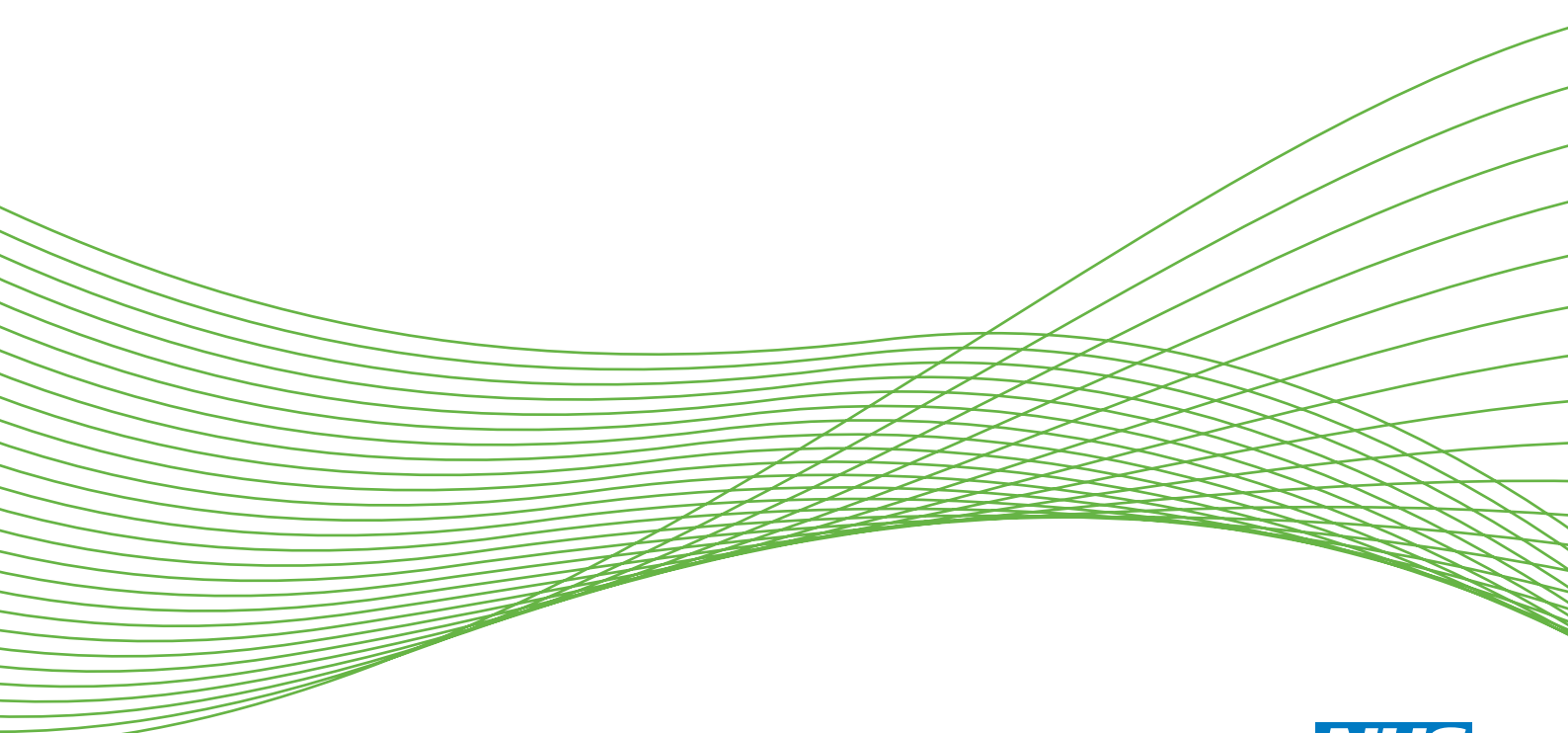


The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes

Chris Bonell, Kelly Dickson, Kate Hinds, GJ Melendez-Torres, Claire Stansfield, Adam Fletcher, James Thomas, Katrina Lester, Elizabeth Oliver, Simon Murphy and Rona Campbell



**National Institute for
Health Research**

The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes

Chris Bonell,^{1*} Kelly Dickson,² Kate Hinds,²
GJ Melendez-Torres,³ Claire Stansfield,²
Adam Fletcher,⁴ James Thomas,² Katrina Lester,^{2†}
Elizabeth Oliver,² Simon Murphy⁴ and Rona Campbell⁵

¹Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, UK

²Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), Social Science Research Unit, University College London Institute of Education, London, UK

³Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK

⁴Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), School of Social Sciences, Cardiff University, Cardiff, UK

⁵Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), School of Social and Community Medicine, University of Bristol, Bristol, UK

*Corresponding author

†In memoriam. The report is dedicated to the memory of Katrina Lester, a very fine researcher and a kind and generous woman.

Declared competing interests of authors: Rona Campbell reports grants from University of Bristol during the conduct of the study and is a Director of DECIPHer Impact Limited, a not-for-profit company wholly owned by the Universities of Bristol and Cardiff whose purpose is to maximise the translation and impact of evidence-based public health improvement research and expertise. It does this by selling goods and services starting with the DECIPHer-ASSIST (A Stop Smoking in Schools Trial) smoking prevention programme. Rona Campbell received modest fees in payment for her work as Director which are paid into an account held at the University of Bristol and used to fund research-related activity. Elizabeth Oliver reports grants from the University College London Institute of Education during the conduct of the study.

Published May 2016

DOI: 10.3310/phr04050

This report should be referenced as follows:

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**(5).

Public Health Research

ISSN 2050-4381 (Print)

ISSN 2050-439X (Online)

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

Editorial contact: nihredit@southampton.ac.uk

The full PHR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/phr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Public Health Research* journal

Reports 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 Research (NIHR), evaluates 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. The Public Health Research programme also complements the NIHR Health Technology Assessment programme which has a growing portfolio evaluating NHS public health interventions.

For more information about the PHR programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/phr>

This report

The research reported in this issue of the journal was funded by the PHR programme as project number 12/153/19. The contractual start date was in October 2013. The final report began editorial review in June 2015 and was accepted for publication in November 2015. 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' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health 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, NETSCC, the PHR programme or the Department of Health. 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, NETSCC, the PHR programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2016. This work was produced by Bonell et al. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Public Health Research Editor-in-Chief

Professor Martin White Director of Research and Programme Leader, UKCRC Centre for Diet and Activity Research (CEDAR), MRC Epidemiology Unit, Institute of Metabolic Science, School of Clinical Medicine, University of Cambridge; Visiting Professor, Newcastle University; and Director, NIHR Public Health Research Programme

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the HTA Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Peter Davidson Director of NETSCC, HTA, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor Elaine McColl Director, Newcastle Clinical Trials Unit, Institute of Health and Society, Newcastle University, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research and Development Group, University of Winchester, UK

Professor John Norrie Health Services Research Unit, University of Aberdeen, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk

Abstract

The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes

Chris Bonell,^{1*} Kelly Dickson,² Kate Hinds,² GJ Melendez-Torres,³ Claire Stansfield,² Adam Fletcher,⁴ James Thomas,² Katrina Lester,^{2†} Elizabeth Oliver,² Simon Murphy⁴ and Rona Campbell⁵

¹Department of Social and Environmental Health Research, London School of Hygiene and Tropical Medicine, London, UK

²Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre), Social Science Research Unit, University College London Institute of Education, London, UK

³Division of Health Sciences, Warwick Medical School, University of Warwick, Coventry, UK

⁴Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), School of Social Sciences, Cardiff University, Cardiff, UK

⁵Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), School of Social and Community Medicine, University of Bristol, Bristol, UK

*Corresponding author chris.bonell@lshtm.ac.uk

†In memoriam. The report is dedicated to the memory of Katrina Lester, a very fine researcher and a kind and generous woman.

Background: Positive Youth Development (PYD) delivered outside school aims to enable young people to develop positive assets such as relationships and confidence, rather than to merely address risk. Existing reviews of PYD effects on substance use or violence are old and unsystematic.

Objectives: To systematically review evidence to answer the following questions: what theories of change inform PYD interventions addressing substance use and violence? What characteristics of participants and contexts are identified as barriers to and facilitators of implementation and receipt in process evaluations of PYD? What is the effectiveness and cost-effectiveness of PYD in reducing substance use and violence? What characteristics of participants and contexts appear to moderate, or are necessary and sufficient for, PYD effectiveness?

Data sources: A total of 21 bibliographic databases; 35 websites and contacting authors.

Review methods: We included reports published in English since 1985 and reporting on theories of change, as well as process, outcome and economic evaluations of PYD targeting 11- to 18-year-olds and addressing substance use or violence. References were screened on title/abstract and, where appropriate, on full report. Data extraction and quality assessment used Critical Appraisal Skills Programme, Evidence for Policy and Practice Information and Co-ordinating Centre and Cochrane tools. Theories of change and process evaluations were qualitatively synthesised. Outcome evaluations were synthesised narratively and meta-analytically.

Results: 32,394 unique references were identified and 48 were included. A total of 16 reports described theories, 13 (10 studies) evaluated processes and 25 (10 studies) evaluated outcomes.

Theories of change: PYD interventions aim to offer opportunities for young people to develop positive 'assets' such as skills and confidence. These are theorised to promote and be promoted by young people's 'intentional self-regulation', which involves reflecting on behaviour; determining goals; using existing resources to pursue these; and redirecting effort when thwarted. This enables 'developmental regulation', namely individuals capitalising on other opportunities to promote personal development. Positive assets thus accrued reduce health risks by reducing the impact on individuals of environmental risk or by ameliorating the impact of such risks. The literature offers limited insights beyond these general ideas.

Process evaluations: Community engagement ensured that programmes were accessible and appealing. Staff capacity and continuity were crucial factors but often challenging when programmes could not offer full-time jobs. Tensions arose between a desire to empower participants to choose activities and a requirement for them to undertake a breadth of activities.

Outcome evaluations: Meta-analyses of all combined outcomes and of short-term alcohol use, illicit drug use and smoking found no significant effects. There were small, statistically significant, short-term effects for an omnibus measure of substance use and for violence. We could not undertake metaregression to assess sociodemographic moderators but narrative synthesis suggested no clear pattern of effects by sex. We found no economic evaluations.

Limitations: Insufficient studies precluded qualitative comparative analyses.

Conclusions: How PYD might promote health is currently undertheorised. Implementation can be challenging. We found little evidence that current PYD interventions delivered outside school reduce substance use or violence. However, these may not constitute a test of the effectiveness of the PYD model, as some included interventions that, although meeting our inclusion criteria, were not exemplars of PYD.

Future work: Further evaluations should assess interventions employing PYD theory of change.

Study registration: This study is registered as PROSPERO CRD42013005439.

Funding: The National Institute for Health Research Public Health Research programme.

Contents

List of tables	xiii
List of figures	xv
List of boxes	xvii
List of abbreviations	xix
Plain English summary	xxi
Scientific summary	xxiii
Chapter 1 Background	1
Health-risk behaviours and young people in the UK	1
Youth programmes to reduce health risks in young people	1
Rationale for this review	2
Review aims and objectives	3
Review questions	3
Chapter 2 Review methods	5
About this chapter	5
Design	5
Criteria for considering studies for this review	5
<i>Date</i>	5
<i>Language</i>	5
<i>Types of participant</i>	5
<i>Types of intervention and setting</i>	6
<i>Types of studies</i>	7
<i>Types of outcomes</i>	7
Search strategy	7
<i>Database search strategy</i>	7
Information management	9
Study selection	10
Data extraction	10
<i>Coding tools</i>	10
<i>Data extraction process</i>	10
<i>Missing data</i>	11
Quality assessment	11
<i>Theory studies</i>	11
<i>Process evaluations</i>	11
<i>Outcome evaluations</i>	12
Synthesis of results	12
<i>Theory</i>	12
<i>Process evaluations</i>	14
<i>Outcome evaluations</i>	14
<i>Economic evaluations</i>	15
User involvement	16
Ethical arrangements	16

Chapter 3 Included studies	17
About this chapter	17
Results of the search	17
Study characteristics	20
<i>Rate of publication</i>	20
<i>Geographical location</i>	20
<i>Target age group of Positive Youth Development programmes</i>	20
<i>Sampled population</i>	20
Chapter 4 Synthesis of theories of change	23
About this chapter	23
Included studies	23
Quality of studies	24
Synthesis of theories of Positive Youth Development	24
<i>Normative theory of what is Positive Youth Development</i>	24
<i>Taxonomy of Positive Youth Development</i>	27
<i>Causal theory of change for Positive Youth Development</i>	28
<i>The generalisability of Positive Youth Development theories of change</i>	30
<i>Summary of theory synthesis</i>	31
<i>Hypotheses arising from the theory synthesis</i>	33
Feedback from consultation on the theory synthesis	33
<i>Policy stakeholders' feedback</i>	33
<i>Young people's feedback</i>	34
Chapter 5 Synthesis of process evaluations	35
About this chapter	35
Included studies	35
Quality of process evaluation methods	37
<i>Reliability of process evaluations</i>	37
<i>Usefulness of qualitative studies</i>	38
<i>Overall weight of evidence</i>	38
Thematic synthesis of process evaluations	39
<i>Theme 1: collaboration with the community</i>	39
<i>Theme 2: young people's relationship with programme providers and peers</i>	41
<i>Theme 3: staff retention</i>	43
<i>Theme 4: youth-led empowerment</i>	44
Summary of synthesis of process evaluations	46
Hypotheses arising from the synthesis of process evaluations	46
Feedback from consultation on the synthesis of process evaluations	46
<i>Feedback from policy and practitioner stakeholders</i>	46
<i>Feedback from consultation with young people</i>	47
Chapter 6 Synthesis of outcome evaluations	49
About this chapter	49
Included studies	49
Evaluation design	52
Risk of bias and quality of evidence	53
<i>Allocation concealment</i>	53
<i>Blinding</i>	53
<i>Incomplete outcome data</i>	55
<i>Selective reporting</i>	55

