health Res* 2021;**31**:1724–37.
266. Beardall NG. *A Program Evaluation Research Study on the Implementation of the Mentors in Violence Prevention Program in a Public High School*. PhD thesis. Cambridge, MA: Lesley University; 2007.
267. Crooks CV, Chiodo D, Zwarych S, Hughes R, Wolfe DA. Predicting implementation success of an evidence-based program to promote healthy relationships among students two to eight years after teacher training. *Can J Commun Ment Health* 2013;**32**:125–38.

268. Payne DL, Lonsway KA, Fitzgerald LF. Rape myth acceptance: exploration of its structure and its measurement using the Illinois Rape Myth Acceptance Scale. *J Res Pers* 1999;**33**:27–68.
269. Price EL, Byers ES, Belliveau N, Bonner R, Caron B, Doiron D, *et al.* The attitudes towards dating violence scales: development and initial validation. *J Fam Violence* 1999;**14**:351–75.
270. Banyard V, Moynihan M, Cares A, Warner R. Readiness to help scale. *APA PsycTests* 2014. <https://doi.org/10.1037/t28157-000>
271. Edwards KM, Rodenhizer KA, Eckstein RP. School personnel's bystander action in situations of dating violence, sexual violence, and sexual harassment among high school teens: a qualitative analysis. *J Interpers Violence* 2020;**35**:2358–69.
272. Arnab S, Brown K, Clarke S, Dunwell I, Lim T, Suttie N, *et al.* The development approach of a pedagogically-driven serious game to support Relationship and Sex Education (RSE) within a classroom setting. *Comput Educ* 2013;**69**:15–30.
273. World Economic Forum. *Global Gender Gap Report 2021*. 2021. URL: [www3.weforum.org/docs/WEF\\_GGGR\\_2021.pdf](http://www3.weforum.org/docs/WEF_GGGR_2021.pdf) (accessed April 2023).
274. Espelage DL, Bub K, Van Ryzin M, Holt MK. *Effects of a Middle School Social-Emotional Learning Program on Bullying, Teen Dating Violence, Sexual Violence, and Substance Use in High School*. Rockville, MD; National Institute of Justice, National Criminal Justice Reference Service; 2017.
275. DeGue S, Niolon PH, Estefan LF, Tracy AJ, Le VD, Vivolo-Kantor AM, *et al.* Effects of Dating Matters® on sexual violence and sexual harassment outcomes among middle school youth: a cluster-randomized controlled trial. *Prev Sci* 2021;**22**:175–85.
276. Taylor BG, Mumford EA, Stein ND. Effectiveness of 'Shifting Boundaries' teen dating violence prevention program for subgroups of middle school students. *J Adolesc Health* 2015;**56**:S20–6.
277. Waterman EA, Edwards KM, Banyard VL, Chang H. Age and sexual orientation moderated the effects of a bystander-focused interpersonal violence prevention program for high school students. *Prev Sci* 2022;**23**:96–107.
278. Welsh Government. *Good Practice Guide: A Whole Education Approach to Violence Against Women, Domestic Abuse & Sexual Violence in Wales*. 2015. URL: <https://gov.wales/sites/default/files/publications/2019-06/good-practice-guide-whole-education-approach-to-violence-against-women-domestic-abuse-and-sexual-violence.pdf> (accessed April 2023).
279. Skivington K, Matthews L, Simpson SA, Craig P, Baird J, Blazeby JM, *et al.* A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ* 2021;**374**:n2061.
280. Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000;**321**:694–6.
281. Craig P, Di Ruggiero E, Frolich KL, Mykhalovskiy E, White M, Campbell R, *et al.* *Taking Account of Context in Population Health Intervention Research: Guidance for Producers, Users and Funders of Research*. Southampton, UK: NIHR Evaluation, Trials and Studies Coordinating Centre; 2018.
282. Langford R, Bonell C, Jones H, Campbell R. Obesity prevention and the Health promoting Schools framework: essential components and barriers to success. *Int J Behav Nutr Phys Act* 2015;**12**:15.
283. Latzman NE, D'Inverno AS, Niolon PH, Reidy DE. Chapter 12 – gender inequality and gender-based violence: extensions to adolescent dating violence. In Wolfe DA, Temple JR, editors. *Adolescent Dating Violence*. San Diego, CA: Academic Press; 2018. pp. 283–314.

284. Bonell C, Jamal F, Melendez-Torres GJ, Cummins S. 'Dark logic': theorising the harmful consequences of public health interventions. *J Epidemiol Community Health* 2015;**69**:95–8.
285. Lameiras-Fernández M, Martínez-Román R, Carrera-Fernández MV, Rodríguez-Castro Y. Sex education in the spotlight: what is working? Systematic review. *Int J Environ Res Public Health* 2021;**18**:2555.
286. Allen L. 'Say everything': exploring young people's suggestions for improving sexuality education. *Sex Educ* 2005;**5**:389–404.
287. Patterson SP, Hilton S, Flowers P, McDaid LM. What are the barriers and challenges faced by adolescents when searching for sexual health information on the internet? Implications for policy and practice from a qualitative study. *Sex Transm Infect* 2019;**95**:462–7.
288. Goldfarb ES, Lieberman LD. Three decades of research: the case for comprehensive sex education. *J Adolesc Health* 2021;**68**:13–27.
289. Gardner EA. Abstinence-only sex education: college students' evaluations and responses. *Am J Sex Educ* 2015;**10**:125–39.
290. Lindberg LD, Maddow-Zimet I, Boonstra H. Changes in adolescents' receipt of sex education, 2006–2013. *J Adolesc Health* 2016;**58**:621–7.
291. Forrest S, Strange V, Oakley A; The Ripple Study Team. What do young people want from sex education? The results of a needs assessment from a peer-led sex education programme. *Cult Health Sex* 2004;**6**:337–54.
292. Ketting E, Friele M, Michielsen K; European Expert Group on Sexuality Education. Evaluation of holistic sexuality education: a European expert group consensus agreement. *Eur J Contracept Reprod Health Care* 2016;**21**:68–80.
293. UNAIDS, UNFPA, UNICEF, UN Women, WHO. *International Technical Guidance on Sexuality Education: An Evidence-Informed Approach*. Paris, France: UNESCO Publishing; 2018.
294. UNFPA. *The Evaluation of Comprehensive Sexuality Education Programmes: A Focus on the Gender and Empowerment Outcomes*. New York, NY: United Nations Population Fund; 2015.
295. World Health Organization. *WHO Recommendations on Adolescent Sexual and Reproductive Health and Rights* Geneva, Switzerland: World Health Organization; 2018. URL: [www.who.int/publications/i/item/9789241514606](http://www.who.int/publications/i/item/9789241514606) (accessed April 2023).
296. World Health Organization. *Health-Promoting Schools: A Healthy Setting for Living, Learning and Working*. Geneva, Switzerland: World Health Organization; 1998. URL: <https://apps.who.int/iris/handle/10665/63868> (accessed April 2023).
297. Home Office. *Tackling Violence Against Women and Girls*. 2021. URL: [www.gov.uk/government/publications/tackling-violence-against-women-and-girls-strategy](http://www.gov.uk/government/publications/tackling-violence-against-women-and-girls-strategy) (accessed April 2023).
298. Welsh Government. *Violence Against Women, Domestic Abuse and Sexual Violence (Wales) Act 2015*. 2015. URL: [www.legislation.gov.uk/anaw/2015/3/contents/enacted](http://www.legislation.gov.uk/anaw/2015/3/contents/enacted)
299. Great Britain. *Domestic Abuse Act 2021*. 2021. URL: [www.legislation.gov.uk/ukpga/2021/17/contents/enacted](http://www.legislation.gov.uk/ukpga/2021/17/contents/enacted) (accessed April 2023).
300. Estyn. *'We Don't Tell Our Teachers': Experiences of Peer-on-Peer Sexual Harassment Among Secondary School Pupils in Wales*. Cardiff: Estyn; 2021.
301. Oringanje C, Meremikwu MM, Eko H, Esu E, Meremikwu A, Ehiri JE. Interventions for preventing unintended pregnancies among adolescents. *Cochrane Database Syst Rev* 2016;**2**:CD005215.

302. Dozois E, Wells L, Crooks C. Chapter 12. Developing a comprehensive, evidence-based strategy for violence prevention: lessons learned from Alberta, Canada. In: Taylor MF, Khan U, Pooley JA, editors. *Crime and Violence Prevention: Moving Beyond Hot-Stove Policing and Perpetrator Rehabilitation*. Hauppauge, NY: Nova Science Publishers; 2016. pp. 215–30.
303. Jones KA, Tancredi DJ, Abebe KZ, Paglisotti T, Miller E. Cases of sexual assault prevented in an athletic coach-delivered gender violence prevention program. *Prev Sci* 2021;**22**:504–8.
304. Watts VB. *Project PRIDE: Engaging High School Students in Reducing Teen Dating Violence in Their School*. PhD thesis. Oxford, OH: Miami University; 2016.
305. Luo F, DeGue S, Le VD. Estimating from the payer perspective the implementation cost of Dating Matters®: a comprehensive teen dating violence prevention model. *J Interpers Violence* 2020;**37**:NP9144–67.
306. Lerner JM. *Sexual Harassment Policies and Procedures: A Case Study of the Implementation Methods of Three School Districts*. PhD thesis. East Whiteland, PA: Immaculata College; 1999.
307. Noonan RK, Emshoff JG, Mooss A, Armstrong M, Weinberg J, Ball B. Adoption, adaptation, and fidelity of implementation of sexual violence prevention programs. *Health Promot Pract* 2009;**10**:59S–70S.
308. Bush JL, Bush HM, Coker AL, Brancato CJ, Clear ER, Recktenwald EA. Total and marginal cost analysis for a high school based bystander intervention. *J Sch Violence* 2018;**17**:152–63.
309. Jones R, Booker EA, Akintobi TH. A media reflection phase II: implementation and evaluation of a media intervention combating mass media's influence on the cycle of disadvantage and disability. *Int J Child Health Hum Dev* 2010;**3**:391–405.
310. Winegust AK. *Pass It On: An Evaluation of a Sexualized Violence Prevention Program for Middle School and High School Students*. Master's thesis. Toronto, Canada: University of Toronto; 2015.
311. Baker CK, Naai R, Mitchell J, Trecker C. Utilizing a train-the-trainer model for sexual violence prevention: findings from a pilot study with high school students of Asian and Pacific Islander descent in Hawai'i. *Asian Am J Psychol* 2014;**5**:106–15.
312. Cutbush S, Gibbs D, Krieger K, Clinton-Sherrod M, Miller S. Implementers' perspectives on fidelity of implementation: 'Teach every single part' or 'Be right with the curriculum?'. *Health Promot Pract* 2017;**18**:275–82.
313. Hall WJ. *The Implementation and Effectiveness of Policy Interventions for School Bullying*. PhD thesis. Chapel Hill, NC: University of North Carolina; 2016.
314. Wilson CP. *Efficacy of the Student-Related Sexual Harassment Policy of a Large, Urban School District in the Southwest*. DEd thesis. Tempe, AZ: Arizona State University; 2005.
315. Murray BL. Sexual health education for adolescents with developmental disabilities. *Health Educ J* 2019;**78**:1000–11.
316. Henshaw MWA. *High School Boys' Experiences of a Violence Against Women Prevention Program*. Master's thesis. London, ON: University of Western Ontario; 2016.
317. Chiodo DG. *A Qualitative Study of the Fidelity of Implementation of an Evidence-Based Healthy Relationships Program*. PhD thesis. London, ON: University of Western Ontario; 2017.
318. Crooks CV, Exner-Cortens D, Siebold W, Rosier M, Baker J. Building capacity to implement teen dating violence prevention: Lessons learned from the Alaska Fourth R initiative. In: Wolfe DA, Temple JR, editors. *Adolescent Dating Violence*. San Diego, CA: Elsevier Academic Press; 2018. pp. 503–21.

319. Dunlop C, Chiodo D, Crooks C. *Fourth R Programs: Are Educators Implementing 1 to 2 Years after Training?*. Public Health Agency of Canada; 2018. URL: [www.csmh.uwo.ca/docs/Implementation-Report.pdf](http://www.csmh.uwo.ca/docs/Implementation-Report.pdf) (accessed April 2023).
320. Exner-Cortens D, Spiric V, Crooks C, Syeda M, Wells L. Predictors of healthy youth relationships program implementation in a sample of Canadian middle school teachers. *Can J School Psychol* 2020;**35**:100–22.
321. Achyut P, Bhatla N, Verma R. Questioning gender norms to promote sexual reproductive health among early adolescents: evidence from a school program in Mumbai, India. In: Djamba YK, Kimuna SR, editors. *Gender-Based Violence: Perspectives from Africa, the Middle East, and India*. Switzerland: Springer International; 2015. pp. 195–213.
322. Brown K, Arnab S, Bayley J, Newby K, Joshi P, Judd B, et al. Tackling sensitive issues using a game-based environment: serious game for relationships and sex education (RSE). *Annu Rev CyberTherapy Telemed* 2012;**181**:165–71.
323. Clarke S, Arnab S, Dunwell I, Brown K. PR:EPARe: a game-based approach to relationship guidance for adolescents. *Procedia Comput Sci* 2012;**15**:38–44.
324. Bando R, Hidalgo N, Land A. *El efecto de la educación en las actitudes de género: Evidencia experimental en educación secundaria en El Salvador*. Washington, DC: Banco Interamericano de Desarrollo; 2018.
325. Gottzén L, Lindberg D, Bruno L, Joelsson T, Eriksson M, Billevik J. *Utvärdering av Mentors in Violence Prevention i Sverige*. Slutrapport: Stockholms Universitet; 2021. URL: [www.diva-portal.org/smash/record.jsf?pid=diva2%3A1516875&dswid=-1309](http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1516875&dswid=-1309) (accessed April 2023).
326. Bush HM, Coker AL, DeGue S, Clear ER, Brancato CJ, Fisher BS. Do violence acceptance and bystander actions explain the effects of Green Dot on reducing violence perpetration in high schools? *J Interpers Violence* 2019;**36**:10753–74.
327. Coker AL, Bush HM, Brancato CJ, Huang Z, Clear ER, Follingstad DR. Longer term impact of bystander training to reduce violence acceptance and sexism. *J Sch Violence* 2020;**19**:525–38.
328. Coker AL, Bush HM, Huang Z, Brancato CJ, Clear ER, Follingstad DR. How does Green Dot bystander training in high school and beyond impact attitudes toward violence and sexism in a prospective cohort? *J Interpers Viol* 2022;**37**:NP13830–53.
329. Crooks CV, Exner-Cortens D, Siebold W, Moore K, Grassgreen L, Owen P, et al. The role of relationships in collaborative partnership success: lessons from the Alaska Fourth R project. *Eval Program Plann* 2018;**67**:97–104.
330. dos Santos KB. *Mobilizando comportamentos de ajuda na rede de amizades: uma estratégia de prevenção à violência no namoro baseada nos pares e na abordagem do espectador*. PhD thesis. Brasilia, Brazil: Universidade de Brasília; 2016.
331. Gage AJ, Honoré JG, Deleon J. Impact of a dating violence prevention curriculum among high school students in Port-au-Prince, Haiti. *J Commun Healthc*. 2016;**9**(3):178–89.
332. Joyce A, Ollis D, Kearney S, Leung L, Foenander E. The influence of contextual factors on implementation fidelity in a whole school approach to prevention of violence against women. *Health Promot J Austr* 2019;**30**:238–45.
333. Joyce A, Green C, Kearney S, Leung L, Ollis D. Alignment and political will: upscaling an Australian respectful relationships program. *Health Promot Int* 2019;**34**:892–901.
334. Kearney S, Leung L, Joyce A, Ollis D, Green C. Applying systems theory to the evaluation of a whole school approach to violence prevention. *Health Promot J Austr* 2016;**27**:230–5.

335. Ollis D, Dyson S. Respectful relationships education: a case study of working in schools. In: Crisp B, Taket A, editors. *Eliminating Gender-Based Violence*. New York, NY: Routledge/Taylor & Francis; 2018. pp. 40–53.
336. Levesque DA. *A Stage-Based Expert System for Teen Dating Violence Prevention*. Final Progress Report. South Kingstown, RI: Pro-Change Behavior Systems, Inc.; 2007.
337. NCT. *A Stage-Based Expert System for Teen Dating Violence Prevention*. ClinicalTrials.gov. 2015.
338. Makleff S, Billowitz M, Garduño J, Cruz M, Silva Márquez VI, Marston C. Applying a complex adaptive systems approach to the evaluation of a school-based intervention for intimate partner violence prevention in Mexico. *Health Policy Plan* 2020;**35**:993–1002.
339. Malo-Juvera V. *The Effect of Young Adult Literature on Adolescents' Rape Myth Acceptance*. Doctoral thesis. Miami, FL: Florida International University; 2012.
340. Ponsford R, Meiksin R, Crichton J, Bragg S. Impact guaranteed? Involving potential recipients, implementers and wider stakeholders in the co-production of two school-based relationships and sex education interventions to prevent teenage pregnancy and dating and relationships violence (Preprint). *Res Sq* 2019.
341. Ponsford R, Meiksin R, Bragg S, Crichton J, Emmerson L, Tancred T, *et al*. Co-production of two whole-school sexual health interventions for English secondary schools: positive choices and project respect. *Pilot Feasibility Stud* 2021;**7**:50.
342. Jaime MC, McCauley HL, Tancredi DJ, Nettiksimmons J, Decker MR, Silverman JG, *et al*. Athletic coaches as violence prevention advocates. *J Interpers Violence* 2015;**30**:1090–111.
343. Miller E, Tancredi DJ, McCauley HL, Decker MR, Virata MCD, Anderson HA, *et al*. One-year follow-up of a coach-delivered dating violence prevention program: a cluster randomized controlled trial. *Am J Prev Med* 2013;**45**:108–12.
344. Muck C, Schiller EM, Kartner J. Prevention of sexual violence in adolescence – evaluation study on the effect of two school-based prevention programs on the willingness for disclosure and the experience of victimization. *Z Soziol Erzieh Sozi* 2018;**38**:154–70.
345. Namy S, Heilman B, Stich S, Edmeades J. *Be A Man, Change the Rules! Findings and Lessons from Seven Years of CARE International Balkans' Young Men Initiative*. Washington, DC: Young Men Initiative, CARE International, International Center for Research on Women; 2014.
346. Young Men Initiative. *The Young Men Initiative: Engaging Young Men in the Western Balkans in Gender Equality and Violence Prevention: A Case Study*. CARE International; 2012. URL: [www.care.org/news-and-stories/resources/the-young-men-initiative-a-case-study](http://www.care.org/news-and-stories/resources/the-young-men-initiative-a-case-study) (accessed April 2023).
347. Ollis D. The role of teachers in delivering education about respectful relationships: exploring teacher and student perspectives. *Health Educ Res* 2014;**29**:702–13.
348. Payne EC, Smith MJ. Safety, celebration, and risk: educator responses to LGBTQ professional development. *Teach Educ* 2012;**23**:265–85.
349. Raible CA. *Adolescent Relationship Abuse; School Nurse-Led Intervention and Prevention*. Master's thesis. Pittsburgh, PA: University of Pittsburgh; 2016.
350. Rogers MM, Wilding MA, Wood ACM. *An Evaluation of the Change Up Programme*. University of Salford; 2018. URL: <https://usir.salford.ac.uk/id/eprint/49501/> (accessed April 2023).
351. Rosenbluth B, Whitaker DJ, Anne Valle L, Ball B. Integrating strategies for bullying, sexual harassment, and dating violence prevention: the Expect Respect elementary school project. In Espelage DL, Swearer SM, editors. *Bullying in North American Schools*. 2nd edn. New York, NY: Routledge; 2011. pp. 262–72.

## REFERENCES

352. Mabin A. *Effectiveness of a Teen Dating Violence Prevention Program in Middle Schools*. MSc thesis. Syracuse, NY: Syracuse University; 2019.
353. Taylor BG, Mumford EA, Weiwei L, Stein N. *Assessing Different Levels and Dosages of the Shifting Boundaries Intervention to Prevent Youth Dating Violence in New York City Middle Schools: A Randomized Control Trial*. Report submitted to the National Institute of Justice. Chicago, IL: University of Chicago; 2015.
354. Crooks C, Zwicker J, Wells L, Hughes R. Estimating costs and benefits associated with evidence-based violence prevention: four case studies based on the Fourth R program. *SPP Research Papers* 2017;**10**:1–25.
355. Crooks C, Exner-Cortens D, Prunella L, Moore K. *Alaska Fourth R Curriculum Evaluation: Findings Report 2011–2014*. Strategic Prevention Solutions; 2014. URL: <https://safesupportive-learning.ed.gov/resources/alaska-fourth-r-curriculum-evaluation-findings-report-2011-2014> (accessed April 2023).