<i>Type of comparator</i>	55
<i>Accounting for clustering</i>	55
<i>Other sources of bias</i>	55
Effects of interventions on substance use	55
<i>Narrative synthesis of effects on substance use</i>	55
<i>Meta-analyses of effects on substance use</i>	59
Effects of interventions on violence	75
<i>Narrative synthesis of effects on violence</i>	75
<i>Meta-analysis of effects on violence</i>	76
Characteristics of participants and contexts appearing to moderate or to be necessary and sufficient for Positive Youth Development effectiveness	79
<i>Overall quality of evidence and Grading of Recommendations, Assessment, Development and Evaluations assessment</i>	79
Feedback from consultation on the synthesis of outcome evaluations	80
<i>Policy stakeholders' feedback</i>	80
Chapter 7 Discussion and conclusions	81
Review limitations	81
<i>Deviations from protocol</i>	81
<i>Other limitations</i>	81
Key results	84
<i>Included studies</i>	84
<i>Taxonomy of Positive Youth Development</i>	84
Conclusions	88
Implications	89
<i>Implications for research</i>	89
<i>Implications for public health</i>	90
Acknowledgements	91
References	93
Appendix 1 Protocol	105
Appendix 2 Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist	123
Appendix 3 Exclusion criteria and guidance	125
Appendix 4 Search strategies for bibliographic databases	127
Appendix 5 Data extraction tool for theory studies	131
Appendix 6 Data extraction tool for extracting process information	133
Appendix 7 Data extraction tool for outcome evaluations	135
Appendix 8 Author contact for missing data	139
Appendix 9 Theory studies: coding templates	141
Appendix 10 Process studies: coding templates	153

Appendix 11 Data transformation	157
Appendix 12 User involvement	159
Appendix 13 Characteristics of theory reports	169
Appendix 14 Characteristics and quality assessment of process evaluations studies	175
Appendix 15 Application of theory synthesis to descriptions of Positive Youth Development interventions included in the process and outcome evaluations	189
Appendix 16 Characteristics and quality risk of bias of outcome evaluations	193
Appendix 17 Policy and young people's report briefs	213

List of tables

TABLE 1 Initial coding template	13
TABLE 2 Overview of interventions subject to process or outcome evaluation included in this review	19
TABLE 3 Reliability and usefulness of findings	38
TABLE 4 Risk-of-bias summary review authors' judgements about each risk-of-bias item for each included study	53
TABLE 5 Measures used in studies reporting substance-use outcomes	56
TABLE 6 Substance-use outcome meta-analyses	60
TABLE 7 Measures used in studies reporting violence outcomes	75
TABLE 8 Violence outcome meta-analyses	76
TABLE 9 Summary of findings for PYD for substance use and violence	79
TABLE 10 Deviations from proposal and protocol	81
TABLE 11 Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist	123
TABLE 12 Exclusion criteria and guidance	125
TABLE 13 Data transformation	157
TABLE 14 Characteristics of theory reports (studies included in the review to answer RQ1)	169
TABLE 15 Quality assessment of theory studies	173
TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2)	175
TABLE 17 Quality assessment of process evaluations	188
TABLE 18 Application of theory synthesis to descriptions of Positive Youth Development interventions included in the process and outcome evaluations	190
TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3)	193
TABLE 20 Quality assessment of outcome evaluations	211

List of figures

FIGURE 1 Flow of studies in the review	18
FIGURE 2 Rate of study publication	21
FIGURE 3 Spread of studies by country (mutually exclusive)	22
FIGURE 4 Target age group of programmes according to study type (mutually exclusive)	22
FIGURE 5 Population sampled in outcome and process evaluations (not mutually exclusive)	22
FIGURE 6 Synthesised theory of change for PYD effects on substance use and violence	32
FIGURE 7 Review authors' judgements about each risk-of-bias item presented as percentages across all included studies	54
FIGURE 8 All substance-use outcomes	61
FIGURE 9 Short-term substance-use outcomes	64
FIGURE 10 Omnibus substance-use outcomes	67
FIGURE 11 Short-term omnibus substance-use outcomes	68
FIGURE 12 Illicit drug-use outcomes	69
FIGURE 13 Short-term illicit drug-use outcomes	70
FIGURE 14 Alcohol outcomes	71
FIGURE 15 Short-term alcohol outcomes	72
FIGURE 16 Smoking outcomes	73
FIGURE 17 Violence outcomes	77
FIGURE 18 Short-term violence outcomes	78

List of boxes

BOX 1 NVivo Node Structure Report: Kate Hinds	141
BOX 2 NVivo Node Structure Report: Chris Bonell	145

List of abbreviations

ALPHA	Advice Leading to Public Health Advancement	OR	odds ratio
ASP	after-school programme	PI	principal investigator
BBBS	Big Brothers Big Sisters of America	PYD	Positive Youth Development
CI	confidence interval	PYDC	Positive Youth Development Collaborative
DECIPHer	Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement	QOP	Quantum Opportunity Program
GED	General Educational Development	RCT	randomised controlled trial
GRADE	Grading of Recommendations, Assessment, Development and Evaluations	RQ	review question
MAP	Maryland after-school programme	SES	socioeconomic status
NGYCP	National Guard Youth Challenge Program	YARP	Youth Action Research for Prevention
		YPDP	Young People's Development Programme

Researchers

AF	Adam Fletcher	KH	Kate Hinds
CB	Chris Bonell	KL	Katrina Lester
CS	Claire Stansfield	LM	Lizzie Oliver
GM	GJ Melendez-Torres	RC	Rona Campbell
JT	James Thomas	SM	Simon Murphy
KD	Kelly Dickson		

Plain English summary

Positive Youth Development (PYD) delivered outside school provides young people with additional education and leisure opportunities to promote overall personal development, for example in terms of self-esteem, confidence, skills and aspirations. PYD is presented as a way in which to reduce substance use (smoking, drinking alcohol and drug use) and violence. We reviewed existing research on this by examining how PYD is intended to reduce substance use and violence; how delivery is affected by context; and the impacts of PYD on substance use and violence. We explicitly defined what evidence we would include and how we would integrate it. We searched thoroughly for relevant research, finding 48 reports, of which 25 examined impacts, mostly in the USA.

Our review of PYD theory suggested that PYD aims to help young people identify their goals and how to achieve these. This should make them less likely to become involved in or harmed by violence and drug use in their neighbourhoods. Existing theories are not clear about exactly how this protection occurs.

Our review of evaluations of PYD delivery suggest that it is most successful when programmes engage with young people's families and communities to gain their trust and support. Providers of PYD need to be skilful, but recruiting and retaining such people can be hard when funding is short term. There are sometimes tensions in PYD programmes between letting young people choose what activities to pursue and ensuring that they participate in diverse activities, including education.

Our pooling of the results of evaluations of PYD impacts on young people found very small reductions in substance use and violence, most of which probably occurred by chance.

Currently, there is no good evidence to support investments in PYD delivered outside school to prevent substance use and violence. However, before any new evaluations are undertaken, PYD providers need to be clearer about how interventions are meant to work.

Scientific summary

Background and rationale

Substance use and violence are very common and damaging to young people's health. There are calls for interventions to address multiple- rather than single-risk behaviours because such behaviours cluster together and such interventions are potentially more efficient. Positive Youth Development (PYD) delivered outside school is one such intervention to address interclustered risk behaviours among young people. The UK's National Youth Agency (NYA) defines such interventions as voluntary educational activities aiming to bring about generalised youth development in terms of positive assets such as skills and confidence, rather than to merely remedy 'problem behaviours'.

Non-systematic reviews of PYD effects on violence and drug use suggest benefits as well as variability, but must be treated with caution given that the majority of these reviews are unsystematic and quite old.

Aim and review questions

This systematic review aimed to systematically search for, appraise the quality of and synthesise evidence on PYD interventions addressing substance use or violence, asking the following review questions (RQs):

RQ1: what theories of change inform PYD interventions delivered to young people aged 11–18 years addressing substance use and violence?

RQ2: what characteristics of participants and contexts are identified as barriers to and facilitators of implementation and receipt in process evaluations of PYD?

RQ3: what is the effectiveness and cost-effectiveness of PYD compared with usual or no treatment in reducing substance use (smoking, alcohol use, drug use) and violence (perpetration and victimisation)?

RQ4: what characteristics of participants and contexts appear to moderate, or are necessary and sufficient for, PYD effectiveness?

Methods

We searched for reports published in English since 1985 on theories of change, as well as process, outcome and economic evaluations of PYD interventions targeting youths aged 11–18 years to reduce substance use or violence, including experimental and quasi-experimental outcome evaluations. We searched 21 bibliographic databases plus websites, clinical trials registers and expert consultation. References were screened on title and abstract, then, where appropriate, on full report. Data extraction and quality assessment of included studies were undertaken by two reviewers, who used a third reviewer when a consensus could not be reached. Quality assessment used existing tools.

Two reviewers used template analysis and thematic content analysis to qualitatively metasynthesise theories of change and process evaluation findings. We developed a narrative and meta-analytic synthesis of outcome evaluations, the latter of which used multilevel meta-analytic models to account for multiple reported effect sizes. Because studies did not report how effects varied by subgroup, we were not able to use metaregression to examine how intervention effects were moderated by sociodemographic characteristics. Insufficient study numbers and quality and low statistical or qualitative heterogeneity of

results precluded metaregression or qualitative comparative analysis to test hypotheses about how intervention effects varied by characteristics of participants and contexts. Economic data were not synthesised as no such studies were found.

We consulted policy/practice stakeholders and young people on completion of our synthesis of theories and processes to inform them of how we used these to develop hypotheses, as well as on completion of our draft report to inform them of our plans for communication and dissemination.

Results

Included studies

We identified 32,394 unique references from searches. Of these, 31,634 were excluded by screening on title and abstract. Of the 760 remaining references, we could obtain full reports of 689. Screening these led to a further 641 studies being excluded. The remaining 48 included reports arose from a total of 30 distinct studies (i.e. a distinct description of theory of change or empirical evaluation). Sixteen reports described theories of change; 13 reports (from 10 distinct studies) evaluated processes; and 25 reports (from 10 distinct studies) evaluated outcomes. Five reports combined theories of change, process evaluation and/or outcome evaluation (see *Table 3* in the main report for an overview and overlap between study reporting and programmes evaluated).

Review question 1: theories of change for Positive Youth Development effects on substance use and violence

Sixteen reports were included. We aimed to assess the quality of these theories by drawing on criteria used previously, but found that it was challenging to apply these consistently to the PYD theoretical literature.

The included literature did not provide sufficient information to develop a comprehensive theory of change for the effects of PYD interventions on substance use and violence. Nonetheless, by filling in some of the gaps in the literature (in a transparent manner) we succeeded in generating a theory of change.

Positive Youth Development interventions are intended to provide a positive environment for young people in terms of positive expectations; enduring and affective relationships with adults; diverse activities and settings; and active participation whereby young people are empowered to choose activities and take on responsibilities. Interventions vary according to whether they aim to enable young people to contribute to or challenge the existing social order; focus on individual or collective development; aim to transform individuals only or also the environments in which they live; provide breadth, depth and duration of activities; and address only positive assets or also risk behaviours.

In providing positive expectations and active participation in diverse activities, PYD aims to offer 'affordances', that is, resources individuals use in the course of their development (e.g. relationships, challenges, education). We interpret that young people can make use of these affordances in learning to apply 'intentional self-regulation' to specific intervention activities such as sports, arts or outdoor activities. Intentional self-regulation comprises 'intentionality' (assessing their current skills); 'selection' (setting goals for what they want to achieve); 'optimisation' (using their existing skills and the new affordances that PYD provides to achieve these goals); and 'compensation' (reviewing and if necessary redirecting actions to meet goals).

Positive Youth Development interventions aim to reward young people when they make progress with activities regarded as pro-social within that programme. Drawing on, but filling in gaps in, the PYD literature, we interpret that PYD interventions enable young people to engage in and learn from diverse, mutually reinforcing intervention activities, so that they develop generalised intentional self-regulation focused on pro-social goals.

As a result of developing intentional self-regulation, young people are better able to develop various 'positive assets' such as the '5 Cs': competence, confidence, connection, character and caring. As these accrue, young people can make better use of the opportunities available in their wider environments, which leads to positive 'developmental regulations', that is, a positive feedback cycle whereby individuals gain more benefit from opportunities in their environment. This in turn enables them to make a positive contribution to their communities and societies, or, as a few authors would argue, enables them to contribute by critiquing and challenging inequities present in the existing social order.

These positive assets may then reduce risk behaviours via 'buffering', whereby risk factors in a young person's environment have less impact on those with positive assets, or 'compensation', whereby even if a young person engages in a risk behaviour, their possession of positive assets ameliorates the impacts of this on their overall health and development. (Note that this use of compensation is quite distinct from that within the description of intentional self-regulation above.) It is also argued that positive assets may reduce risk via 'molecular' mechanisms, whereby a specific asset exerts specific protection against a specific risk, or via 'pile-up', whereby accumulation of multiple assets is protective regardless of the specific assets involved.

However, the theoretical literature synthesised here offers only limited insights beyond these general ideas. It suggests that engagement with pro-social peer groups or institutions might reduce antisocial behaviours via exposure and adherence to pro-social norms. It also suggests that improved emotional self-regulation, social skills and self-efficacy might contribute to better decision-making to avoid violence and substance use. But this falls short of a systematic theory of how the accrual of assets in particular or in general contributes to reductions in substance use or violence.

Review question 2: characteristics of participants and contexts that affect implementation and receipt of Positive Youth Development

Of the 10 included studies, eight were conducted in the USA, one was conducted in Australia and one was conducted in England. Study quality ranged from high reliability and usefulness to low reliability and usefulness.

A number of themes emerged from synthesis. Community engagement was a key to ensuring that programmes were culturally sensitive, accessible and appealing to young people and their parents, as well as to the wider community. Employing community members could be pivotal to successful implementation and providing role models. However, volunteers could be unreliable, for example in acting as mentors. Collaboration with other community agencies could be important, particularly in expanding the range of activities offered, but could lead to a move away from the original approaches.

Another theme was young people's relationships with providers and peers. Providers should relate to young people in a calm and nurturing but authoritative way. Skilled providers could bridge social differences between participants, but this could be undermined by poor training or retention. Retention of staff was challenging where programmes could not offer full-time positions. A final theme concerns the challenges to ensuring that young people are empowered to make decisions about programme activities while also requiring them to engage in diverse activities, including vocational or academic activities.

Review question 3: effectiveness and cost-effectiveness in reducing substance use and violence

We found 13 study reports of 10 distinct outcome evaluations and included 12 study reports of nine distinct outcome evaluations in our meta-analyses. All but one study was conducted in the USA (one was conducted in the UK). We could not categorise interventions as per the taxonomy derived from our theory synthesis because of the lack of detail reported. We found no economic evaluations and thus could not assess

cost-effectiveness. Four studies were randomised controlled trials; five were non-randomised trials with prospectively matched control groups; and one included both randomised trial and non-randomised components which were analysed together. Overall quality of evidence for our analyses of substance-use and violence outcomes was rated 'very low'. Only one study described a theory of change for how its PYD-type intervention components might lead to reductions in substance use or violence.

Our meta-analyses found a small overall effect for substance-use outcomes that was significant neither statistically nor in terms of public health relevance. Meta-analyses found no effect for an omnibus measure of substance use across all time points but did find a small, statistically significant, effect for this outcome in the short term (0–4 months post intervention). Meta-analyses of illicit drug-use and alcohol outcomes found no significant effects either across all time points or in the short term. Our meta-analysis for smoking at all time points included only four studies and so its finding of no significant effects should be treated with caution. Meta-analysis of short-term smoking was not undertaken, as only two studies could have been included. It was not possible to undertake metaregressions looking at effects by sex. Narrative synthesis of three studies that examined subgroup effects found that these were mixed with no clear pattern. We aimed to examine effects by socioeconomic status, ethnicity or area deprivation but no studies reported these. One evaluation presented a subgroup analyses by combinations of sex and ethnicity, reporting that minority ethnic status was not a moderator of effects.

Meta-analyses suggested that PYD interventions did not have a statistically significant effect on violence outcomes across all time points, but that there was a beneficial effect on short-term outcomes. However, this finding was only marginally significant in sensitivity analysis and should be treated with caution. Metaregression to assess subgroup effects was not possible. Only one evaluation reported effects by sex and found that this did not moderate effects. It also examined presented subgroup analyses by categories defined by sex and ethnicity, reporting that these did not moderate effects.

Review question 4: characteristics of participants and contexts that determine effectiveness

We aimed to examine the participant characteristics and contexts that appear to moderate, or are necessary and sufficient for, PYD effectiveness. A synthesis of PYD theories of change and process evaluations suggested several hypotheses:

- Interventions that offer a breadth of activities may be more effective for younger adolescents, whereas those that emphasise depth may be more effective for older adolescents.
- Interventions that combine prevention and positive development may be less effective than those that focus only on positive development.
- Interventions of more than 1 year's duration may be more effective than those of shorter duration.
- Interventions may be more effective for participants with low or moderate levels of baseline risk, as there is more scope for stimulating 'intentional self-regulation'.
- Interventions that have specific methods to engage communities will be more effective.
- Projects that engage with schools will achieve better recruitment.
- Interventions that are delivered by well-trained staff will be more effective.
- Interventions that have better staff retention will be more effective.
- Interventions that offer some choices but require some engagement with educational components will be more effective.

However, the limited number of studies and very low level of statistical heterogeneity or qualitative differences in the effects that these studies reported precluded examination of these hypotheses via metaregression or qualitative comparative analyses.

Conclusions

Theories of change for how PYD might impact on substance use or violence are currently inadequate. These neither fully describe how interventions enable participants to develop 'intentional self-regulation' nor explain how the development of positive assets enables young people to reduce their risk of engaging in substance use and violence. Any future evaluations of the effects of PYD interventions on these outcomes must clarify the intended mechanisms of action.

Tensions can arise in how PYD is implemented in practice. Programmes are often intended to empower young people's choice of activity and to ensure that they are engaged in diverse activities to develop multiple assets. Programmes usually aim to use skilled providers who can engage participants in long-term affective relationships but this can be undermined by agencies' inability to offer secure, full-time jobs. Use of volunteers can engender community support and offer role models, but volunteers may also be hard to retain.

Our meta-analyses do not offer evidence that existing PYD interventions delivered outside school have effects of public health importance in reducing substance use and violence among young people. Our conclusions with regard to smoking should be treated with caution because of the paucity of included studies. Our review may not constitute a test of the effectiveness of the PYD model, as the interventions evaluated in included studies, although meeting our inclusion criteria, were commonly implemented variously between sites and often emphasised explicit risk reduction as much as, if not more than, positive development, so these may not be exemplars of PYD. PYD aiming to reduce substance use and violence should be implemented only in the context of rigorous evaluation. We found no studies examining the cost-effectiveness of PYD.

Study registration

This study is registered as PROSPERO CRD42013005439.

Funding

Funding for this study was provided by the Public Health Research programme of the National Institute for Health Research.

Chapter 1 Background

Health-risk behaviours and young people in the UK

The health of young people in the UK is among the worst in Europe, with marked inequalities across the social scale.^{1,2} Health-risk behaviours increase during adolescence³ and lead to high rates of later chronic disease and other problems, and to substantial economic costs.⁴ Child poverty is currently increasing, raising the possibility of upwards trends in young people's risk behaviours, with worrying implications for future chronic disease rates and NHS costs.⁵ Smoking, drinking alcohol and using illicit drugs (henceforth termed substance use), as well as violence, are highly prevalent and damaging to young people's long-term health. Rates of youth smoking, drinking and illicit drug use in the UK are among the highest in Europe.^{6,7} Despite significant declines since the 1990s, 17% of 15-year-old boys and 20% of 15-year-old girls are regular or occasional smokers, and around one-quarter of 15-year-olds drink alcohol every week, while 15% of 15-year-old boys and 11% of 15-year-old girls report drug use in the past month.⁸ Recent estimates suggest that more than 11,500 under-18-year-olds access drug treatment services each year.⁹ There are short- as well as long-term public health threats arising from young people's substance use. For example, adolescent use of cannabis is associated in the short term with increased sexual risk behaviour and injury.¹⁰ Young people's use of substances is associated with social disadvantage across studies,¹¹ and reinforces existing socioeconomic inequalities in health across the life course. This raises key equity considerations: substance use is most prevalent among socially disadvantaged young people and frequent use at a young age is strongly associated with more harmful use and chronic illness in adulthood.¹²⁻¹⁴ Aggression and violence are similarly challenging, and preventing youth violence continues to be a public health, education and criminal justice priority.¹⁵⁻¹⁷ One survey reports that by 15-16 years of age, one-quarter of young people have carried a weapon, and 19% reported attacking someone with the intention to hurt them seriously.¹⁸ Violence is subject to marked social inequalities¹⁹ and is associated with an increased risk of physical health problems;²⁰ engaging in other health-risk behaviours such as substance use;²¹⁻²³ long-term emotional, behavioural and mental health problems;^{20,24,25} and self-harm and suicide.²⁶ Moreover, gang involvement is associated with acute health risks and strongly correlated with later-life offending and serious, adverse mental health outcomes in longitudinal studies.²⁷ The economic costs associated with youth substance use and aggression are extremely high.^{12,28,29}

Youth programmes to reduce health risks in young people

Primary prevention interventions to reduce health risks in adolescence are potentially highly cost-effective.³⁰ There are increasing calls for adolescent health interventions to address multiple- rather than single-risk behaviours because such behaviours cluster together^{31,32} and because such interventions are potentially more feasible and efficient.³³ Positive Youth Development (PYD) is one such intervention to address interclustered risk behaviours among young people. PYD is the dominant paradigm in youth work in the UK.

The National Youth Agency (NYA), the major youth-work organisation in the UK, defines such interventions as voluntary and informal educational activities aiming to bring about generalised youth development rather than merely remedying 'problem behaviours'. Such development is defined in terms of the promotion of positive skills, attitudes, relationships and identities.³⁴ A literature review published by the NYA developed a complex definition of PYD in terms of philosophy, constructs, domains and processes but similarly emphasised young people's positive attributes and competencies through structured voluntary activities.³⁵

Similarly, in the USA, PYD is defined in terms of its goal of developing a range of positive development assets, such as bonding, resilience, social, emotional, cognitive, behavioural or moral competence, self-determination, spirituality, self-efficacy, clear and positive identity, belief in the future, recognition for

positive behaviour, opportunities for pro-social involvement and/or pro-social norms,³⁰ academic, cognitive or vocational skills, confidence, connections to peers and adults, character in terms of self-control, respect and morality and caring for others.³⁶

Although one aim of this review is to synthesise existing literature on the theory of change underlying PYD interventions, it was apparent to us in planning the review that PYD has the potential to reduce substance use and violence through various complex pathways. First, PYD can address some of the underlying social determinants of these outcomes, such as disengagement from education, lack of social support and low aspirations for the future.³⁰ Second, PYD can divert young people away from substance use and violence through engaging them in more positive forms of recreation.³⁶ Third, PYD can promote social and emotional competences, some of which are important protective factors against adolescent health-risk behaviours.⁵ Fourth, PYD providers can provide credible health messages and signpost health services.³⁷

Even in the context of public-sector cuts, there is major investment in such interventions. The UK government's Positive for Youth³⁸ report announced a multimillion pound investment in youth work, youth centres, the National Citizen Service and other youth volunteering projects. The most recent public health White Paper³⁹ cited such work as a key element in promoting young people's health. The Mayor of London and local government across the UK are also investing millions of pounds in various PYD interventions.⁴⁰ The devolved governments in Scotland and Wales also emphasise these principles and promote investment in PYD.^{41,42}

However, despite this widespread investment and potential, the evidence base for the public health benefits of such interventions is unclear. Although a systematic review examining non-health outcomes⁴³ reported benefits for self-confidence and self-esteem, school bonding, positive social behaviours, school grades and achievement test scores, the review did not systematically examine health effects. Systematic reviews of health outcomes have, so far, focused only on sexual health^{44,45} and have not attempted to meta-analyse evidence of effects. They have, however, reported sustained effects, albeit with considerable unexplained variability between programmes. For example, the Children's AID Society Carrera programme reduced teenage pregnancy in some US sites but not others,³⁷ whereas two evaluations of PYD interventions in the UK suggested adverse and no effects on sexual health, respectively.^{46,47} US researchers have argued that some youth programmes that target 'delinquent' young people and that are insufficiently well structured may actually reinforce violence and antisocial behaviours via peer deviancy training.⁴⁸ Others have disputed this, referring to meta-analyses of interventions addressing youth delinquency⁴⁹ which suggest that the targeting and structure of sessions do not moderate effects. However, no systematic review focusing on PYD interventions has examined these questions. Non-systematic reviews of PYD effects on violence and drug use^{30,50} have reported benefits as well as variability, but their findings must be treated with caution, given that they were unsystematic and are now quite old.

Rationale for this review

This review aims to fill two timely and important knowledge gaps and to provide important evidence to local government commissioners of youth services and public health. First, it aims to synthesise evidence on the effectiveness and cost-effectiveness of PYD interventions delivered outside school as a means of primary prevention in reducing substance use and violence. Second, it aims to examine how effects vary according to the characteristics of participants, in order to assess what works, for whom and in what settings (to inform assessments of generalisability), as well as to estimate effects on health inequalities. Addressing intervention effects is important because, as described above, UK young people have among the worst health in Europe, with marked inequalities across the social scale. Moreover, PYD interventions are receiving significant policy attention and investment, despite a lack of evidence of health benefits from systematic reviews. Addressing moderators of effects is also important given the possibility discussed above that PYD effects will vary and given our interest in assessing the potential of PYD to reduce health inequalities.

Review aims and objectives

The overarching aim of this systematic review is to systematically search for, appraise the quality of and synthesise evidence on PYD programmes that address substance use or violence and that examine the extent to which these effects vary and/or are moderated by characteristics of participants and contexts.