# Appendix 1 Process documentation

## Protocol amendments

TABLE 32 Protocol amendments

Version	Amendment	Rationale	Date	Submitted to NIHR	Submitted to PROSPERO
1	Minor adjustments to search strings to account for wildcards and free-text terms	Ensure consistency across databases	4 June 2020	5 June 2020	5 June 2020
1.1	Acknowledgement of funding, disclaimer	NIHR request	3 July 2020	3 July 2020	N/A
2.0	<p><b>MAJOR</b>            Addition of new research question (RQ5) relating to network meta-analysis            Reordering of research questions to better reflect flow of analysis            Editing methods for RQ6 to better reflect the size of the evidence base</p> <p><b>MINOR</b>            Minor edits to qualitative synthesis methods (RQ1) to better reflect approaches to evidence            Editing outcomes to clarify inclusion of DRV- and GBV-related behaviours, for example bystander outcomes            Clarifying narrative synthesis of mediation evidence            Editing timelines to reflect contract variation</p>	Contract variation to undertake network meta-analysis; reflecting evidence included	9 November 2021	9 November 2021	9 November 2021

## Protocol deviations and clarifications

TABLE 33 Protocol deviations and clarifications

Deviation or clarification	Rationale
<b>Clarification.</b> Update searches used revised strategies to improve precision, and targeted databases where studies were previously identified.	Update searches were informed by learning from the original round of screening and supplementary author searches, to maximise search effectiveness.
<b>Deviation.</b> We did not search OpenGrey.eu. for grey literature.	OpenGrey closed in early 2021 and is no longer updated. We searched alternative sources to identify grey literature, including checks of publication lists on key websites and Google Scholar.
<b>Deviation.</b> Test sets used 100 abstracts instead of 50.	This was done to improve reliability of screening given the number of records located.
<b>Deviation.</b> We used Covidence instead of EPPI-Reviewer.	Due to the distributed locations of the reviewer team and the challenges of the pandemic, Covidence was a more appropriate software.

continued

TABLE 33 Protocol deviations and clarifications (*continued*)

Deviation or clarification	Rationale
<b>Deviation.</b> For extraction of study characteristics (but not for appraisal), one reviewer extracted each study, with comprehensive (i.e. 100%) quality assurance by two reviewers.	Given the volume of evidence identified, full double extraction of study characteristics (but not of appraisals) was judged to be an inefficient use of resources.
<b>Deviation.</b> In the intervention components analysis, descriptors were applied to all included interventions, but intervention components were only labelled for outcome evaluations.	Because of the volume of evidence identified and systematic differences between outcome evaluations and other forms of evidence included in this review in the depth of intervention description, this was viewed as producing a more reliable and useful analysis for later metaregressions.
<b>Deviation.</b> In the synthesis of intervention theories, we prioritised breadth rather than intervention type in initial coding.	This was done to ensure that our initial coding appropriately represented the range of interventions; also, based on the evidence identified, analysing within type first would have limited our analysis.
<b>Clarification.</b> As described in <a href="#">Chapter 6</a> , synthesis of intervention theories unfolded in a staged approach.	This was done to ensure the relevance and reliability of the synthesis, and to provide a specific rationale and audit trail for the sequence of coding.
<b>Deviation.</b> We used standardised mean differences (Cohen's <i>d</i> ) to meta-analyse findings of knowledge, attitudes and behaviours.	This was appropriate given the majority of estimates included in these outcome domains drew from continuous scales (cf victimisation and perpetration, which were frequently reported as binary variables).
<b>Deviation.</b> We did not sensitivity-analyse meta-analyses by imputed ICC.	Our estimate of 0.05 was considered highly probable on the basis of ICCs reported in included studies.
<b>Deviation.</b> We did not restrict meta-analyses by risk of bias	Given the results of our appraisal, where only one study was rated as low risk of bias, this analysis would have been uninformative.
<b>Deviation.</b> We did not organise synthesis interventions by type, outcome and time category, instead using outcome, type and time category.	This was done to improve intelligibility of meta-analyses but did not affect the number or type of analyses undertaken.
<b>Clarification.</b> Meta-analyses for specific types of victimisation and perpetration outcomes were only undertaken overall rather than by intervention type.	This was done to reduce the number of specious analyses given the overall high number of meta-analyses undertaken.
<b>Deviation.</b> Pairwise meta-analyses including only two studies were estimated using multilevel meta-analysis with a within-study correlation equal to 0.8.	Robust variance estimation grows increasingly unstable as analysis degrees of freedom approaches 1.
<b>Clarification.</b> Synthesis of mediation and moderation evidence did not stratify by time category.	This would have been unhelpful given sparseness of identified evidence; moreover, many analyses were longitudinal, including both short-term and long-term data.
<b>Deviation.</b> Moderation evidence was not stratified by intervention type.	This would have been unhelpful given sparseness of identified evidence.
<b>Deviation.</b> We undertook QCA by outcome category, instead of examining 'joint effectiveness' on DRV and GBV; and did not analyse long-term GBV victimisation or perpetration.	We did not have enough overlap in included studies for this analysis to be possible.
<b>Clarification.</b> We undertook exploratory NMAs by component type only for DRV outcomes.	The quantity of GBV evidence identified was too narrow to make this analysis useful.

TABLE 34 Overview of studies included in the review

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Active interventions							
Acquaintance rape education program	Fay 2006 <sup>198</sup>	Fay 2006 <sup>198</sup>	Fay 2006 <sup>198</sup>		Fay 2006 <sup>198</sup>		North America, USA
Adult-led sexual harassment classes	Sabella 1995 <sup>158</sup>		Sabella 1995 <sup>158</sup>				North America, USA
AHYR strategy		Dozois 2016 <sup>302</sup>					North America, Canada
AAL	Harrington 2019 <sup>216</sup>	Harrington 2019 <sup>216</sup>					North America, USA
Battered Women: Violence Behind Closed Door video	Walther 1996 <sup>183</sup>		Walther 1986 <sup>183</sup>				North America, USA
Benzies and Batchies	de Lijster 2016 <sup>168</sup>		de Lijster 2016 <sup>168</sup>		de Lijster 2016 <sup>168</sup>		Europe, The Netherlands
BVV audio-drama intervention	Cockcroft 2019 <sup>134</sup>	Cockcroft 2019 <sup>134</sup>					Africa, Botswana
BRIGHT single + 5-session boost		Cascardi 2014 <sup>80</sup>					North America, USA
BITB-HSC	Edwards 2019 <sup>197</sup>	Edwards 2019 <sup>197</sup>	Edwards 2019 <sup>197</sup>		Edwards 2019 <sup>197</sup>		North America, USA
BRIGHT single session		Cascardi 2014 <sup>80</sup>					North America, USA
Bystander intervention curriculum	Lee 2018 <sup>87</sup>	Lee 2018 <sup>87</sup>	Lee 2018 <sup>87</sup>				Asia, Taiwan
Bystander-focused violence prevention curriculum	Edwards 2019 <sup>208</sup>	Edwards 2019 <sup>208</sup>					North America, USA
Bystanders		Jalušič 2019 <sup>117</sup>					Europe, Slovenia
C&C	Elias-Lambert 2015 <sup>213</sup>	Elias-Lambert 2015 <sup>213</sup>					North America, USA
CAMPAIGN		Achyut 2011 <sup>78</sup>	Achyut 2011 <sup>78</sup>				Asia, India
CARE face to face		McGinn 2017 <sup>123</sup>					Europe, Ireland
CARE VW		McGinn 2017 <sup>123</sup>					Europe, Ireland
CPVFK		Cameron 2007 <sup>177</sup>					North America, Canada
Change up	Rogers 2018 <sup>127</sup>	Rogers 2018 <sup>127</sup>					Europe, England

continued

TABLE 34 Overview of studies included in the review (continued)

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Chesterfield Relate (Relationship Education Leading Adolescents toward Empowerment) Program	Cramer 2015 <sup>249</sup>	Cramer 2015 <sup>249</sup>					North America, USA
CBIM	Miller 2012 <sup>77</sup>	Miller 2012 <sup>77</sup>	Miller 2012 <sup>77</sup>			Jones 2021 <sup>303</sup>	North America, USA
CBIM	Miller 2020 <sup>180</sup>		Miller 2020 <sup>180</sup>		Miller 2020 <sup>180</sup>	Jones 2021 <sup>303</sup>	North America, USA
CBIM: coach delivery		Jaime 2016 <sup>178</sup>					North America, USA
CAAR		Cameron 2007 <sup>177</sup>					North America, Canada
Community activities	Foshee 1998 <sup>36</sup>		Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>		North America, USA
CBPR project to reduce DRV	Watts 2016 <sup>304</sup>	Watts 2016 <sup>304</sup>					North America, USA
Connect with Respect	Cahill 2019 <sup>132</sup>	Cahill 2019 <sup>132</sup>					Africa, Namibia, South Africa, South Sudan, eSwatini, Tanzania, Zambia and Zimbabwe
Connections: Dating and Emotions curriculum		Gardner 2005 <sup>209</sup>					North America, USA
DAT-E Adolescence	Sanchez-Jimenez 2018 <sup>49</sup>	Sanchez-Jimenez 2018 <sup>49</sup>	Sanchez-Jimenez 2018 <sup>49</sup>				Europe, Spain
Dating Matters	Niolon 2019 <sup>166</sup>		Niolon 2019 <sup>166</sup>		Niolon 2019 <sup>166</sup>	Luo 2020 <sup>305</sup>	North America, USA
Dating violence prevention program	Avery-Leaf 1997 <sup>196</sup>		Avery-Leaf 1997 <sup>196</sup>				North America, USA
Dating violence prevention program	Macgowan 1997 <sup>89</sup>	Macgowan 1997 <sup>89</sup>	Macgowan 1997 <sup>89</sup>		Macgowan 1997 <sup>89</sup>		North America, USA
Dating Violence Prevention Project Curriculum (Avery-Leaf <i>et al.</i> 1997)		Elias-Lambert 2010 <sup>217</sup>					North America, USA
Drama intervention	Brunk 1993 <sup>250</sup>	Brunk 1996 <sup>250</sup>					North America, Canada
DRV advocate delivery		Jaime 2016 <sup>178</sup>					North America, USA
DRV curriculum			Gage 2016 <sup>165</sup>		Gage 2016 <sup>165</sup>		North America, Haiti
DRV intervention	dos Santos 2019 <sup>163</sup>		dos Santos 2019 <sup>163</sup>				South America, Brazil

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
DRV pilot intervention	Filho 2017 <sup>84</sup>		Filho 2017 <sup>84</sup>				South America, Brazil
DRV prevention	Filho 2017 <sup>84</sup>	Filho 2017 <sup>84</sup>	Filho 2017 <sup>84</sup>				South America, Brazil
DRV prevention intervention		Murta 2016 <sup>145</sup>					South America, Brazil
ECPVG	Bando 2019 <sup>184</sup>		Bando 2019 <sup>184</sup>		Bando 2019 <sup>184</sup>		North America, El Salvador
Educational CD-ROM	Yom 2005 <sup>170</sup>		Yom 2005 <sup>170</sup>				Asia, South Korea
Ending violence	Jaycox 2006 <sup>185</sup>	Jaycox 2006 <sup>86</sup>	Jaycox 2006 <sup>185</sup>		Jaycox 2006 <sup>185</sup>		North America, USA
Evaluation of existing sexual harassment policy in three participating districts	Lerner 1999 <sup>306</sup>	Lerner 1999 <sup>306</sup>					North America, USA
Expect Respect	Roberts 2009 <sup>95</sup>	Roberts 2009 <sup>95</sup>	Roberts 2009 <sup>95</sup>		Roberts 2009 <sup>95</sup>		North America, USA
Expect Respect	Rosenbluth 2004 <sup>159</sup>		Rosenbluth 2004 <sup>159</sup>		Rosenbluth 2004 <sup>159</sup>		North America, USA
Expect Respect		Ball 2009 <sup>218</sup>					North America, USA
Expect Respect		Ball 2015 <sup>223</sup>					North America, USA
Expect Respect and MOST Clubs		Noonan 2009 <sup>307</sup>					North America, USA
FAIR Education Act		Garces-Foley 2017 <sup>259</sup>					North America, USA
Familias En Nuestra Escuela (Families in our School)	Enriquez 2012 <sup>220</sup>	Enriquez 2012 <sup>220</sup>					North America, USA
Filles et Garçons, en route pour l'Égalité	Hale 2012 <sup>105</sup>	Hale 2012 <sup>105</sup>					Europe, UK, France, Spain and Malta
FSACC		Cameron 2007 <sup>177</sup>					North America, Canada
FSP-VGMS Project		Sarr 2019 <sup>140</sup>					Africa, Cameroon, Senegal and Togo
GEA	ICRW 2017 <sup>85</sup>	ICRW 2017 <sup>85</sup>					Asia, Bangladesh
GEA + CAMPAIGN	Achyut 2011 <sup>78</sup>	Achyut 2011 <sup>78</sup>	Achyut 2011 <sup>78</sup>				Asia, India
GEMS	ICRW 2017 <sup>85</sup>	ICRW 2017 <sup>85</sup>	ICRW 2017 <sup>85</sup>		ICRW 2017 <sup>85</sup>		Asia, India, Vietnam
GPE		Ollis 2017 <sup>152</sup>					Australasia, Australia

continued

TABLE 34 Overview of studies included in the review (continued)