These aims have been addressed by focusing on the following objectives:

1. conducting electronic and other searches for studies of PYD interventions
2. screening references and reports for inclusion in the review
3. extracting data from and assessing the quality of included studies
4. synthesising thematically theories of change relating to PYD interventions to produce a taxonomy and theory of change for PYD interventions
5. synthesising process evaluations of PYD interventions
6. consulting with policy-makers and practitioners and youth to validate the resultant taxonomy and theory of change
7. synthesising outcome and economic evaluation data and undertaking metaregression and qualitative comparative analyses
8. drawing on these syntheses to draft a report addressing our review questions (RQs)
9. consulting with policy-makers and practitioners and young people on the draft report to inform amendments and dissemination.

Review questions

The following RQs were addressed:

RQ1: what theories of change inform PYD interventions delivered to young people aged 11–18 years addressing substance use and violence?

RQ2: what characteristics of participants and contexts are identified as barriers and facilitators of implementation and receipt in process evaluations of PYD?

RQ3: what is the effectiveness and cost-effectiveness of PYD compared with usual or no treatment in reducing substance use (smoking, alcohol, drugs) and violence (perpetration and victimisation)?

RQ4: what characteristics of participants and contexts appear to moderate, or are necessary and sufficient for, PYD effectiveness?

Chapter 2 Review methods

About this chapter

This section outlines the methods used in this systematic review. They were described a priori in a research protocol⁵¹ (see *Appendix 1*). The study is registered as PROSPERO CRD42013005439 (see www.crd.york.ac.uk/PROSPERO/). Although there are no checklists for a complex, multimethod review such as the one undertaken, we have adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidance (the PRISMA checklist can be found in *Appendix 2*).⁵²

Design

The project is a multimethod systematic review of known existing research. This chapter describes the flow of studies through the review as well as the characteristics of included studies. This is followed by four chapters presenting our various syntheses:

1. A thematic synthesis of the literature describing the theory of change of PYD interventions.
2. A thematic synthesis of process evaluations.
3. A narrative synthesis and meta-analysis of experimental [randomised controlled trials (RCTs)] and quasi-experimental studies (employing non-randomised prospective comparison groups) of the effectiveness of PYD interventions on substance use and violence outcomes. We found no economic evaluations.
4. An overview bringing together all three syntheses outlined above.

Criteria for considering studies for this review

The criteria and definitions used for considering which studies to include in this review are outlined below. These inclusion criteria were operationalised into exclusion criteria to inform our screening of studies found (see *Appendix 3*). The results of this screening process are detailed in *Chapter 3*.

Date

We included only studies published in and after 1985, as this is when PYD interventions first began to be developed.^{30,44} Our original proposal did not restrict included studies by date; we added this exclusion criterion at an early stage in the review, including this in our registered protocol.⁵¹

Language

We included only studies published in English because PYD interventions appear to be a phenomenon particular to anglophone countries.^{30,44} Our original proposal did not exclude studies published in other languages; we added this exclusion criterion at an early stage in the review, including this in our registered protocol.⁵¹

Types of participant

We included studies in which the majority of participants were aged 11–18 years. Although the World Health Organization defines adolescents, the target group for PYD, as those aged 10–19 years,⁵³ to increase this review's UK policy relevance, we chose 11–18 years as our age range as this encompasses those engaged in secondary education in the UK. We excluded studies of populations targeted on the basis of pre-defined physical and mental health conditions (because we are interested in PYD as primary prevention) but not those that targeted participants on the basis of pre-existing risk behaviour or other forms of targeting (e.g. area-level deprivation). We also excluded interventions that targeted parents/carers alongside young people in order to focus on family functioning.

Types of intervention and setting

Informed by existing theoretical frameworks,^{30,44} PYD interventions were defined as programmes that involve voluntary education with the aim not merely of preventing problem behaviour but also of promoting generalised (beyond health) and positive (beyond avoiding risk) development, which were defined as promoting:

- bonding (developing the child's relationship with a healthy adult, positive peers, school, community, or culture)
- resilience (strategies for adaptive coping responses to change and stress and for promoting psychological flexibility and capacity)
- social competence (developmentally appropriate interpersonal skills and rehearsal strategies for practising these skills, including communication, assertiveness, refusal and resistance, conflict-resolution and interpersonal negotiation strategies for use with peers and adults)
- emotional competence (identifying feelings in self or others, skills for managing emotional reactions or impulses, or skills for building the youth's self-management strategies, empathy, self-soothing or frustration tolerance)
- cognitive competence (cognitive abilities, processes or outcomes including academic performance, logical and analytic thinking, problem-solving, decision-making, planning, goal-setting and self-talk skills)
- behavioural competence (skills and reinforcement for effective verbal, non-verbal and other actions)
- moral competence (empathy, respect for cultural or societal rules and standards, a sense of right and wrong or a sense of moral or social justice)
- self-determination (capacity for empowerment, autonomy, independent thinking or self-advocacy, or their ability to live and grow by self-determined internal standards and values, which may or may not include group values)
- spirituality (beliefs in a higher power, internal reflection or meditation; supporting youth in exploring a spiritual belief system or sense of spiritual identity, meaning or practice)
- self-efficacy (personal goal-setting, coping and mastery skills, or techniques to change negative self-efficacy expectancies or self-defeating cognitions)
- clear and positive identity (healthy identity formation and achievement in youth, including positive identification with a social or cultural subgroup that supports young people's healthy development of a sense of self)
- belief in the future (belief in his or her future potential, goals, options, choices or long-range hopes and plans were classified as promoting belief in the future, including guaranteed tuition to post-secondary institutions, school-to-work linkages, future employment opportunities or future financial incentives to encourage continued progress on a pro-social trajectory; or optimism about a healthy and productive adult life)
- recognition for positive behaviour (response systems for rewarding, recognising or reinforcing children's pro-social behaviours were classified as using recognition for positive behaviour)
- opportunities for pro-social involvement (activities and events in which youths could actively participate, make a positive contribution and experience positive social exchanges) and/or
- pro-social norms (clear and explicit standards for behaviour that minimised health risks and supported pro-social involvement).

Included interventions either needed to address at least one of these forms of asset but could be applied to different domains such as family, community, school, or needed to address more than one of these assets in a single domain.

Our original funding proposal defined PYD interventions in terms of voluntary education provided by youth workers and addressing generalised, positive development in terms of vocational, academic, social or cognitive skills; self-confidence; positive identities, attitudes and aspirations; and/or relationships with adults or peers. However, we used the more theoretically informed definition listed above from an earlier stage in the review, which was included in our registered protocol.⁵¹

We included studies in which interventions were provided in community settings (which could include schools) outside of normal school time. Our definition excluded PYD delivered in school time because this involves a distinctive theory of change and has been the subject of recent reviews.^{54,55} It also excludes interventions delivered in custodial or probationary settings, clinical settings or employment training for school leavers, again because such interventions will involve distinctive theories of change, and, in the case of clinical and employment training settings, will feature participants not meeting our inclusion criteria.

Types of studies

We included multiple types of studies based on whether or not they could answer the individual RQs. In order to address RQ1, we included studies describing a PYD intervention theory of change in relation to our outcomes. We defined theory in the same way as in our previous National Institute for Health Research (NIHR)/Public Health Research (PHR)-funded review of the effects of schools and school-environment interventions on health.⁵⁶ Included studies either dealt exclusively with theory of change or addressed it alongside the reporting of empirical data.

In order to address RQ2, we included studies reporting on process evaluations of PYD intervention. Included studies reported on how the planning, delivery, receipt or causal pathways of PYD varied or were influenced by characteristics of place or person using quantitative and/or qualitative data. These studies either reported exclusively on process evaluations or reported process data alongside outcome or economic data. In order to address RQ3, we included studies reporting on outcome and economic evaluations of PYD interventions. We included experimental (RCTs) and quasi-experimental studies (employing non-randomised prospective comparison groups). Control groups needed to receive usual care or no treatment. Economic studies addressing RQ3 were defined in terms of their comparison of the costs and consequences of two or more interventions or, where there was good reason to believe that outcomes were similar, involved cost-minimisation analyses. In order to address RQ4, we have drawn on the syntheses of all of the above study types.

Types of outcomes

This review included studies addressing substance use (i.e. smoking, alcohol use and/or drug use) or violence (i.e. perpetration and/or victimisation involving physical violence aimed at person(s) as opposed to damage to property).

Informed by existing systematic reviews that focus on substance use and violence among young people,⁵⁷⁻⁶⁰ outcome measures could draw on either dichotomous or continuous variables and/or self-report or observational data. They could use measures of frequency (monthly, weekly or daily), the number of episodes of use or an index constructed from multiple measures. Alcohol measures could examine alcohol consumption or problem drinking. Drug outcomes could examine general or specific illicit drug use. Measures of violent and aggressive behaviour could examine the perpetration or victimisation of physical violence including violent crime.

Search strategy

Database search strategy

Search terms

A sensitive search strategy using both indexed and free-text terms was developed and tested by an experienced information scientist (CS). These searches were run on 7 November 2013. Key search terms were determined by the RQ and the inclusion criteria and were developed and tested against papers already known to the research team in writing the research proposal. The search strategy involved developing strings of terms and synonyms to capture three core concepts in the review:

1. Concept 1: population (e.g. youth or young people or adolescents).
2. Concept 2: intervention (e.g. after-school clubs or community-based programme or informal education).
3. Concept 3: population and intervention (e.g. youth work or youth development or youth club).

These concepts were combined in searches as follows: concept 1 and (concept 2 or concept 3). The initial free-text search terms generated for concepts 1 and 2 were broad and could identify non-relevant literature (e.g. volunteer, development, etc.). Thus, to balance specificity with sensitivity, where possible we required that intervention terms were adjacent or near to population terms (e.g. after 'school' N12 'young people' OR 'adolescen*' OR 'youth'). An example of a search string from PsycINFO can be found in *Appendix 4*.

Databases

Searches were undertaken on the following 21 electronic bibliographic databases:

- Applied Social Science Abstracts via ProQuest
- Australian Educational Index via ProQuest
- BiblioMap health promotion research via EPPI-Centre
- British Educational Index via ProQuest
- Child Data via Ovid
- Cochrane Central Register of Controlled Trials via The Cochrane library
- Cumulative Index to Nursing and Allied Health Literature via EBSCOhost
- Database of Abstracts of Reviews of Effects
- Database of Promoting Health Effectiveness Reviews via EPPI-Centre
- Health Technology Assessment Database
- EconLit via Ovid
- Education Research Index Citations via ProQuest
- International Bibliography Social Sciences via ProQuest
- MEDLINE via EBSCOhost
- NHS Economic Evaluation Database
- PsycINFO via EBSCOhost
- Sociological Abstracts via ProQuest
- Social Sciences Citation Index via Web of Knowledge
- Social Policy and Practice via Ovid
- Social care online via Ovid The Health Management Information Consortium via Ovid
- Trials Register of Promoting Health Interventions (via EPPI-Centre).

Other search sources

The following websites were also searched to identify relevant studies:

- The Campbell Library
- International Clinical Trials Registry Platform
- OpenGrey
- The Health and Safety Information Centre of the International Labour Office
- Drug and Alcohol Findings Effectiveness Bank
- Dissertation Abstracts/Index to Theses
- Schools and students health education unit research archive
- Google Scholar (Google Inc., Mountain View, CA, USA)
- Google (Google Inc., Mountain View, CA, USA)
- US Center for Substance Abuse Prevention
- Northern Ireland Online Research Base
- US Food and Drug Administration–Centers for Disease Control and Prevention Youth Tobacco Prevention
- Tobacco Use Behaviour Research
- UK Clinical Research Network Study Portfolio
- Personal Social Services Research Unit Discussion Papers (see www.pssru.ac.uk/)
- Cost-effectiveness Analysis Registry
- National Youth Agency

- Social Policy Digest
- Drug database (see www.drug.org.au/)
- Smoking and Health Resource Library
- Child and Adolescent Health Research Unit
- Children in Scotland
- Children in Wales
- Children's Research Centre
- Welsh Government Social Research
- Scottish Government website
- Young Minds
- Social Issues Research Centre
- European Commission – Community Research and Development Information Service library
- Centre for Prevention Research and Development
- Joseph Rowntree Foundation
- Childhoods Today
- The Children's Society
- National Youth Agency
- National Criminal Justice Reference Service Abstracts.

Website searches were run between 7 and 16 January 2014. Depending on the functionality of each website interface, searches were undertaken using a combination of medical subject headings (if available) and/or free-text search terms to capture the key concepts of the intervention (e.g. 'positive youth development'; 'youth work'; 'youth clubs'). Citations were screened online based on their title, title and abstract or full text when available. Potentially includable studies were cross-referenced with the electronic searches imported to EPPI-Reviewer version 4.0 (Evidence for Policy and Practice Information and Co-ordinating Centre, University of London, London, UK) to identify any unique references. As is customary with searching of this type, only included references were recorded.