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
GBV Prevention Education Programme		Maphosa 2018 <sup>136</sup>					Africa, Zimbabwe
GSDP	Jordan 2018 <sup>147</sup>	Jordan 2018 <sup>147</sup>					Australasia, New Zealand
Go for gold relationship education program	Grimm 2011 <sup>210</sup>	Grimm 2011 <sup>210</sup>					North America, USA
GGI		Schwandt 2016 <sup>141</sup>					Africa, Botswana, Malawi, and Mozambique
Green Acres High	Sorbring 2015 <sup>129</sup>	Sorbring 2015 <sup>129</sup>					Europe, Sweden
Green Dot	Coker 2017 <sup>82</sup>	Coker 2017 <sup>82</sup>	Coker 2017 <sup>82</sup>	Coker 2017 <sup>82</sup>	Coker 2017 <sup>82</sup>	Bush 2018 <sup>308</sup>	North America, USA
Health belief model educational program	Garmaroudi 2016 <sup>172</sup>		Garmaroudi 2016 <sup>172</sup>				Asia, Iran
Healthy relationships		Farrelly 2021 <sup>102</sup>					Europe, England
Healthy relationships drama program	Bell 2006 <sup>110</sup>	Bell 2006 <sup>110</sup>					Europe, England
Healthy Relationships for Youth/Rural Youth Education Project		Tutty 2011 <sup>179</sup>					North America, Canada
HIV/STD risk-reduction intervention	Jemmott 2018 <sup>176</sup>		Jemmott 2018 <sup>176</sup>	Jemmott 2018 <sup>176</sup>	Jemmott 2018 <sup>176</sup>		Africa, South Africa
HRE curriculum		Mills 1998 <sup>150</sup>					Australasia, Australia
IMPower	Decker 2018 <sup>174</sup>		Decker 2018 <sup>174</sup>				Africa, Malawi
IMPower/50 : 50	Baiocchi 2017 <sup>160</sup>		Baiocchi 2017 <sup>160</sup>				Africa, Kenya
Interaction curriculum			Taylor 2010 <sup>200</sup>		Taylor 2010 <sup>157</sup>		North America, USA
ITP intervention	Fernandez-Gonzalez 2020 <sup>169</sup>		Fernandez-Gonzalez 2020 <sup>169</sup>				Europe, Spain
It's your game ... keep it real (IYG)	Peskin 2014 <sup>189</sup>		Peskin 2014 <sup>189</sup>		Peskin 2014 <sup>189</sup>		North America, USA
JOVEN	Gonzalez-Guarda 2015 <sup>193</sup>		Gonzalez-Guarda 2015 <sup>193</sup>		Gonzalez-Guarda 2015 <sup>193</sup>		North America, USA
Katie Brown Educational Program	Joppa 2016 <sup>186</sup>		Joppa 2016 <sup>186</sup>	Joppa 2016 <sup>186</sup>	Joppa 2016 <sup>186</sup>		North America, USA
La Mascara del Amor		Tello 2013 <sup>113</sup>					Europe, Spain

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
La Máscara del Amor		Genovés 2009 <sup>116</sup>					Europe, Spain
La Máscara del Amor		Hale 2012 <sup>105</sup>					Europe, UK, France, Spain and Malta
Law and justice curriculum			Taylor 2010 <sup>200</sup>		Taylor 2010 <sup>200</sup>		North America, USA
Lei Maria da Penha vai às escolas		Kelly de Albuquerque 2020 <sup>104</sup>					South America, Brazil
Lights4Violence	Perez-Marco 2020 <sup>125</sup>	Perez-Marco 2020 <sup>125</sup>					Europe, Spain
Literary instruction	Malo-Juvera 2014 <sup>195</sup>		Malo-Juvera 2014 <sup>195</sup>		Malo-Juvera 2014 <sup>195</sup>		North America, USA
LDH	Hertel 2020 <sup>262</sup>	Hertel 2020 <sup>262</sup>					North America, USA
Love hurts		McElwee 2020 <sup>122</sup>					Europe, England
Making Waves/Vague par vague		Tutty 2011 <sup>179</sup>					North America, Canada
MW/Vpv		Cameron 2007 <sup>177</sup>					North America, Canada
Me and You	Peskin 2019 <sup>188</sup>		Peskin 2019 <sup>188</sup>		Peskin 2019 <sup>188</sup>		North America, USA
Media aware	Scull 2018 <sup>98</sup>	Scull 2018 <sup>98</sup>	Scull 2018 <sup>98</sup>				North America, USA
Media aware	Scull 2021 <sup>97</sup>	Scull 2021 <sup>97</sup>	Scull 2021 <sup>97</sup>				North America, USA
Media intervention	Jones 2010 <sup>309</sup>	Jones 2010 <sup>309</sup>					North America, USA
MEMO4LOVE	Racionero-Plaza 2020 <sup>126</sup>	Racionero-Plaza 2020 <sup>126</sup>					Europe, Spain
Mentors in violence		Bruno 2020 <sup>112</sup>					Europe, Sweden
MVP Program		Beardall 2008 <sup>266</sup>					North America, USA
My voice, my choice	Rowe 2015 <sup>96</sup>	Rowe 2015 <sup>96</sup>	Rowe 2015 <sup>96</sup>		Rowe 2015 <sup>96</sup>		North America, USA
NDVP		AVA 2013 <sup>109</sup>					Europe, England
OTR		AVA 2013 <sup>109</sup>					Europe, England
Papo reto		Oliveira 2016 <sup>146</sup>					South America, Brazil

continued

TABLE 34 Overview of studies included in the review (continued)

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Parallel retreats program		Madsen 1994 <sup>233</sup>					North America, Canada
Pass it on	Winegust 2015 <sup>310</sup>	Winegust 2015 <sup>310</sup>					North America, Canada
Peer educator DRV drama workshops	Walton 2007 <sup>130</sup>	Walton 2007 <sup>130</sup>					Europe, England
PLT		Dos Santos 2019 <sup>142</sup>					South America, Brazil
Peer-led	Sabella 1995 <sup>158</sup>		Sabella 1995 <sup>158</sup>				North America, USA
Popular opinion leader groups to reduce LGBTQQ aggression	Singh 2013 <sup>207</sup>	Singh 2013 <sup>207</sup>					North America, USA
Power Up, Speak Out!	Genereux 2020 <sup>236</sup>	Genereux 2020 <sup>236</sup>					North America, USA
PR:EPARe game	Arnab 2012 <sup>79</sup>	Arnab 2012 <sup>79</sup>	Arnab 2012 <sup>79</sup>				Europe, England
PR:EPARe game vs. waitlist		Arnab 2012 <sup>79</sup>					Europe, England
Practitioner Program	Muck 2018 <sup>94</sup>	Muck 2018 <sup>94</sup>	Muck 2018 <sup>94</sup>		Muck 2018 <sup>94</sup>		Europe, Germany
Practitioner Program		Muck 2018 <sup>94</sup>					Europe, Germany
Precede-Proceed Model DRV intervention vs. no intervention	Ekhtiari 2013 <sup>171</sup>	Ekhtiari 2013 <sup>171</sup>	Ekhtiari 2013 <sup>171</sup>				Asia, Iran
PREPARE	Mathews 2016 <sup>90</sup>	Mathews 2016 <sup>90</sup>	Mathews 2016 <sup>90</sup>		Mathews 2016 <sup>90</sup>		Africa, South Africa
PREPARE		Mathews 2016 <sup>90</sup>					Africa, South Africa
PRA		Jorba 2012 <sup>111</sup>					Europe, Spain
Preventative DRV intervention		Matos 2006 <sup>120</sup>					Europe, Portugal
PREVIO			Munoz-Rivas 2019 <sup>164</sup>		Munoz-Rivas 2019 <sup>164</sup>		Europe, Spain
Project Connect: A Coordinated Public Health Initiative to Prevent Violence Against Women		Raible 2017 <sup>201</sup>					North America, USA
Project Respect	Meiksin 2020 <sup>35</sup>	Meiksin 2020 <sup>35</sup>	Meiksin 2020 <sup>35</sup>			Meiksin 2020 <sup>35</sup>	Europe, England
Promotores educativos		García Escobar 2020 <sup>143</sup>					South America, Peru
Prosocial video game	Boduszek 2019 <sup>161</sup>		Boduszek 2019 <sup>161</sup>				North America, Barbados
POY clubs		Chipeta 2019 <sup>133</sup>					Africa, Malawi



Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
R4Respect	Struthers 2019 <sup>154</sup>	Struthers 2019 <sup>154</sup>					Australasia, Australia
Rape education program	Kershner 1996 <sup>194</sup>		Kershner 1996 <sup>194</sup>		Kershner 1996 <sup>194</sup>		North America, USA
RSIS	Payne 2018 <sup>260</sup>	Payne 2018 <sup>260</sup>					North America, USA
Relaciones Romaticas Constructivas		Gonzalez 2014 <sup>144</sup>					South America, Colombia
Relationships without Fear	Hale 2012 <sup>105</sup>	Hale 2012 <sup>105</sup>					Europe, UK, France, Spain and Malta
RWV		Fawson 2016 <sup>214</sup>					North America, USA
Respect	Baker 2014 <sup>311</sup>	Baker 2014 <sup>148</sup>					North America, Hawaii
Respectful RE	Keddie 2020 <sup>103</sup>	Keddie 2020 <sup>103</sup>					Australasia, Australia
Respectful RE	Ollis 2011 <sup>151</sup>	Ollis 2011 <sup>151</sup>					Australasia, Australia
RREIS	Kearney 2016 <sup>148</sup>	Kearney 2016 <sup>148</sup>					Australasia, Australia
Respond		AVA 2013 <sup>109</sup>					Europe, England
Riot Youth drama intervention	Wernick 2013 <sup>206</sup>	Wernick 2013 <sup>206</sup>					North America, USA
Safe Dates	Foshee 1998 <sup>36</sup>		Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>		North America, USA
Safe Dates			Niolon 2019 <sup>166</sup>		Niolon 2019 <sup>166</sup>	Luo 2020 <sup>305</sup>	North America, USA
Safe Dates		Cutbush 2017 <sup>312</sup>					North America, USA
Safe Schools	Shevlin 2020 <sup>153</sup>	Shevlin 2020 <sup>153</sup>					Australasia, Australia
Safe Schools Coalition Australia	Louden 2016 <sup>149</sup>	Louden 2016 <sup>149</sup>					Australasia, Australia
Safe schools program for gay and lesbian students		Ouellett 1998 <sup>222</sup>					North America, USA
Safe schools program for gay and lesbian students	Watson 2012 <sup>202</sup>	Watson 2012 <sup>202</sup>					North America, USA
SAISIR	Chamberland 2014 <sup>167</sup>		Chamberland 2014 <sup>167</sup>		Chamberland 2014 <sup>167</sup>		Europe, France
SWOVA Respectful Relationship program (R + R)		Tutty 2011 <sup>179</sup>					North America, Canada

continued

TABLE 34 Overview of studies included in the review (continued)

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
SHARP	Miller 2015 <sup>187</sup>	Miller 2015 <sup>93</sup>	Miller 2015 <sup>187</sup>		Miller 2015 <sup>187</sup>		North America, USA
SVPA of 2009		Hall 2016 <sup>313</sup>					North America, USA
School-based youth-driven teen dating violence prevention project	Kervin 2010 <sup>224</sup>	Kervin 2010 <sup>224</sup>					North America, USA
Scientist-Practitioner Program	Muck 2018 <sup>94</sup>	Muck 2018 <sup>94</sup>	Muck 2018 <sup>94</sup>		Muck 2018 <sup>94</sup>		Europe, Germany
Second Step	Espelage 2013 <sup>192</sup>		Espelage 2013 <sup>192</sup>	Espelage 2013 <sup>192</sup>			North America, USA
Secondary Prevention project (modified version of the Massachusetts Teen Dating Violence Prevention and Intervention Program)	Silverman 2000 <sup>99</sup>	Silverman 2000 <sup>99</sup>	Silverman 2000 <sup>99</sup>		Silverman 2000 <sup>99</sup>		North America, USA
Self-led	Sabella 1995 <sup>158</sup>		Sabella 1995 <sup>158</sup>				North America, USA
Sexual coercion prevention program	Pacifici 2001 <sup>191</sup>		Pacifici 2001 <sup>191</sup>		Pacifici 2001 <sup>191</sup>		North America, USA
Sexual harassment and child abuse policy	Wilson 2005 <sup>314</sup>	Wilson 2005 <sup>40</sup>					North America, USA
Sexual harassment intervention	Durand 1997 <sup>83</sup>	Durand 1997 <sup>83</sup>	Durand 1997 <sup>83</sup>		Durand 1997 <sup>83</sup>		North America, USA
Sexual health programme for YP with developmental disabilities		Murray 2019 <sup>315</sup>					North America, Canada
Sexuality education program		Makleff 2019 <sup>205</sup>					North America, Mexico
Shifting boundaries	Taylor 2017 <sup>156</sup>		Taylor 2017 <sup>156</sup>		Taylor 2017 <sup>156</sup>		North America, USA
Shifting boundaries: Building + classroom	Taylor 2011 <sup>100</sup>	Taylor 2011 <sup>100</sup>	Taylor 2011 <sup>100</sup>		Taylor 2011 <sup>100</sup>		North America, USA
Shifting boundaries: Building + classroom vs. building only vs. classroom only vs. control group	Taylor 2011 <sup>100</sup>	Taylor 2011 <sup>100</sup>					North America, USA
Shifting boundaries: Building only	Taylor 2011 <sup>100</sup>	Taylor 2011 <sup>100</sup>	Taylor 2011 <sup>100</sup>		Taylor 2011 <sup>100</sup>		North America, USA
Shifting boundaries: Classroom only	Taylor 2011 <sup>100</sup>	Taylor 2011 <sup>100</sup>	Taylor 2011 <sup>100</sup>		Taylor 2011 <sup>100</sup>		North America, USA
Skhokho	Jewkes 2019 <sup>155</sup>		Jewkes 2019 <sup>155</sup>		Jewkes 2019 <sup>155</sup>		Africa, South Africa

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Skhokho + caregivers	Jewkes 2019 <sup>155</sup>		Jewkes 2019 <sup>155</sup>		Jewkes 2019 <sup>155</sup>		Africa, South Africa
SKILLZ Street		Merrill 2018 <sup>137</sup>					Africa, South Africa
Sources	Yoder 2020 <sup>101</sup>	Yoder 2020 <sup>101</sup>					North America, USA
Southall Black Sisters		AVA 2013 <sup>109</sup>					Europe, England
Start Strong Bronx (adaptation of The Fourth R)	Cissner 2014 <sup>81</sup>	Cissner 2014 <sup>81</sup>	Cissner 2014 <sup>81</sup>		Cissner 2014 <sup>81</sup>	Cissner 2014 <sup>81</sup>	North America, USA
Stay in love+		Kempes 2010 <sup>118</sup>					Europe, Netherlands
Stay strong/Safe Dates		Gibbs 2016 <sup>263</sup>					North America, USA
TakeCARE	Jouriles 2019 <sup>181</sup>		Jouriles 2019 <sup>181</sup>		Jouriles 2019 <sup>181</sup>		North America, USA
TakeCARE	Sargent 2017 <sup>38</sup>		Sargent 2017 <sup>38</sup>		Sargent 2017 <sup>38</sup>		
Teen Choices	Levesque 2016 <sup>182</sup>	Levesque 2016 <sup>182</sup>	Levesque 2016 <sup>182</sup>		Levesque 2016 <sup>182</sup>		North America, USA
Teen VIP	Miller 1998 <sup>162</sup>		Miller 1998 <sup>162</sup>				North America, USA
Tender		AVA 2013 <sup>109</sup>					Europe, England
Tender drama and art intervention		DMSS Research & Consultancy 2012 <sup>115</sup>					Europe, UK
Texas Team's Teen Dating Violence Awareness and Prevention Toolkit	Busch-Armendariz 2008 <sup>235</sup>	Busch-Armendariz 2008 <sup>235</sup>					North America, USA
Text message campaign	Guillot-Wright 2018 <sup>215</sup>	Guillot-Wright 2018 <sup>215</sup>					North America, USA
The 'Respect' project	Henderson 2002 <sup>106</sup>	Henderson 2002 <sup>106</sup>					Europe, UK
The 5 W's approach to Bullying	Merrell 2004 <sup>92</sup>	Merrell 2004 <sup>92</sup>	Merrell 2004 <sup>92</sup>		Merrell 2004 <sup>92</sup>		North America, USA
The Father's Day Breakfast		Henshaw 2016 <sup>316</sup>					North America, Canada
The Fourth R		Chiodo 2017 <sup>317</sup>					North America, Canada
The Fourth R		Crooks 2013 <sup>267</sup>					North America, Canada
The Fourth R		Crooks 2018 <sup>318</sup> , Crooks 2014 <sup>355</sup>					North America, Canada
The Fourth R		Dunlop 2018 <sup>319</sup>					North America, Canada

continued

TABLE 34 Overview of studies included in the review (continued)

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
The Fourth R		Exner-Cortens 2020 <sup>320</sup>					North America, Canada
The Fourth R	Gale 2011 <sup>244</sup>	Gale 2011 <sup>244</sup>					North America, USA
The Fourth R		Tutty 2011 <sup>179</sup>					North America, Canada
The Fourth R: Skills for Youth Relationship	Wolfe 2009 <sup>190</sup>	Wolfe 2009 <sup>190</sup>	Wolfe 2009 <sup>190</sup>		Wolfe 2009 <sup>190</sup>	Wolfe 2009 <sup>190</sup>	North America, USA
The GENER@T Program		Mateos Inchaurredo 2020 <sup>119</sup>					Europe, Spain
The Good School Toolkit	Devries 2017 <sup>175</sup>		Devries 2017 <sup>175</sup>		Devries 2017 <sup>175</sup>		Africa, Uganda
The HRP for LGBTQ + Youth	Lapointe 2018 <sup>256</sup>	Lapointe 2018 <sup>256</sup>					North America, Canada
The MVP program	Williams 2017 <sup>131</sup>	Williams 2017 <sup>131</sup>					Europe, Scotland
The Modified Integrated Thematic Instructional Model for Sexual Harassment Prevention		Cheney 1998 <sup>264</sup>					North America, USA
The SDTP		Turner 2006 <sup>212</sup>					North America, USA
The Safe Schools Program	Knowles 1997 <sup>232</sup>	Knowles 1997 <sup>232</sup>					North America, USA
The SwV Programme including GBV component	UNICEF 2016 <sup>107</sup>	UNICEF 2016 <sup>107</sup>					Europe, Serbia
THE SKILLZ STREET PLUS PROGRAMME		Cooper 2017 <sup>135</sup>					Africa, South Africa
The TANESA guardian programme		Mgalla 1998 <sup>138</sup>					Africa, Tanzania
The Tender Healthy Relationship project		Sanders-McDonagh 2015 <sup>128</sup>					Europe, England
The TRUST project		CRG Research 2016 <sup>114</sup>					Europe, UK
Theatre intervention	Belknap 2013 <sup>219</sup>	Belknap 2013 <sup>219</sup>					North America, USA
Train the trainer sexual violence prevention program implemented by the SATC	Weingarten 2018 <sup>108</sup>	Weingarten 2018 <sup>108</sup>					North America, Hawaii

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Twilight book		Lynch 2014 <sup>88</sup>	Lynch 2014 <sup>88</sup>				Europe, England
Victim Support Cornwall (the SAFE Project)		AVA 2013 <sup>109</sup>					Europe, England
Web-based DV prevention program	Jung 2013 <sup>173</sup>		Jung 2013 <sup>173</sup>				Asia, South Korea
WOMANKIND WSA		Maxwell 2010 <sup>121</sup>					Europe, England and Wales
WOMEN's group	Diegel 1999 <sup>203</sup>	Diegel 1999 <sup>134</sup>					North America, USA
You-Me-Us	Coyle 2019 <sup>199</sup>		Coyle 2019 <sup>199</sup>				North America, USA
Young men initiative		Namy 2015 <sup>124</sup>					Europe, Bosnia and Herzegovina, Croatia, Serbia
YES-HR	Eisman 2019 <sup>221</sup>	Eisman 2019 <sup>221</sup>					North America, USA
Zero Tolerance School Alliance vs. no intervention		Nicholson 2018 <sup>139</sup>					Africa, South Africa
Control interventions							
Active control			Baiocchi 2017 <sup>160</sup>				Africa, Kenya
Active control			Fernandez-Gonzalez 2020 <sup>169</sup>				Europe, Spain
Active control			Gage 2016 <sup>165</sup>		Gage 2016 <sup>165</sup>		North America, Haiti
Active control			Jemmott 2018 <sup>176</sup>	Jemmott 2018 <sup>176</sup>	Jemmott 2018 <sup>176</sup>		Africa, South Africa
Active control			Jouriles 2019 <sup>181</sup>		Jouriles 2019 <sup>181</sup>		North America, USA
Active control		Lee 2018 <sup>87</sup>	Lee 2018 <sup>87</sup>				Asia, Taiwan
Active control			Levesque 2016 <sup>182</sup>		Levesque 2016 <sup>182</sup>		North America, USA
Active control			Miller 1998 <sup>162</sup>				North America, USA
Active control			Walther 1986 <sup>183</sup>				North America, USA
Active control			Yom 2005 <sup>170</sup>				Asia, South Korea
Active control/waitlist		Merrell 2004 <sup>92</sup>	Merrell 2004 <sup>92</sup>		Merrell 2004 <sup>92</sup>		North America, USA

continued

TABLE 34 Overview of studies included in the review (continued)