We also hand-searched those journals that (1) contained studies that we included; (2) were found only via reference checking; and (3) were not indexed on databases that we had searched. We hand-searched these initially for the past 5 years and if these elicited > 1 new included study we hand-searched for a further 5 years. Our original proposal stated that we would hand-search the five journals that yield the highest numbers of studies meeting the inclusion criteria. However, we decided very early on in the review that this approach would have low specificity in identifying studies not already identified by our other methods of searching and was thus not a good use of resources. We therefore amended this aspect of our search so that hand-searching was much more targeted towards studies that were unlikely to have been identified through other means, including this in our registered protocol.⁵¹

We also sought to contact subject experts to identify unpublished or ongoing research. We pursued this by contacting the authors of all included studies to seek their advice on other studies that we might consider for inclusion. As stated above, the International Clinical Trials Registry Platform was also searched to identify any relevant ongoing and unpublished trials on PYD. Finally, we searched reference lists of all included studies for further relevant studies.

Information management

All citations identified by our searches were uploaded and managed during the review process using the EPPI-Centre's specialist online review software, EPPI-Reviewer version 4.0 software.⁶¹ This software records the bibliographic details of each study; where studies were found and how; reasons for their inclusion or exclusion; descriptive and quality assessment codes; text about each included study; and the data used and produced during synthesis. The software also enables us to store and track electronic documents (e.g. portable document format files).

Study selection

An exclusion criteria worksheet, informed by our inclusion criteria and with guidance notes (see *Appendix 3*), was prepared and piloted by four reviewers (CB, KD, KD, LM), who screened 100 references in pairs, on title and abstracts. We increased the number of studies on which to pilot the screening from 50 studies as originally set out in our proposal to 100, in order to facilitate greater discussion on the application of the exclusion criteria by ensuring that we had a wider variety of citations to compare. Pilot screening results were discussed by pairs of reviewers involved in screening to ensure consistency in applying the criteria. A 90% agreement rate was required before proceeding to independent screening of the full set of references. Between submitting the proposal and finalising the protocol⁵¹ we piloted the search. Initial yields indicated that the use of sensitive and comprehensive terms to capture literature seeking to answer theory, process and outcome RQs would lead to a high number of studies to screen (e.g. above 10,000). It was decided that double screening of the search would therefore not be feasible and a better use of resources would be to pilot the exclusion criteria and screen independently, an approach taken in most EPPI-Centre reviews.⁶² The remaining references were divided between five reviewers (CB, KD, KH, KL, LM) with each reference being screened independently by one reviewer. If a single reviewer could not reach a decision regarding inclusion of a specific article, judgement for selection was referred to a second reviewer. If these reviewers could not reach a consensus then a third reviewer was consulted. Full reports were obtained for those references judged to meet our inclusion criteria based on title and abstract or where there was insufficient information from the title and abstract to judge inclusion. References thus passing this first round of title and abstract screening were subject to a second round of screening using the same approach, but based on full study reports in order to determine which studies were included in the review. The principal investigator [PI (CB)] made a final check of all studies identified as potential included in the review, as a final check to determine inclusion, to identify which RQ they answered and to determine cases in which multiple reports were 'linked', that is, reporting on the same study.

Data extraction

Coding tools

Data were extracted using coding tools developed for the review components relating to each RQ (see *Appendices 5–7*). Each drew on and supplemented the codes used in the EPPI-Centre classification system for health promotion and public health research.⁶³ For studies describing a theory of change (either as a purely theoretical study or an empirical study addressing a theory of change), we extracted data on aim; description of theory of change; links to other theories; description of how PYD is intended to act on the individual or their environment; and description of how PYD is intended to reduce substance use or violence. For both process and outcome evaluations, we extracted data on study location; intervention/ components; intervention development and delivery; timing of intervention and evaluation; provider characteristics; target population; sampling and sample characteristics (where relevant by wave of follow-up); data collection and analysis. For process evaluations, we also extracted data on findings relevant to our review, including verbatim qualitative data plus author descriptions and interpretations. For studies reporting on outcome evaluations, we also extracted data on the nature of the control group(s); research design; unit of allocation; generation and concealment of allocation; blinding; adjustment/control of clustering and confounding; outcomes; and effect sizes overall and by age, sex, socioeconomic status (SES) and ethnic subgroup. No economic evaluations were found, so no economic data were extracted.

Data extraction process

Data extraction tools were piloted on two studies. Reviewers met to compare extraction and to identify any differences that might inform refinements of the coding tools or how these were applied. All study reports were then extracted by two reviewers who worked independently before meeting to discuss and agree their coding, to ensure quality and consistency in their interpretations. If reviewers could not reach a consensus, judgement was referred to a third reviewer. The data extraction tool for theory was adapted slightly after it had been piloted on two papers.

Missing data

Where missing data might have affected our ability to assess the quality of studies or synthesise findings, we contacted study authors to request additional information (see *Appendix 8*). When authors were not traceable or sought information was not forthcoming within 2 months of contact, we recorded that the study information was missing on the data extraction form, and this was captured in our risk-of-bias assessment of the study.

Quality assessment

The quality of each study was assessed independently by two reviewers, with differences in opinion resolved by discussion without the need for recourse to a third reviewer. Given that no economic studies were found, no economic quality appraisal occurred.

Theory studies

The quality of studies reporting on theory was assessed using a new tool which was informed by a tool used in a previous NIHR review⁵⁶ as well as by other recent work on theory synthesis.^{64,65} Quality was assessed in terms of clarity of constructs, clarity of relationships between constructs, testability, parsimony and generalisability. We pre-specified how we would apply each of these criteria as follows:

1. Construct clarity: are the constructs clear, either because they are well-established constructs such as self-esteem or because they are more novel theoretical constructs such as cultural pride or spirituality for which the authors provide a definition?
2. Relationships between constructs: do the authors explain the relationships between their constructs? Do they describe each step in the theory of change?
3. Testability: would it appear possible in principle to test the theory of change within our review by drawing on experimental or quasi-experimental evaluations of intervention that were informed by these theories of change? For example, a theory of change involving improving young people's social competence to reduce violence would be testable, whereas a theory of change involving improvements to the way young people are treated in society overall would not be.
4. Parsimony: is the theory of change as simple as it can be without losing its value? This is a subjective decision but we could probably agree that some theories of change involve so many constructs and/or have so many interconnections between these constructs that it detracts from their ability either to inform an intervention or an empirical evaluation of whether an intervention informed by that theory of change is effective.
5. Generalisability: does the theory of change apply across: (1) behaviours; (2) populations; and (3) contexts, or is it relevant only to smoking but not drug use, or only to young men or Native American peoples but not to women or Asian peoples, or only to the USA but no other country, for example?

Process evaluations

Process evaluations were assessed using standard Critical Appraisal Skills Programme and EPPI-Centre tools for qualitative studies.⁶⁶ Quality tools for qualitative studies address the rigour of sampling, data collection, data analysis, the extent to which the study findings are grounded in the data, whether or not the study privileges the perspectives of participants, the breadth of findings and the depth of findings. A final step in the quality assessment of qualitative studies was to assign studies two types of 'weight of evidence'. First, reviewers assigned a weight (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). Second, reviewers assigned an additional weight (low, medium, high) to rate the usefulness of the findings for shedding light on factors relating to the RQs. Guidance was given to reviewers to help them reach an assessment on each criterion and the final weight of evidence.

Outcome evaluations

Outcome evaluations were assessed for risk of bias using the tool modified from the questions suggested in the Cochrane Handbook for Systematic Reviews of Interventions.⁶⁷ For each study, two reviewers independently judged the likelihood of bias in seven domains: sequence generation; allocation concealment; blinding (of participants, providers or outcome assessors); incomplete outcome data; selective outcome reporting; and other sources of bias (e.g. recruitment bias in cluster-randomised studies); and intensity/type of comparator. Each study was subsequently allocated a score of 'high risk', 'low risk' or 'unclear risk' within each domain.

We used the Grading of Recommendations, Assessment, Development and Evaluations (GRADE) approach as described in the Cochrane Handbook for Systematic Reviews of Interventions⁶⁸ to present the quality of evidence and 'Summary of findings' tables. 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; imprecision of results; and publication bias. The GRADE approach specifies four levels of quality (high, moderate, low and very low).

Synthesis of results

Theory

We synthesised descriptions of theory of change using a form of thematic content analysis known as template analysis.⁶⁹⁻⁷² Theory synthesis is not the same as synthesis of qualitative research for two reasons. First, theory synthesis analyses theoretical constructs rather than interpretive concepts. Theoretical concepts have been developed anew by theorists and are not necessarily based on prior concepts used by those who are the objects of theory. In contrast, interpretive concepts are interpretations of the concepts used by participants in qualitative research. Thus, theoretical concepts are first order, whereas qualitative interpretations are second order. Second, syntheses of theory bring together theoretical concepts that are often described using fairly consistent terminology across included reports and that may often be informed by similar assumptions. In contrast, syntheses of qualitative research bring together interpretive concepts which originate from quite diverse data and study designs, employ different terminology and are often informed by different epistemological and ontological assumptions. For these reasons, it would not be appropriate to use metaethnographic methods in synthesising theory. We instead used thematic content analysis. It may even be inaccurate to discuss the elements of the emerging synthesis in terms of 'themes', as this word suggests a label that has been given to help interpret a diverse array of qualitative data, whereas theory synthesis will more often need to proceed via the development of more tightly defined constructs and inter-relations. Nonetheless, we still believed that it would be useful to employ thematic content analysis to synthesise the theoretical literature, because the goal was to compare and contrast how constructs and the inter-relations between constructs are described across diverse sources in order to build theory that goes beyond what is present in each individual source.

The coding template was developed in three stages.

First, two reviewers developed an initial coding template based on theoretical concepts that were already apparent within the protocol or which occurred to them during their data extraction of the theoretical literature (*Table 1*). The codes were arranged hierarchically, going from broad themes to more specific codes and subcodes.

Second, using NVivo (QSR International, Warrington, UK) to manage data, the two reviewers independently used the initial template to code the same two theory papers (which were chosen on the basis that both researchers had earlier assessed these as high quality). In applying the coding template to code these two papers in turn, the researchers could each refine or add to the template as new theoretical 'themes' arose or existing ones were reformulated. Where codes were refined or added, the reviewers

TABLE 1 Initial coding template

Themes	Codes and subcodes
Definition of PYD interventions	PYD vs. prevention science/traditional youth programmes Definition in terms of developmental assets vs. programme atmosphere Characteristics of programmes
Taxonomy	Individual vs. environment/community emphasis Pro-social development vs. critical conscious raising General 'pile up' of assets vs. specific 'molecular' effects of specific assets on specific outcomes
Mechanism of action	Action on risk of substance use/violence Action on thriving
Possible moderation by context	Moderation by person/population Moderation by setting

wrote short 'memos' explaining their thinking. Each kept a record of the coding template as it stood at the end of coding each paper. The reviewers also noted how each study contributed to their development and refinement of the coding template. The reviewers then discussed the refined coding template that they had each produced, developing an agreed new common template.

Third, the reviewers then coded the rest of the theory papers using the template refined in the previous stage, coding the most recently published study reports first. As they did so, they could again refine or add to the coding template as new themes arose. In the course of analysing the papers, each reviewer wrote an overarching memo which presented the emerging overall synthesis as they saw it. This reflected the evolving coding template, expanding on it to describe emerging ideas about how PYD is defined and categorised and its intended mechanisms of action, as well as to hypothesise about possible moderators of intervention effects. Each reviewer recorded how each paper contributed to the emerging overall synthesis. Again, as reviewers coded each paper they kept a record of the coding template and the overall synthesis as it stood at the end of coding each paper.

In this way the researchers gradually developed a synthesis of PYD theory which attempted to develop a common and overarching definition, taxonomy, theorised mechanism of action and set of hypothesised contextual moderators for PYD. When the two researchers finished coding all the theory papers, they wrote up a summary drawing on their individual templates and memos. The text was organised under the four headings of the initial coding template, namely, definition of PYD, mechanism of action, moderation by context and taxonomy.

The reviewers then combined their efforts, producing an overall summary of their analysis by discussion and iterative drafting. The resulting text was then edited in response to verbal and written comments from co-investigators. These processes of refining the analysis did not fundamentally change the substance of the synthesis but did enable it to be presented in a conceptually clearer manner. For the initial templates that two reviewers created, see *Appendix 9*; these have been included to enable readers to judge the extent to which our final report of the synthesis tallies with our initial attempts at coding the theoretical literature.

Process evaluations

We aimed to produce a qualitative metasynthesis of process evaluations using thematic synthesis methods.⁷³⁻⁷⁵ Qualitative metasynthesis aims to develop interpretative explanations and understanding from multiple cases of a given phenomenon by using qualitative research reports. Two reviewers independently read and reread all study reports to gain a detailed understanding of the findings and then undertook line-by-line coding of the findings sections using EPPI-Reviewer version 4 software.⁶¹ They first applied *in vivo* codes to what Schutz⁷⁶ termed first-order (verbatim quotes from participations) and second-order constructs (authors' descriptions or interpretation of the data). Reviewers wrote memos to summarise their interpretations of these first- and second-order concepts. Axial codes were then applied which aimed to make connections between *in vivo* codes in order to deepen analysis, with reviewers writing memos to describe emerging 'metathemes'. Each reviewer developed an emerging coding template, a hierarchical organisation of the codes that were applied in the course of the analysis. The two reviewers then met to compare their coding templates and to agree a common template which would form the basis for the drafting of the synthesis, which represented a set of third-order constructs developed by the reviewers.