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Control			Coyle 2019 <sup>199</sup>				North America, USA
Control		Genovés 2009 <sup>116</sup>					Europe, Spain
Control		Gonzalez 2014 <sup>144</sup>					South America, Colombia
Control		Muck 2018 <sup>94</sup>					Europe, Germany
Community activities	Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>	Foshee 1998 <sup>36</sup>		North America, USA
No intervention		Achyut 2011 <sup>78</sup>	Achyut 2011 <sup>78</sup>				Asia, India
No intervention		ICRW 2017_India <sup>85</sup>	ICRW 2017_India <sup>85</sup>		ICRW 2017_India <sup>85</sup>		Asia, India, Vietnam
No intervention		ICRW 2017_Vietnam <sup>85</sup>	ICRW 2017_Vietnam <sup>85</sup>		ICRW 2017_Vietnam <sup>85</sup>		Asia, India, Vietnam
No intervention			Avery-Leaf 1997 <sup>196</sup>				North America, USA
No intervention			Boduszek 2019 <sup>161</sup>				North America, Barbados
No intervention		Cissner 2014 <sup>81</sup>	Cissner 2014 <sup>81</sup>		Cissner 2014 <sup>81</sup>		North America, USA
No intervention			Coker 2017 <sup>82</sup>	Coker 2017 <sup>82</sup>	Coker 2017 <sup>82</sup>		North America, USA
No intervention			dos Santos 2019 <sup>163</sup>				South America, Brazil
No intervention			Edwards 2019 <sup>197</sup>		Edwards 2019 <sup>197</sup>		North America, USA
No intervention		Ekhtiari 2013 <sup>171</sup>	Ekhtiari 2013 <sup>171</sup>				Asia, Iran
No intervention		Fay 2006 <sup>198</sup>	Fay 2006 <sup>198</sup>		Fay 2006 <sup>198</sup>		North America, USA
No intervention			Filho 2017 <sup>84</sup>				South America, Brazil
No intervention			Jewkes 2019 <sup>155</sup>		Jewkes 2019 <sup>155</sup>		Africa, South Africa
No intervention			Jung 2013 <sup>173</sup>				Asia, South Korea
No intervention		Mathews 2016 <sup>90</sup>	Mathews 2016 <sup>90</sup>		Mathews 2016 <sup>90</sup>		Africa, South Africa
No intervention			Miller 2012 <sup>77</sup>				North America, USA
No intervention			Miller 2020 <sup>180</sup>		Miller 2020 <sup>180</sup>		North America, USA
No intervention		Munoz-Rivas 2019 <sup>164</sup>	Munoz-Rivas 2019 <sup>164</sup>		Munoz-Rivas 2019 <sup>164</sup>		Europe, Spain

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
No intervention			Rosenbluth 2004 <sup>159</sup>		Rosenbluth 2004 <sup>159</sup>		North America, USA
No intervention		Rowe 2015 <sup>96</sup>	Rowe 2015 <sup>96</sup>		Rowe 2015 <sup>96</sup>		North America, USA
No intervention			Taylor 2010 <sup>200</sup>		Taylor 2010 <sup>200</sup>		North America, USA
No intervention		Achyut 2011 <sup>78</sup>					Asia, India
No intervention		ICRW 2017 <sup>85</sup>					Asia, India, Vietnam
No intervention		Bruno 2020 <sup>112</sup>					Europe, Sweden
No intervention		Cameron 2007 <sup>177</sup>					North America, Canada
No intervention		Cascardi 2014 <sup>80</sup>					North America, USA
No intervention		Cissner 2014 <sup>81</sup>					North America, USA
No intervention		Coker 2017 <sup>82</sup>					North America, USA
No intervention	Gardner 2005 <sup>209</sup>	Gardner 2005 <sup>209</sup>					North America, USA
No intervention		Mathews 2016 <sup>90</sup>					Africa, South Africa
No intervention		Mgalla 1998 <sup>138</sup>					Africa, Tanzania
No intervention		Nicholson 2018 <sup>139</sup>					Africa, South Africa
No intervention		Rowe 2015 <sup>96</sup>					North America, USA
No intervention	Schwandt 2016 <sup>141</sup>	Schwandt 2016 <sup>141</sup>					Africa, Botswana, Malawi, and Mozambique
No intervention		Tello 2013 <sup>113</sup>					Europe, Spain
Usual practice			Bando 2019 <sup>184</sup>		Bando 2019 <sup>184</sup>		North America, El Salvador
Usual practice			Decker 2018 <sup>174</sup>				Africa, Malawi
Usual practice		Durand 1997 <sup>83</sup>	Durand 1997 <sup>83</sup>		Durand 1997 <sup>83</sup>		North America, USA
Usual practice		Jaycox 2006 <sup>185</sup>	Jaycox 2006 <sup>185</sup>		Jaycox 2006 <sup>185</sup>		North America, USA
Usual practice			Joppa 2016 <sup>186</sup>	Joppa 2016 <sup>186</sup>	Joppa 2016 <sup>186</sup>		North America, USA
Usual practice		Meiksin 2020 <sup>35</sup>	Meiksin 2020 <sup>35</sup>	Meiksin 2020 <sup>35</sup>	Meiksin 2020 <sup>35</sup>	Meiksin 2020 <sup>35</sup>	Europe, England

continued

TABLE 34 Overview of studies included in the review (continued)

Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Usual practice		Miller 2015 <sup>187</sup>	Miller 2015 <sup>187</sup>		Miller 2015 <sup>187</sup>		North America, USA
Usual practice			Peskin 2014 <sup>189</sup>		Peskin 2014 <sup>189</sup>		North America, USA
Usual practice			Peskin 2019 <sup>188</sup>		Peskin 2019 <sup>188</sup>		North America, USA
Usual practice		Scull 2018 <sup>98</sup>	Scull 2018 <sup>98</sup>				North America, USA
Usual practice		Silverman 2000 <sup>99</sup>	Silverman 2000 <sup>99</sup>		Silverman 2000 <sup>99</sup>		North America, USA
Usual practice		Wolfe 2009 <sup>190</sup>	Wolfe 2009 <sup>190</sup>		Wolfe 2009 <sup>190</sup>		North America, USA
Usual practice		Baker 2014 <sup>311</sup>					North America, Hawaii
Usual practice		Durand 1997 <sup>83</sup>					North America, USA
Usual practice		Meiksin 2020 <sup>91</sup>					Europe, England
Usual practice		Miller 2015 <sup>93</sup>					North America, USA
Usual practice		Scull 2018 <sup>98</sup>					North America, USA
Usual practice		Jaycox 2006 <sup>86</sup>					North America, USA
Waitlist		Filho 2017 <sup>84</sup>					South America, Brazil
Waitlist		Macgowan 1997 <sup>89</sup>					North America, USA
Waitlist		Arnab 2012 <sup>79</sup>	Arnab 2012 <sup>79</sup>				Europe, England
Waitlist			Chamberland 2014 <sup>167</sup>		Chamberland 2014 <sup>167</sup>		Europe, France
Waitlist			de Lijster 2016 <sup>168</sup>		de Lijster 2016 <sup>168</sup>		Europe, The Netherlands
Waitlist			Devries 2017 <sup>175</sup>		Devries 2017 <sup>175</sup>		Africa, Uganda
Waitlist		Filho 2017 <sup>84</sup>	Filho 2017 <sup>84</sup>				South America, Brazil
Waitlist			Garmaroudi 2016 <sup>172</sup>				Asia, Iran
Waitlist			Gonzalez-Guarda 2015 <sup>193</sup>		Gonzalez-Guarda 2015 <sup>193</sup>		North America, USA
Waitlist			Kershner 1996 <sup>194</sup>		Kershner 1996 <sup>194</sup>		North America, USA



Name of intervention	Theory (author, year)	Process evaluation (author, date)	Outcome evaluation (author, date)	Mediation	Moderation	Economic outcomes (author, date)	Location
Waitlist	Lynch 2014 <sup>88</sup>	Lynch 2014 <sup>88</sup>	Lynch 2014 <sup>88</sup>				Europe, England
Waitlist		Macgowan 1997 <sup>89</sup>	Macgowan 1997 <sup>89</sup>		Macgowan 1997 <sup>89</sup>		North America, USA
Waitlist			Malo-Juvera 2014 <sup>195</sup>		Malo-Juvera 2014 <sup>195</sup>		North America, USA
Waitlist		Muck 2018 <sup>94</sup>	Muck 2018 <sup>94</sup>		Muck 2018 <sup>94</sup>		Europe, Germany
Waitlist			Pacifici 2001 <sup>191</sup>		Pacifici 2001 <sup>191</sup>		North America, USA
Waitlist		Roberts 2009 <sup>95</sup>	Roberts 2009 <sup>95</sup>		Roberts 2009 <sup>95</sup>		North America, USA
Waitlist		Sanchez-Jimenez 2018 <sup>49</sup>	Sanchez-Jimenez 2018 <sup>49</sup>				Europe, Spain
Waitlist			Scull 2021 <sup>97</sup>				North America, USA
Waitlist		Lynch 2014 <sup>88</sup>					Europe, England
Waitlist		Roberts 2009 <sup>95</sup>					North America, USA
Waitlist		Sanchez-Jimenez 2018 <sup>49</sup>					Europe, Spain
Waitlist		Scull 2021 <sup>97</sup>					North America, USA
Waitlist		Yoder 2020 <sup>101</sup>					North America, USA
Waitlist/Stories of us			Espelage 2013 <sup>192</sup>	Espelage 2013 <sup>192</sup>			North America, USA

BRIGHT, Building Relationships in Greater Harmony Together; CAAR, Coalition against abuse in relationships; CBPR, Community-based participatory research; CPVFK, Centre de prevention de la violence familiale de kent; ECPVG, Educación Comunitaria para la Prevención de la Violencia de Género; FSACC, Fredericton sexual assault crisis centre; FSP-VGMS, Fonds de solidarité prioritaire: violences de genre en milieu scolaire; GGI, Go Girls! Initiative; GPE, gender and positive education program; GSDP, Girls' Self Defence Project; HIV, human immunodeficiency virus; HRE, human relationships education; HRP, healthy relationships program; ITP, incremental theory of personality; LDH, love doesn't hurt; MOST, men of strength; MVP, mentors in violence prevention; MW, making waves; NDVP, Nottingham Domestic Violence Project; OTR, off the record; PLT, peer leader training; PRA, Prevencio de Relaciones Abusives; POY, protect our youth; RSIS, reduction of stigma in schools; RWV, relationships without violence; SATC, Sex Abuse Treatment Center; SDTP, safe dating theater project; STD, sexually transmitted disease; SVPA, School Violence Prevention Act; SWOVA, Saltspring Women Opposed to Violence and Abuse; SwV, school without violence; Vpv, vague par vague; WOMEN, Working On Meaningful Relationships Now; VW, virtual world; YES-HR, youth empowerment solutions for healthy relationships.

TABLE 35 Full list of included publications

Lead study	Linked papers
Achyut 2011 <sup>78</sup>	Achyut 2011; <sup>78</sup> Achyut 2015 <sup>321</sup>
Achyut 2016 <sup>231</sup>	Achyut 2016; <sup>231</sup> ICRW 2017 <sup>85</sup>
Ainoa 2020 <sup>119</sup>	Ainoa 2020 <sup>119</sup>
Arnab 2012 <sup>79</sup>	Arnab 2012; <sup>322</sup> Arnab 2013; <sup>272</sup> Brown 2012; <sup>322</sup> Clarke 2012 <sup>323</sup>
AVA 2013 <sup>109</sup>	AVA 2013 <sup>109</sup>
Avery-Leaf 1997 <sup>196</sup>	Avery-Leaf 1997 <sup>196</sup>
Baiocchi 2017 <sup>160</sup>	Baiocchi 2017 <sup>160</sup>
Baker 2014 <sup>311</sup>	Baker 2014 <sup>311</sup>
Ball 2009 <sup>218</sup>	Ball 2009 <sup>218</sup>
Ball 2015 <sup>223</sup>	Ball 2015 <sup>223</sup>
Bando 2019 <sup>184</sup>	Bando 2018; <sup>324</sup> Bando 2019 <sup>184</sup>
Beardall 2008 <sup>266</sup>	Beardall 2008 <sup>266</sup>
Belknap 2013 <sup>219</sup>	Belknap 2013 <sup>219</sup>
Bell 2006 <sup>110</sup>	Bell 2006 <sup>110</sup>
Boduszek 2019 <sup>161</sup>	Boduszek 2019 <sup>161</sup>
Brunk 1993 <sup>250</sup>	Brunk 1993 <sup>250</sup>
Bruno 2020 <sup>112</sup>	Bruno 2020; <sup>112</sup> Gottzén 2021 <sup>325</sup>
Busch-Armendariz 2008 <sup>235</sup>	Busch-Armendariz 2008 <sup>235</sup>
Cahill 2019 <sup>132</sup>	Cahill 2019 <sup>132</sup>
Cameron 2007 <sup>177</sup>	Cameron 2007 <sup>177</sup>
Cascardi 2014 <sup>80</sup>	Cascardi 2014 <sup>80</sup>
Chamberland 2014 <sup>167</sup>	Chamberland 2014 <sup>167</sup>
Cheney 1998 <sup>264</sup>	Cheney 1998 <sup>264</sup>
Chiodo 2017 <sup>317</sup>	Chiodo 2017 <sup>317</sup>
Chipeta 2019 <sup>133</sup>	Chipeta 2019 <sup>133</sup>
Cissner 2014 <sup>81</sup>	Cissner 2014 <sup>81</sup>
Cockcroft 2019 <sup>134</sup>	Cockcroft 2019 <sup>134</sup>
Coker 2017 <sup>82</sup>	Bush 2018; <sup>308</sup> Coker 2017; <sup>326</sup> Coker 2019; <sup>238</sup> Coker 2020; <sup>327</sup> Coker 2020; <sup>255</sup> Coker 2021; <sup>328</sup> Davidov 2019 <sup>258</sup>
Cooper 2017 <sup>135</sup>	Cooper 2017 <sup>135</sup>
Coyle 2019 <sup>199</sup>	Coyle 2019 <sup>199</sup>
Cramer 2015	Cramer 2015 <sup>249</sup>
CRG 2016 <sup>114</sup>	CRG 2009 <sup>114</sup>
Crooks 2013 <sup>267</sup>	Crooks 2013 <sup>267</sup>
Crooks 2018 <sup>318</sup>	Crooks 2018; <sup>318</sup> Crooks 2018; <sup>329</sup> Crooks 2014 <sup>355</sup>

TABLE 35 Full list of included publications (*continued*)

Lead study	Linked papers
Cutbush 2017 <sup>312</sup>	Cutbush 2017 <sup>312</sup>
de Lijster 2016 <sup>168</sup>	deLijster 2016 <sup>168</sup>
Decker 2018 <sup>174</sup>	Decker 2018 <sup>174</sup>
Devries 2017 <sup>175</sup>	Devries 2017 <sup>203</sup>
Diegel 1999 <sup>203</sup>	Diegel 1999 <sup>115</sup>
DMSS 2012 <sup>115</sup>	DMSS Research & Consultancy 2012 <sup>115</sup>
dos Santos 2019 <sup>163</sup>	dos Santos 2016; <sup>330</sup> dos Santos 2019 <sup>142</sup>
Dozois 2016 <sup>302</sup>	Dozois 2016 <sup>302</sup>
Dunlop 2018 <sup>319</sup>	Dunlop 2018 <sup>319</sup>
Durand 1997 <sup>83</sup>	Durand 1997 <sup>83</sup>
Edwards 2019 <sup>197</sup>	Edwards 2019; <sup>197</sup> Edwards 2020; <sup>230</sup> Edwards 2021; <sup>204</sup> Waterman 2021 <sup>277</sup>
Edwards 2019 <sup>208</sup>	Edwards 2019 <sup>208</sup>
Eisman 2019 <sup>221</sup>	Eisman 2019 <sup>221</sup>
Ekhtiari 2014 <sup>234</sup>	Ekhtiari 2014; <sup>234</sup> Ekhtiari 2013 <sup>171</sup>
Elias-Lambert 2010 <sup>217</sup>	Elias-Lambert 2010 <sup>217</sup>
Elias-Lambert 2015 <sup>213</sup>	Elias-Lambert 2015 <sup>213</sup>
Enriquez 2012 <sup>220</sup>	Enriquez 2012 <sup>220</sup>
Espelage 2013 <sup>192</sup>	Espelage 2013; <sup>192</sup> Espelage 2015; <sup>37</sup> Espelage 2015; <sup>239</sup> Espelage 2017 <sup>274</sup>
Exner-Cortens 2020 <sup>320</sup>	Exner-Cortens 2020 <sup>320</sup>
Farrelly 2020 <sup>102</sup>	Farrelly 2020 <sup>102</sup>
Fawson 2016 <sup>214</sup>	Fawson 2016 <sup>214</sup>
Fay 2006 <sup>198</sup>	Fay 2006 <sup>198</sup>
Fernandez-Gonzalez 2020 <sup>169</sup>	Fernandez-Gonzalez 2020 <sup>169</sup>
Filho 2017 <sup>84</sup>	Filho 2017 <sup>84</sup>
Foshee 1998 <sup>36</sup>	Foshee 1998; <sup>36</sup> Foshee 1998; <sup>243</sup> Foshee 2000; <sup>242</sup> Foshee 2004; <sup>242</sup> Foshee 2005 <sup>241</sup>
Gage 2016 <sup>165</sup>	Gage 2016; <sup>165</sup> Gage 2016 <sup>331</sup>
Gale 2011 <sup>244</sup>	Gale 2011 <sup>244</sup>
Garces-Foley 2017 <sup>259</sup>	Garces-Foley 2017 <sup>259</sup>
Garcia-Escobar 2020 <sup>143</sup>	GarciaEscobar 2020 <sup>143</sup>
Gardner 2005 <sup>209</sup>	Gardner 2005 <sup>209</sup>
Garmaroudi 2016 <sup>172</sup>	Garmaroudi 2016 <sup>172</sup>
Genereux 2020 <sup>236</sup>	Genereux 2020 <sup>236</sup>
Genoves 2009 <sup>116</sup>	Genoves 2009 <sup>116</sup>
Gibbs 2016 <sup>263</sup>	Gibbs 2016 <sup>263</sup>

continued

TABLE 35 Full list of included publications (*continued*)

Lead study	Linked papers
Gonzalez 2014 <sup>144</sup>	Gonzalez 2014 <sup>144</sup>
Gonzalez-Guarda 2015 <sup>193</sup>	Gonzalez-Guarda 2015 <sup>193</sup>
Grimm 2011 <sup>210</sup>	Grimm 2011 <sup>210</sup>
Guillot-Wright 2018 <sup>215</sup>	Guillot-Wright 2018 <sup>215</sup>
Hale 2012 <sup>105</sup>	Hale 2012 <sup>105</sup>
Hall 2016 <sup>313</sup>	Hall 2016 <sup>216</sup>
Harrington 2019 <sup>216</sup>	Harrington 2019 <sup>216</sup>
Henderson 2002 <sup>106</sup>	Henderson 2002 <sup>106</sup>
Henshaw 2016 <sup>316</sup>	Henshaw 2016 <sup>316</sup>
Hertel 2020 <sup>262</sup>	Hertel 2020 <sup>262</sup>
Jaime 2016 <sup>178</sup>	Jaime 2016 <sup>178</sup>
Jalušič 2019 <sup>117</sup>	Jalušič 2019 <sup>117</sup>
Jaycox 2006 <sup>185</sup>	Jaycox 2006; <sup>185</sup> Jaycox 2006 <sup>86</sup>
Jemmott 2018 <sup>176</sup>	Jemmott 2018 <sup>176</sup>
Jewkes 2019 <sup>155</sup>	Jewkes 2019 <sup>155</sup>
Jones 2010 <sup>309</sup>	Jones 2010 <sup>309</sup>
Joppa 2016 <sup>186</sup>	Joppa 2016 <sup>186</sup>
Jorba 2012 <sup>111</sup>	Jorba 2012 <sup>111</sup>
Jordan 2018 <sup>147</sup>	Jordan 2018 <sup>147</sup>
Jouriles 2019 <sup>181</sup>	Jouriles 2019 <sup>181</sup>
Jung 2013 <sup>173</sup>	Jung 2013 <sup>173</sup>
Kearney 2016 <sup>148</sup>	Joyce 2019; <sup>332</sup> Joyce 2019; <sup>333</sup> Kearney 2016; <sup>148</sup> Kearney 2016; <sup>334</sup> Ollis 2018 <sup>335</sup>
Keddie 2020 <sup>103</sup>	Keddie 2020 <sup>103</sup>
KellydeAlbuquerque 2020 <sup>104</sup>	KellydeAlbuquerque 2020 <sup>104</sup>
Kempes 2010 <sup>118</sup>	Kempes 2010 <sup>118</sup>
Kershner 1996 <sup>194</sup>	Kershner 1996 <sup>194</sup>
Kervin 2010 <sup>224</sup>	Kervin 2010 <sup>224</sup>
Knowles 1997 <sup>232</sup>	Knowles 1997 <sup>232</sup>
Lapointe 2018 <sup>256</sup>	Lapointe 2018 <sup>256</sup>
Lee 2018 <sup>87</sup>	Lee 2018 <sup>87</sup>
Lerner 1999 <sup>306</sup>	Lerner 1999 <sup>306</sup>
Levesque 2016 <sup>182</sup>	Levesque 2007; <sup>336</sup> Levesque 2016; <sup>182</sup> NCT 2015 <sup>337</sup>
Levesque 2017 <sup>211</sup>	Levesque 2017 <sup>211</sup>
Louden 2016 <sup>149</sup>	Louden 2016 <sup>149</sup>
Lynch 2014 <sup>88</sup>	Lynch 2014 <sup>88</sup>
Macgowan 1997 <sup>89</sup>	Macgowan 1997 <sup>89</sup>