As the coding template was developed, the reviewers referred to tables summarising the methodological quality of each study, so that this could be taken account of in the synthesis, with findings from more reliable or more useful studies being given more weight.

The two reviewers met to compare and contrast their draft coding templates and memos in order to develop consensus about the development of a single, coherent, overall synthesis. One reviewer then drafted the overall synthesis drawing on each reviewer's coding template, memos and the assessment of the quality of each study. This was then commented upon by the second reviewer, with comments incorporated by the first (see *Appendix 10*).

Outcome evaluations

We undertook both narrative and meta-analytic synthesis of the results of outcome evaluations. Effect sizes from included study reports concerning substance use (smoking, alcohol or drugs) or violence as defined in the protocol⁵¹ were extracted into a Microsoft Excel (Microsoft Corporation, Redmond, WA, USA) spreadsheet and converted into standardised mean differences (Cohen's *d*) using all available information as presented for each study. As recommended by the Cochrane Handbook,⁶⁷ when the evaluation was designed as a RCT, we extracted the 'least adjusted' effect size estimates from each evaluation (i.e. uncontrolled estimates, or estimates controlling for baseline scores). When the evaluation was a matched or otherwise non-randomised design, we extracted the most adjusted effect size estimates (i.e. estimates in which the full vector of control variables was included). In interpreting the results of meta-analyses, we followed the standard rule for interpretation of Cohen's *d* that 0.2 is a small effect, 0.5 is a medium effect and 0.8 is a large effect. Positive effect sizes indicate an effect favouring the intervention.

Because of the variation in reporting across studies, some degree of data transformation and imputation was necessary (see *Appendix 11*). Where the method of analysis was a linear probability model, as was the case with the National Guard Youth Challenge Program (NGYCP)⁷⁷ and the Quantum Opportunity Program (QOP),⁷⁸ we took the group means on each outcome and used the sample sizes for each group to estimate odds ratios (ORs). In one case, an evaluation of after-school programmes (ASPs) reported by Tebes *et al.*,⁷⁹ we used gain scores to adjust for baseline differences. In this case, we estimated that the pre-test to post-test correlation was $r = 0.5$ and sensitivity analysed our estimates with $r = 0.1$ and $r = 0.9$. This was done to render all effect estimates in the same metric as recommended by Morris and DeShon.⁸⁰ Although we also intended to impute intracluster correlation coefficients for studies in which clustering was not appropriately addressed, we were unable to do so because of the small number of studies and the lack of any comparable estimate from other reviews. We note in the results where studies did not adequately address clustering in analyses.

Most studies reported several substance use and violence outcomes at several measurement time points. As indicated in the protocol⁵¹ we intended to use multivariate meta-analysis or another method to synthesise effect sizes in this situation. However, this was not possible because of the heterogeneity of outcomes and the lack of availability of a variance–covariance matrix for reported outcomes. Instead, we used multilevel meta-analysis as set out by Cheung⁸¹ with random effects at both the outcome and study level. Multilevel meta-analysis accounts for dependencies between outcomes from the same study by partitioning the variance (τ^2 -statistic) between outcomes into a within-study level and a between-study level. The final effect size estimate includes all information that the multiple effect size estimates contribute while correcting for the non-independence of multiple effect size estimates from each study. This method indicated that using a random-effects model to synthesise the evidence was appropriate.

We used a standard three-level model, with level one being the ‘hypothetical’ participants who contributed to the effect sizes, level two being the within-study outcome-specific effect size estimates with sampling error, and level three being the ‘between-study’ level [two programmes (NGYCP,⁷⁷ QOP⁷⁸) having multiple study reports contributing to the analysis]. We did not run a four-level model to account for clustering within study reports because multiple reports from the same study involved similar personnel and methods.

Because we did not know the covariance between violence and substance-use outcomes in this context, we ran two separate sets of analyses. First, as specified in the protocol⁵¹ we ran an overall model capturing all violence outcomes, a model capturing all substance-use outcomes, and individual models for smoking, alcohol and illicit drug-use outcomes. Second, because studies often reported ‘omnibus’ outcomes covering a variety of substance-use behaviours across smoking, alcohol and illicit drug use, we also examined these generic substance-use outcomes in a separate model.

In our protocol⁵¹ we planned to examine subgroup effects according to the timing of follow-up assessment: post intervention to 3 months, 3 months to 1 year and > 1 year. However, this would have been inappropriate given the right-skewed distribution of the measurement time points in included studies, so we decided to examine a subset of short-term outcomes captured between post intervention and 4 months in separate models. We did not examine outcomes beyond 4 months because this would have included an extremely broad range of follow-ups.

For each model, we estimated an overall effect size expressed as a standardised mean difference with a 95% confidence interval (CI). We tested overall heterogeneity using Cochran’s Q, and we estimated I^2 for the outcome level and the study level of the model using formulae published elsewhere.⁸¹ Interpretation of I^2 at the level of the study is most comparable to interpretations of I^2 in ‘standard’ meta-analyses that include one effect size per study.

We intended to estimate metaregression models both to examine how intervention effects varied by participants’ SES, sex and ethnicity, to examine how intervention effects varied by area deprivation, and to test hypotheses on other moderators of effects. These other hypotheses were derived from the syntheses of theory and process evaluations, and consultations with young people and policy/practitioner stakeholders. However, such analyses were not possible because of the absence of meaningful heterogeneity in effects between studies as well as the lack of consistency of reporting of subgroup effects within studies. We also intended to run a qualitative comparative analysis to examine the causal combinations of conditions that predict intervention effectiveness. However, because of the absence of meaningful differences across our qualitative assessments of the outcomes of each intervention, this was not possible. We also found insufficient studies (≥ 10 per outcome) to draw funnel plots to assess the presence of possible publication bias. Full details of these methods may be found in our protocol.⁵¹

Economic evaluations

Given that no economic evaluations were found, we did not synthesise economic data.

User involvement

We consulted with policy-makers, practitioners and young people during the course of the review. We convened a policy advisory group of the following stakeholders: Public Health England (Eustace de Sousa, Deputy Director, Children, Young People and Families); Department of Health (Geoff Dessent, Deputy Director, Health and Wellbeing); National Youth Agency (Jessica Urwin, Head of Policy); Association for Young People's Health (Ann Hagell, Research Lead); and Project Oracle/London Metropolitan University (Georgie Parry Crooke, Professor of Social Research and Evaluation). Young people were also consulted via the Advice Leading to Public Health Advancement (ALPHA) young people's public health research advisory group based in the Centre for the Development and Evaluation of Complex Interventions for Public Health Improvement (DECIPHer), a research collaboration between the universities of Cardiff, Bristol and Swansea.

Both groups were consulted at project inception as part of the protocol-development stage to gauge their interest in the work and identify additional priorities.⁵¹ This consultation informed our view that synthesising evidence on PYD was a priority and that this should include assessments of process alongside outcomes in order to consider the feasibility and acceptability of interventions and what contextual factors affect transferability across settings. This consultation also informed our decision to prioritise evidence of effects on substance use (smoking, alcohol and drugs) and violence (perpetration and victimisation).

Feedback from the policy advisory group and young people was also obtained in October 2014 and April 2015. In October 2014, the policy advisory group was provided with information about the review's aims, RQs and methods together with a draft synthesis of theory and process evaluations plus a list of potential hypothesis generated from these (see *Appendix 12*). Because a date for a meeting that everyone could attend proved impossible to find, consultation was undertaken via bilateral telephone conversations with a reviewer or via written e-mail feedback. Stakeholders' views were sought on the clarity of the draft syntheses, whether or not these resonated with their experiences in the UK and their views on the hypotheses that we had developed on the basis of our syntheses. The ALPHA group also met in October 2014, in person. A total of eight young people aged between 15 and 19 years took part in a group discussion, facilitated by two researchers, one of whom took notes. After providing a Microsoft PowerPoint® (Microsoft Corporation, Redmond, WA, USA) presentation distilling the main findings of the theory and process syntheses, the group provided feedback and engaged in a ranking exercise about the importance of various aspects of PYD interventions (see *Appendix 12*). In March 2015, both groups were contacted during the write-up of the final report to inform how the research outputs were structured and to inform the dissemination strategy.

Ethical arrangements

This project was approved by the research ethics committee of the Institute of Education's Faculty of Children and Learning (ethics approval reference number FCL 544). The project complied with the Social Research Association's ethical guidelines⁸² and guidance from the National Coordinating Centre for Public Engagement.⁸³

Chapter 3 Included studies

About this chapter

This chapter reports the results of our systematic search and screening process and gives a brief overview of the included studies.

Results of the search

A total of 44,445 references were identified from the searches. Of these, 12,051 (27%) were identified as duplicates. The remaining 32,394 references were screened on title and abstract and, of these, 31,634 (97.6%) were excluded using the criteria listed in *Appendix 3*.

When piloting the process for screening on title and abstract, initial screening agreements between reviewers varied slightly but were consistently above 90% on whether or not a study should be excluded. Agreement was lower on the question of which particular criterion should be cited in excluding a particular reference, varying from 45% to 77% among five different pairs of reviewers. Discussion between reviewers established that this reflected the multiple criteria that could be used to exclude many studies and, therefore, the choice of which particular criterion to cite in each case was somewhat arbitrary. Given that agreements were above 90% on whether to exclude or not, we moved to a system of one reviewer independently screening each reference, as set out in our protocol.⁵¹

Of the 760 remaining references included at this stage, we were able to obtain the full-text reports of 689, with the remainder not accessible online or in local libraries. Piloting the application of same exclusion criteria used at the title and abstract screening stage on full-text documents, screening agreements between two pairs of reviewers also reached above 90% on whether a study should be excluded or not. The agreements on which criterion to apply still remained lower, varying between 37% (KD and CB) to 66% (KD and KH) for the same reason of multiple possible criteria to apply in many cases. However, similar to the first round of screening, as agreement overall was above 90%, individual reviewers (KD, KH, CB) moved to independent screening of documents, consulting with a second reviewer when a decision could not be easily reached. The PI (CB) also made final checks of all the studies included by reviewers. This process led to a further 641 studies being excluded at this stage in the review screening process.

The remaining 48 reports deemed eligible for inclusion in the review were coded according to which RQ they answered. A total of 30 distinct research studies were included in the review. Five research studies contributed more than one study report to our set of included reports, together providing 23 reports that we needed to link to other reports from the same study. Furthermore, five study reports provided answers to more than one RQ. In presenting the number of studies and study reports, we have not double counted those that address more than one of our review questions. Where appropriate, we clarify whether or not a study or report addressed more than one of our RQs. *Figure 1* summarises the flow of references, reports and studies through the review, providing a breakdown of the exclusion criteria at both title and abstract and full document stages and the number of studies included in each synthesis. *Table 2* provides an overview of the interventions that were subject to process or outcome evaluation included in this review.

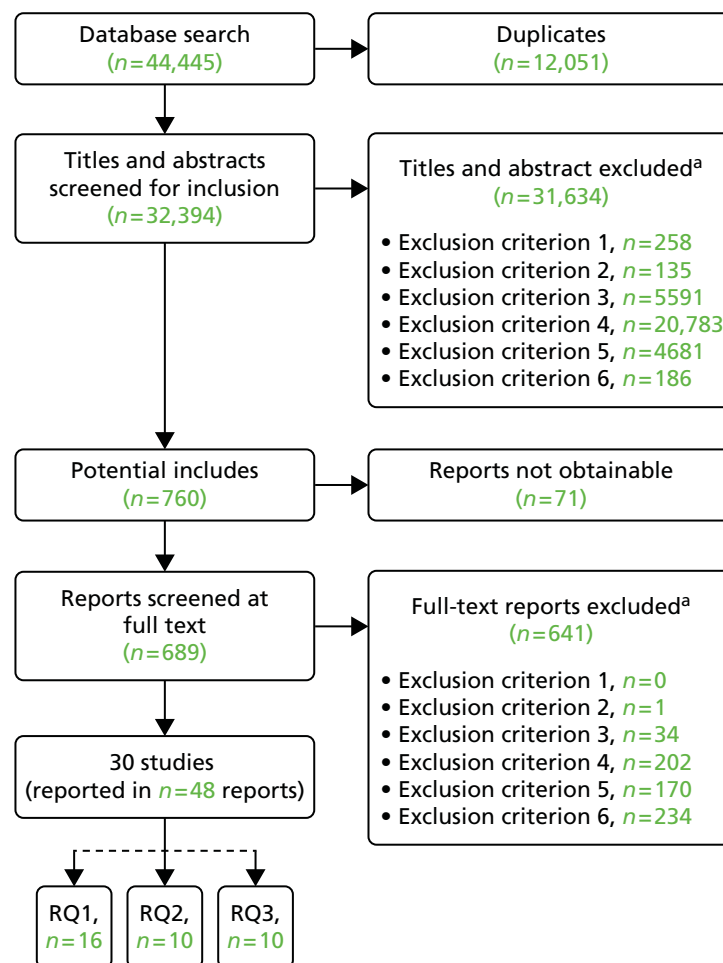


FIGURE 1 Flow of studies in the review. a, The exclusion criteria 1–6 are listed in *Appendix 3*.