TABLE 35 Full list of included publications (*continued*)

Lead study	Linked papers
Madsen 1994 <sup>233</sup>	Madsen 1994 <sup>233</sup>
Makleff 2019 <sup>205</sup>	Makleff 2019; <sup>205</sup> Makleff 2020; <sup>338</sup> Makleff 2020 <sup>265</sup>
Malo-Juvera 2012 <sup>339</sup>	Malo-Juvera 2012; <sup>339</sup> Malo-Juvera 2014 <sup>195</sup>
Maphosa 2018 <sup>136</sup>	Maphosa 2018 <sup>136</sup>
Mathews 2016 <sup>90</sup>	Fonn 2017; <sup>261</sup> Mathews 2016 <sup>90</sup>
Matos 2006 <sup>120</sup>	Matos 2006 <sup>120</sup>
Maxwell 2010 <sup>121</sup>	Maxwell 2010 <sup>121</sup>
McElwee 2020 <sup>122</sup>	McElwee 2020 <sup>122</sup>
McGinn 2017 <sup>123</sup>	McGinn 2017 <sup>123</sup>
Meiksin 2020 <sup>35</sup>	Bragg 2020; <sup>257</sup> Meiksin 2020; <sup>35</sup> Meiksin 2020; <sup>91</sup> Ponsford 2019; <sup>340</sup> Ponsford 2021 <sup>341</sup>
Merrell 2004 <sup>92</sup>	Merrell 2004 <sup>92</sup>
Merrill 2018 <sup>137</sup>	Merrill 2018 <sup>137</sup>
Mgalla 1998 <sup>138</sup>	Mgalla 1998 <sup>138</sup>
Miller 1998 <sup>162</sup>	Miller 1998 <sup>162</sup>
Miller 2012 <sup>77</sup>	Jaime 2015; <sup>342</sup> Jaime 2018; <sup>245</sup> Jones 2021; <sup>303</sup> Miller 2012; <sup>77</sup> Miller 2013 <sup>343</sup>
Miller 2015 <sup>187</sup>	Miller 2015; <sup>187</sup> Miller 2015 <sup>93</sup>
Miller 2020 <sup>180</sup>	Miller 2020 <sup>180</sup>
Mills 1998 <sup>150</sup>	Mills 1998 <sup>150</sup>
Muck 2018 <sup>94</sup>	Muck 2018; <sup>94</sup> Muck 2018 <sup>344</sup>
Munoz-Rivas 2019 <sup>164</sup>	Munoz-Rivas 2019 <sup>145</sup>
Murray 2019 <sup>315</sup>	Murray 2019 <sup>315</sup>
Murta 2016 <sup>145</sup>	Murta 2016 <sup>145</sup>
Namy 2015 <sup>124</sup>	Namy 2014; <sup>345</sup> Namy 2015; <sup>124</sup> Young Men Initiative 2012 <sup>346</sup>
Nicholson 2018 <sup>139</sup>	Nicholson 2018 <sup>139</sup>
Niolon 2019 <sup>166</sup>	DeGue 2020; <sup>275</sup> Luo 2020; <sup>305</sup> Niolon 2019 <sup>166</sup>
Noonan 2009 <sup>307</sup>	Noonan 2009 <sup>307</sup>
Oliveira 2016 <sup>146</sup>	Oliveira 2016 <sup>146</sup>
Ollis 2011 <sup>151</sup>	Ollis 2011; <sup>151</sup> Ollis 2014 <sup>347</sup>
Ollis 2017 <sup>152</sup>	Ollis 2017 <sup>152</sup>
Ouellett 1998 <sup>222</sup>	Ouellett 1998 <sup>222</sup>
Pacifici 2001 <sup>191</sup>	Pacifici 2001 <sup>191</sup>
Payne 2018 <sup>260</sup>	Payne 2012; <sup>348</sup> Payne 2018 <sup>260</sup>
Perez-Marco 2020 <sup>125</sup>	Perez-Marco 2020 <sup>125</sup>
Peskin 2014 <sup>189</sup>	Peskin 2014 <sup>189</sup>
Peskin 2019 <sup>188</sup>	Peskin 2019 <sup>188</sup>

continued

TABLE 35 Full list of included publications (*continued*)

Lead study	Linked papers
Racionero-Plaza 2020 <sup>126</sup>	Racionero-Plaza 2020 <sup>126</sup>
Raible 2017 <sup>201</sup>	Raible 2016; <sup>349</sup> Raible 2017 <sup>201</sup>
Roberts 2009 <sup>95</sup>	Roberts 2009 <sup>95</sup>
Rogers 2018 <sup>127</sup>	Rogers 2018; <sup>127</sup> Rogers 2018 <sup>350</sup>
Rosenbluth 2004 <sup>159</sup>	Rosenbluth 2004; <sup>159</sup> Rosenbluth 2011; <sup>351</sup> Whitaker 2004 <sup>237</sup>
Rowe 2015 <sup>96</sup>	Rowe 2015 <sup>96</sup>
Sabella 1995 <sup>158</sup>	Sabella 1995 <sup>158</sup>
Sanchez-Jimenez 2018 <sup>49</sup>	Munoz-Fernandez 2019; <sup>247</sup> Sanchez-Jimenez 2018 <sup>49</sup>
Sanders-McDonagh 2015 <sup>128</sup>	Sanders-McDonagh 2015 <sup>128</sup>
Sargent 2017 <sup>38</sup>	Sargent 2017 <sup>38</sup>
Sarr 2019 <sup>140</sup>	Sarr 2019 <sup>140</sup>
Schwandt 2016 <sup>141</sup>	Schwandt 2016 <sup>141</sup>
Scull 2018 <sup>98</sup>	Scull 2018 <sup>98</sup>
Scull 2021 <sup>97</sup>	Scull 2021 <sup>97</sup>
Shevlin 2020 <sup>153</sup>	Shevlin 2020 <sup>153</sup>
Silverman 2000 <sup>99</sup>	Silverman 2000 <sup>99</sup>
Singh 2013 <sup>207</sup>	Singh 2013 <sup>207</sup>
Sorbring 2015 <sup>129</sup>	Ryding 2013; <sup>251</sup> Sorbring 2015 <sup>129</sup>
Taylor 2008 <sup>248</sup>	Taylor 2008; <sup>248</sup> Taylor 2010; <sup>157</sup> Taylor 2010 <sup>200</sup>
Taylor 2011 <sup>100</sup>	Mabin 2019; <sup>352</sup> Taylor 2011; <sup>100</sup> Taylor 2013; <sup>40</sup> Taylor 2015 <sup>276</sup>
Taylor 2015 <sup>353</sup>	Taylor 2015; <sup>353</sup> Taylor 2017 <sup>156</sup>
Tello 2013 <sup>113</sup>	Tello 2013 <sup>113</sup>
Tilbury 2019 <sup>154</sup>	Tilbury 2019 <sup>154</sup>
Turner 2006 <sup>212</sup>	Turner 2006 <sup>212</sup>
Tutty 2011 <sup>179</sup>	Tutty 2011 <sup>179</sup>
UNICEF 2016 <sup>107</sup>	UNICEF 2016 <sup>107</sup>
Walther 1986 <sup>183</sup>	Walther 1986 <sup>183</sup>
Walton 2007 <sup>130</sup>	Walton 2007 <sup>130</sup>
Watson 2012 <sup>202</sup>	Watson 2012 <sup>202</sup>
Watts 2016 <sup>304</sup>	Watts 2016 <sup>304</sup>
Weingarten 2018 <sup>108</sup>	Weingarten 2018 <sup>108</sup>
Wernick 2013 <sup>206</sup>	Wernick 2013 <sup>206</sup>
Williams 2017 <sup>131</sup>	Williams 2017 <sup>131</sup>
Wilson 2005 <sup>314</sup>	Wilson 2005 <sup>314</sup>
Winegust 2015 <sup>310</sup>	Winegust 2015 <sup>310</sup>
Wolfe 2009 <sup>190</sup>	Crooks 2017; <sup>354</sup> Wolfe 2009 <sup>190</sup>
Yoder 2020 <sup>101</sup>	Yoder 2020 <sup>101</sup>
Yom 2005 <sup>170</sup>	Yom 2005 <sup>170</sup>

## Appendix 2 Extra information about review methods

### Overview of included studies

An overview of the studies included in the review, including which research questions they presented evidence for, is provided in [Table 34](#). A full list of included studies with all linked papers is provided in [Table 35](#).

**TABLE 36** Overview of data points extracted across included studies

Study design	Population	Intervention	Data collection and analysis	Outcome data
Trial name	Cluster inclusion/exclusion criteria	Background support	Description of data collection methods (quantitative and qualitative)	Description of outcome
Linked publications	Participant inclusion/exclusion criteria	Intervention name	Description of data analysis methods (quantitative and qualitative)	Measure used (e.g. a validated scale)
Location of trial sites	Total student sample size	Intervention target (DRV/GBV/both)	Type of comparison	Outcome type (count, scale, frequency)
Study design	Total staff sample size	Intervention facilitator	Approach to analysis (ITT, ITT with imputation, per protocol)	Follow-up time description
Sample size and number of clusters	Student age	Method of delivery	Covariates included in analysis	Follow-up time category ( $\leq 1$ year vs. $> 1$ year)
Time frame of recruitment or data collection	Student sex/gender	Allocated clusters	Outcome category	Direction of outcome
Method of recruitment	Student sexuality	Allocated participants	ICC	
Method of randomisation	Student ethnicity	Duration of intervention		
Use of blinding	Student or school SES			
Power justification	Student relationship status			
School setting information	Student abuse history			
Aims of evaluation	Staff description			
Sampling strategy	Comparability of groups at baseline			

## Data extraction tool

A DET was developed a priori in Microsoft Excel® to extract study-level characteristics for studies reporting process, implementation, efficacy, moderation, mediation, cost and resource use data. The DET was piloted prior to use and was amended following feedback from stakeholders. The DET involved extraction of a wide range of data points across multiple tabs and is not reported in full here. An overview of data points extracted across studies is shown in [Table 36](#); a selection of these data were extracted as relevant to each research question.





EME  
HSDR  
HTA  
PGfAR  
**PHR**

Part of the NIHR Journals Library  
[www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

*This report presents independent research funded by the National Institute for Health and Care Research (NIHR).  
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the  
Department of Health and Social Care*

***Published by the NIHR Journals Library***