TABLE 2 Overview of interventions subject to process or outcome evaluation included in this review

Interventions examined in the review	Included theory studies	Included process evaluations	Included outcome evaluations
Supervised ASP		Armstrong and Armstrong ⁸⁴	
A violence, delinquency and substance abuse prevention programme		Baker <i>et al.</i> ⁸⁵	
YARP	Berg <i>et al.</i> ⁸⁶	Berg <i>et al.</i> ⁸⁶	Berg <i>et al.</i> ⁸⁶
Chicano Latino Youth Leadership Institute		Bloomberg <i>et al.</i> ⁸⁷	
Stand Up Help Out: leadership development ASP		Bulanda and McCrea ⁸⁸	
All Stars prevention curriculum: an enhanced ASP		Cross <i>et al.</i> ⁸⁹	Cross <i>et al.</i> ⁹⁰
		Gottfredson <i>et al.</i> ⁹¹	Gottfredson <i>et al.</i> ⁹²
			Gottfredson <i>et al.</i> ⁹¹
			Gottfredson <i>et al.</i> ⁹³
MAPs			
Community Youth Initiative		Lee <i>et al.</i> ⁹⁴	
Cool Girls, Inc.			Kuperminc <i>et al.</i> ⁹⁵
BBBS			Rhodes <i>et al.</i> ⁹⁶
			Grossman and Tierney ⁹⁷
			Tierney ⁹⁸
QOP		Maxfield <i>et al.</i> ⁹⁹	Rodriguez-Planas ⁷⁸
		Maxfield <i>et al.</i> ¹⁰⁰	Maxfield <i>et al.</i> ¹⁰⁰
			Rodriguez-Planas ¹⁰¹
			Rodriguez-Planas ¹⁰²
			Schirm <i>et al.</i> ¹⁰³
			Schirm and Rodriguez-Planas ¹⁰⁴
			Schirm <i>et al.</i> ¹⁰⁵
NGYCP		Schwartz <i>et al.</i> ⁷⁷	Schwartz <i>et al.</i> ⁷⁷
		Bloom <i>et al.</i> ¹⁰⁶	Millenky <i>et al.</i> ¹⁰⁷
			Millenky <i>et al.</i> ¹⁰⁸
			Millenky <i>et al.</i> ¹⁰⁹
			Perez-Arce <i>et al.</i> ¹¹⁰
PYDC			Tebes <i>et al.</i> ⁷⁹
Stay SMART programme			St Pierre and Kaltreider ¹¹¹
YPDP		Wiggins <i>et al.</i> ¹¹²	Wiggins <i>et al.</i> ¹¹²
			Wiggins <i>et al.</i> ⁴⁶

BBBS, Big Brothers Big Sisters of America; MAP, Maryland after-school programme; PYDC, Positive Youth Development Collaborative; YARP, Youth Action Research for Prevention; YPDP, Young People's Development Programme.

Study characteristics

A descriptive overview of the 30 studies and 48 study reports included in the review is provided below. It includes details of the rate of publication, geographical location of study publication, the age group targeted by PYD programmes and the age of actual sampled populations.

Rate of publication

Figure 2 provides a breakdown of rate of publication, based on the 48 reports included in the review according to which RQ they answered. The data indicate that there was an increase in studies being published in this area from 2002–2011 ($n = 37$) reaching a peak in 2003 ($n = 5$) and another in the period 2009–2011 ($n = 6$ per year) with few studies published before 2001 ($n = 6$).

Geographical location

Figure 3 provides a summary of the geographical origin of studies. Of the 30 included studies, the vast majority were written and published in the USA ($n = 26$), reflecting the origins of the programme. Only a few studies were conducted in other countries; one each from the UK, Australia, Canada and the Hong Kong province of China.

Target age group of Positive Youth Development programmes

The majority of process and outcome evaluation studies ($n = 15$) focused on programmes delivered to children and young people under the age of 15 years ($n = 8$) rather than to those over 15 years of age ($n = 4$) (Figure 4). Three process evaluations did not explicitly report the age group of their target population of interest, simply stating that programmes were eligible to all young people or for those young people considered to be 'at risk' in their community.

Sampled population

Figure 5 summarises the populations sampled. The outcome evaluations sampled the same population targeted by the programmes under investigation. However, as would be expected, the studies containing process data sampled both programme providers ($n = 7$) and young people ($n = 7$). Of the seven studies with process data sampling young people, three sampled those under 15 years of age and four sampled those aged over 15 years. The remaining two studies failed to provide a breakdown of the age of young people in their sample.

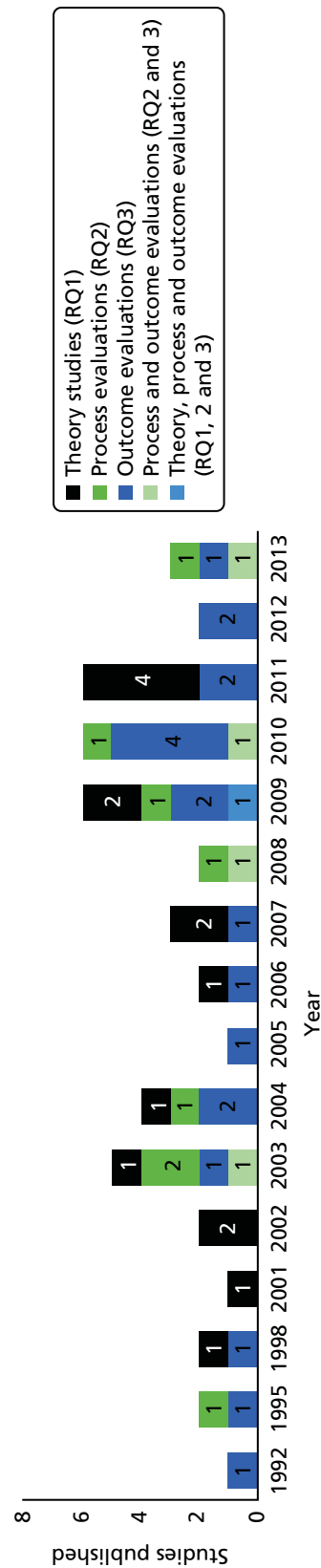


FIGURE 2 Rate of study publication.

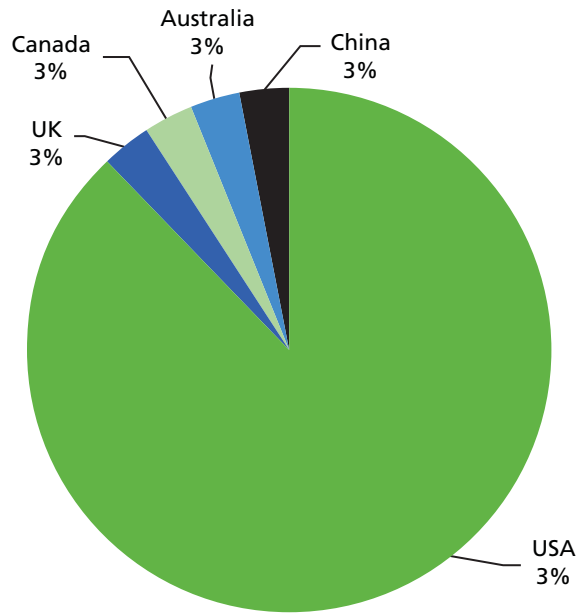


FIGURE 3 Spread of studies by country (mutually exclusive).

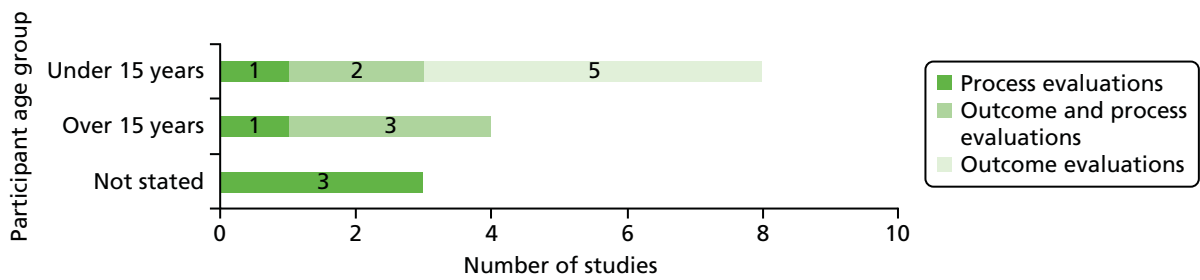


FIGURE 4 Target age group of programmes according to study type (mutually exclusive).

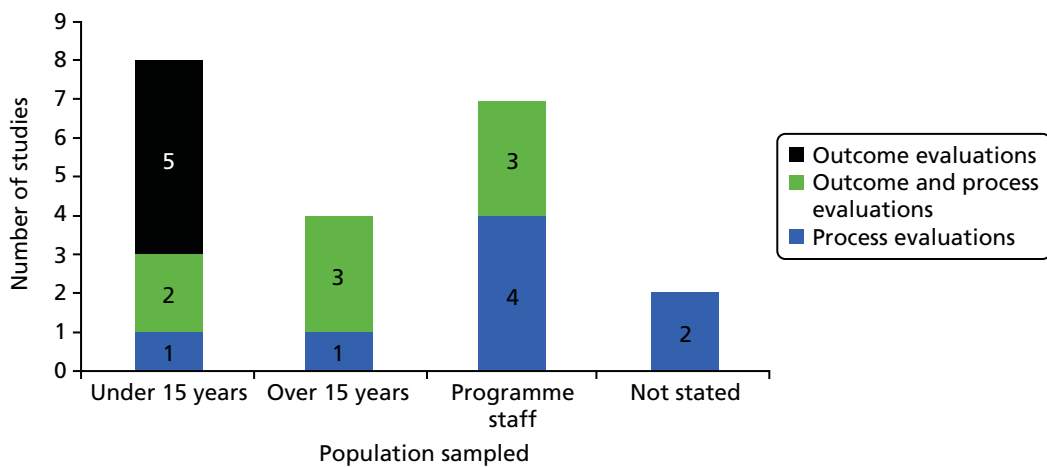


FIGURE 5 Population sampled in outcome and process evaluations (not mutually exclusive).

Chapter 4 Synthesis of theories of change

About this chapter

This chapter describes and reports the quality of the included studies which describe theories of change for PYD interventions. It also reports our thematic synthesis of these studies, which aimed to produce a taxonomy and theory of change of PYD interventions and which uses this synthesis to develop some hypotheses about the factors that might moderate the effectiveness of PYD interventions. Characteristics of theory studies and an assessment of their quality are tabulated in *Appendix 13*.

Included studies

A total of 16 reports were judged to set out a theory of change for PYD and were thus included.^{50,86,113–126} One report was led by Canadian authors¹¹⁷ and one by authors from Hong Kong.¹²² All other reports were led by authors from the USA.^{50,86,113–116,118–121,123–126} One report which described theory of change was also included in our review of process and outcome evaluations,⁸⁶ but all other theory reports did not report on empirical evaluations of PYD interventions. Four reports were led by Benson^{113–116} and two by Lerner and Lerner.^{123,124}

During the process of coding studies, reviewers determined that the included reports engaged in two types of theorising, namely 'normative' (what things should look like) and causal (how things are causally interconnected). Nine of the reports largely focused on developing normative theory,^{113–116,119,122–125} which asserted the value of PYD in focusing on the development of assets rather than merely the prevention of risk behaviours, albeit with some discussion of causality too. Six reports largely focused on causal theory, which aimed to describe the mechanisms by which PYD might promote benefits,^{50,117,118,120,121,126} and one report gave attention to both normative and causal elements.⁸⁶ Rather than limit synthesis to causal theories, we decided to broaden this to include normative theorising because we felt that this enabled a clearer specification of the assumptions that lie behind PYD interventions. Most of the reports drew on established psychological theories to describe and/or assert the value of PYD. These included bio-ecological theory,¹²⁷ referenced by Benson *et al.*;¹¹³ social learning theory,¹²⁸ referenced by Berg *et al.*;⁸⁶ and identity development theory,^{129,130} referenced by Busseri *et al.*¹¹⁷ We identified only one novel theory that was explicitly presented as a theory of PYD, namely that developed by Benson *et al.*¹¹³ called 'developmental assets theory'. This is predominantly a normative theory rather than a causal theory, in that it asserts the importance of developing positive youth assets rather than describing a causal theory of change explaining how PYD interventions might achieve their impacts.

Although all studies gave some attention to how PYD might reduce risk behaviours, only nine studies did so in any depth,^{50,86,115,118,120–122,125,126} and, of these, only Catalano *et al.*¹¹⁸ presented a comprehensive assessment of the causal mechanisms involved. Thus, many studies only narrowly met the inclusion criteria for this element of our review and did not present detailed causal pathways by which PYD interventions might reduce substance use and violence. Nonetheless, our synthesis of these studies succeeded in drawing on these disparate reports to develop a synthesis which offers a somewhat more comprehensive analysis of how PYD intervention might aim to reduce these risk behaviours than is presented in any single included study.

Quality of studies

Assessing the quality of theory reports proved challenging despite the use of criteria informed by previous literature.^{64,131} As shown in *Appendix 13*, there was very little agreement on the scores given by the two reviewers and, thus, with the exception of three studies, we decided not to develop an overall agreed score. The utility of the assessment of the quality of theory reports is considered further in *Chapter 7*.

Synthesis of theories of Positive Youth Development

The two reviewers differed in how they coded and interpreted the theoretical literature (see coding templates in *Appendix 9*). One attempted to bring the theories together to create an overarching theory of change from PYD intervention to risk reduction, even where the lack of clarity and attention to the question of risk reduction in primary sources did not make this task very easy. The other reviewer focused on highlighting that the PYD theoretical literature gives more attention to asserting what PYD is and why it is superior to prevention science, rather than on theorising the causal pathways by which PYD interventions promote positive assets or how these then help to reduce risks. This reviewer also focused on how the PYD literature did not put forward a novel theory of change but rather relied on a number of existing psychological theories. Nonetheless, it was possible to produce an overall analysis which drew on these different elements, and this is summarised below.

The original focus for the synthesis was to define, produce a taxonomy of and identify a theory of change for PYD interventions. However, our thematic analysis led us to expand our examination of PYD definitions to develop a synthesis of what we term a 'normative' theory of PYD in addition to our focus on causal theories of change. Normative theories set out ideal conditions, or as Sayer¹³² argues, what alternative, better, social conditions might be possible. In addition, our synthesis examined the potential generalisability of PYD and the factors that might moderate its impacts.

Normative theory of what is Positive Youth Development

Within the theme of normative theory, we identified subthemes including the principles and assumptions which underlie PYD; a normative vision of what positive development should look like for young people and their environments; and the required characteristics of PYD interventions.¹³³

Principles/assumptions that underlie Positive Youth Development

Included reports suggested a number of principles and assumptions that appeared to underpin PYD interventions.

All young people have potential

The first assumption of the PYD literature is that all young people have the potential to develop healthily rather than this being dependent on innate or other fixed factors. This assumption was at the heart of a number of concepts and assertions made in included reports. Lerner and Lerner,¹²³ for example, suggest that PYD 'conveys the adults' beliefs in youth as resources to be developed rather than as problems to be managed' (Roth and Brooks-Gunn,¹³⁴ cited in Lerner and Lerner¹²³), a sentiment echoed by Kim *et al.*¹²¹ and Schwartz *et al.*¹²⁶ In addition, Schwartz *et al.*¹²⁶ and Lerner *et al.*^{123,124} assert that PYD is possible because young people's development is subject to 'plasticity'.¹²⁶ Benson *et al.*¹¹⁴ and Benson and Scales¹¹⁵ similarly set out an assumption that all youth have the inherent capacity for optimal development given appropriate opportunities and ecologies. Berg *et al.*,⁸⁶ Kia Keating *et al.*,¹²⁰ Lerner and Lerner,¹²³ Roth and Brooks-Gunn⁵⁰ and Lee¹²² all describe how hope is central to the atmosphere of PYD programmes. Lerner and Lerner¹²³ and Kim *et al.*¹²¹ stress that the optimism expressed by adults in the strengths and potentials of young people is an important contributor to the fostering of hope. This might stand in contrast to traditional prevention programmes that explicitly present to young people the belief that they may experience adverse outcomes.

Amelioration is required because of deteriorating environments

Another assumption, apparent within reports by Benson *et al.*¹¹⁴ and Kim *et al.*,¹²¹ is that PYD interventions are necessary because of deteriorations in young people's environments. Addressing this matter from a conservative political position, Benson *et al.*¹¹⁴ refer to the erosion since the 1960s of what they describe as the 'traditional' supports for youth development, such as stable families and communities characterised by intergenerational relationships and good-quality school systems.¹¹⁴ Kim *et al.*¹²¹ similarly refer to deficits in parenting arising from social mobility and transiency which erode neighbourhood stability.¹²¹ Roth and Brooks-Gunn⁵⁰ state that 'when circumstances prevent . . . families, schools, and communities from providing their youth with fundamental . . . resources, youth development programs offer one avenue for fulfilling these needs'. Moreover, Kim *et al.*¹²¹ suggest that the separation of youth from adults in modern society is denying young people adult models and potentially increasing the alienation and isolation of young people.

In considering deteriorating social contexts, only a minority of authors emphasise increasing social inequalities; Berg *et al.*⁸⁶ and Ginwright and Cammarota¹¹⁹ both describe young people's environment as toxic to healthy development because of inequalities relating to SES, ethnicity and sex. Ginwright and Cammarota¹¹⁹ criticise what they regard as mainstream ideas of youth development in their neglect of the inequalities present in existing social arrangements. These authors maintain that it is not sufficient to understand challenges to young people's development only in terms of the individual, their family and community and that attention must also be given to the larger social and economic forces that impact on and limit healthy development.

Normative vision of what positive development should involve

Another subtheme within our coding template concerns what positive development should involve.

Thriving and positive assets

The major emphasis across nearly all the literature was on the importance of enabling young people not merely to reduce problem behaviours and adverse outcomes but to achieve 'normal development'.⁵⁰ This is emphasised for example by Roth and Brooks-Gunn,⁵⁰ Schwartz *et al.*¹²⁶ and Lerner and Lerner.¹²³ Kim *et al.*,¹²¹ Catalano *et al.*,¹¹⁸ Busseri *et al.*¹¹⁷ and Roth and Brooks-Gunn,⁵⁰ for example, all state that 'problem free is not fully prepared'⁵⁰ – in other words that a young person needs more than merely to avoid risk in order to thrive. Lerner and Lerner¹²³ and Kim *et al.*¹²¹ all contrast PYD with the risk factor approach characteristic of prevention science, arguing that PYD defines the positive assets that should be promoted and not merely the risks and problems that should be avoided. Benson *et al.*¹¹⁴ describes PYD as a 'strength-based approach':

The theory and research undergirding developmental assets and asset-building community are designed, in part, to reframe the targets and pathways of human development around images of strength and potential. We posit that this shift is crucial for mobilising both personal and collective efficacy on behalf of child and adolescent development. By so doing, we ultimately seek to balance paradigms so that communities pursue deficit reduction and asset building with equal vigour.

*Benson et al.*¹¹³

Kia-Keating *et al.*¹²⁰ argue that PYD is also distinct from the resilience literature in that PYD focuses more on what is actually required for thriving (whether adversity is experienced or not), whereas the resilience literature often continues to emphasise the avoidance of harms in the face of adversity. Roth and Brooks-Gunn,⁵⁰ Busseri *et al.*,¹¹⁷ Lerner *et al.*,^{123,124} Schwartz *et al.*¹²⁶ and Perkins *et al.*¹²⁵ all build on this idea of thriving by suggesting what particular 'assets' PYD might aim to develop, terming these the '5 Cs':

1. competence in academics, social, emotional and vocational areas
2. confidence in who one is becoming (identity)
3. connection to self and others
4. character that comes from positive values, integrity and a strong sense of morals
5. caring and compassion.¹²⁵

The Search Institute, Benson *et al.*^{113,114,116} and Roth and Brooks-Gunn⁵⁰ propose an alternative categorisation of 40 assets. These comprise 20 'internal' assets (in four groups) which relate to young people's positive development and 20 'external' assets (also in four groups) which describe the assets that environments should possess to enable young people's positive development.

Internal (individual) assets include:

- 'commitment to learning (achievement motivation, school engagement, homework, bonding to school, reading for pleasure)
- positive values (caring, equality and social justice, integrity, honesty, responsibility, restraint)
- social competencies (planning and decision-making, interpersonal competence, cultural competence, resistance skills, peaceful conflict resolution skills) and
- positive identity (personal power, self-esteem, sense of purpose, positive view of personal future)¹¹³

External (environmental) assets include:

- 'Support (family support; positive family communication; other adult relationships; caring neighbourhood; caring school climate; parent involvement in schooling)
- Empowerment (community values youth; youth as resources; service to others; safety)
- Boundaries and expectations (family boundaries; school boundaries; neighbourhood boundaries; adult role models; positive peer influence; high expectations)
- Constructive use of time (creative activities; youth programmes; religious community; time at home)^{113,114,116}

Contribution or challenge to social order?

Within the subtheme of what positive development should look like, we found a major division in the literature. Most authors, including Benson *et al.*¹²³ and others from the Search Institute, argue that positive development should enable young people to make positive contributions to their communities and to society as they currently exist, which are considered broadly benign. Kim *et al.*,¹²¹ drawing on the social control theory of Hirschi,¹³⁵ emphasises the importance of the young person bonding to the conventional social order.

However, Berg *et al.*⁸⁶ and Ginwright and Cammarota¹¹⁹ instead promote the ideal young person as developing an understanding of, and challenging, the inequalities present in the existing social order, in order to transform both themselves and their environment. They propose that the 'social justice model of youth development'¹¹⁹ should promote youth's agency to challenge oppression, enabling young people to develop 'critical consciousness' informed by the work of Freire¹³⁶ on pedagogies of resistance.

The required characteristics of Positive Youth Development interventions

Another subtheme within our theme of normative theory is the required characteristics of PYD interventions.

Affective relationships with adults

Several authors, such as Roth and Brooks-Gunn,⁵⁰ Benson and Scales,¹¹⁵ Lerner and Lerner¹²³ and Kim *et al.*,¹²¹ argue that, unlike what they regard as conventional services for young people, PYD providers must develop 'affective' and not merely 'instrumental' relationships with participants.¹²¹ Roth and Brooks-Gunn⁵⁰ argue for the importance of adults creating a 'family-like atmosphere'.

Diverse activities and settings

Many of the included reports asserted that PYD interventions should offer diverse activities and settings for participants to engage with. Lerner and Lerner¹²³ and Benson *et al.*¹¹⁴ suggest that such activities are important in enabling young people to develop skills. Roth and Brooks-Gunn⁵⁰ argue that activities also provide opportunities for recognition. Benson *et al.*¹¹⁴ suggest that PYD activities should involve 'synergies' and 'redundancies'; in other words, PYD interventions should provide young people with activities that are diverse and that enable the development of common assets in the course of different activities, so that these are complementary and mutually reinforcing in promoting participants' positive development.

Active participation

Benson *et al.*¹¹⁴ and Lerner and Lerner¹²³ both emphasise young people's active participation in their lists of key features of PYD. Roth and Brooks-Gunn⁵⁰ identify youth empowerment as a key objective, whereas Lerner and Lerner¹²³ argue that PYD should provide opportunities for youth leadership. Lerner and Lerner,¹²³ Benson *et al.*¹¹⁴ and Roth and Brooks-Gunn⁵⁰ also emphasise young people's active involvement in providing services to others via 'service learning'.⁵⁰ The more radical strand within the literature, demonstrated for example by Ginwright and Cammarota,¹¹⁹ emphasises that young people should become active in developing their own critical analysis of the environment, understanding how institutions might better serve their communities and developing campaigns to make these more responsive to local needs.

Taxonomy of Positive Youth Development

The final stage of our synthesis of 'normative theories' was the development of a new taxonomy of different models of PYD based on the similarities and differences identified via the thematic analysis. No included report offered an explicit overall taxonomy, but various authors identified particular forms of intervention, which we also drew on in developing our own taxonomy.

Challenging or contributing to the existing social order

Following on from the earlier subtheme of contributing to or challenging the existing social order, we might identify a distinction between 'radical' and 'mainstream' PYD programmes. For example, Berg *et al.*⁸⁶ draw on the theorists such as Gramsci,¹³⁷ Bourdieu,¹³⁸ Foucault¹³⁹ and Freire¹⁴⁰⁻¹⁴² to develop an intervention which uses a participatory action research approach to enable young people to develop a critical consciousness of, and to challenge, oppressive forces such as discrimination and racism within society. This stands in contrast to the majority of descriptions of PYD which emphasise the promotion of young people's bonding to, and contribution towards, reproducing current societal arrangements.

Collectivist or individualistic orientation

In addition, informed by our earlier subtheme of PYD addressing individual and/or collective assets, we can contrast PYD where it is depicted as an intervention to promote individual development versus characterisations of PYD as a collective endeavour. For example, Schwartz *et al.*¹²⁶ argue that, for some communities, particularly in some non-western cultures, interventions need to place less emphasis on individual development and more emphasis on group empowerment. Interventions that emphasise individual agency may be less useful in some cultural contexts than others. As noted above, Berg *et al.*⁸⁶ make a similar argument, but one informed by a belief in the political importance of collective action.

Focusing on the individual and/or the environment

Roth and Brooks-Gunn⁵⁰ contrast intervention activities that aim to modify both an individual young person as well as the context within which they live with those that focus only on the former. Referring to the asserted importance of PYD compensating for limitations in the environment within which young people develop, Roth and Brooks-Gunn⁵⁰ acknowledge that the capacity for compensation by any one PYD programme is limited and what is needed is for PYD principles to inform the transformation of wider environments and services. They contrast PYD which aims to 'prepare adolescents for the world by ensuring that they possess the 5 Cs, [versus that which aims] to shape a better world for youth by also increasing the supports available to them at home, school, and in their community'. Making a similar point from within the more radical strand of PYD theorising, Ginwright and Cammarota¹¹⁹ argue that a lack of focus on social, economic and political forces is a major limitation in mainstream PYD.

Breadth, depth and duration of activities

Another means of categorising PYD interventions is in terms of the breadth, depth and duration of activities. Roth and Brooks-Gunn⁵⁰ report that programmes vary in which, and how many, assets they seek to enhance. Busseri *et al.*¹¹⁷ distinguish interventions that offer a breadth of participation in many activities and those that offer more in-depth participation in fewer activities. Busseri *et al.*¹¹⁷ argue that breadth of activities might be more important in early adolescence when young people are experimenting with

different identities, whereas in-depth participation in selected activities may become more important in later adolescence as identities begin to stabilise during the transition to adulthood.

Prevention and promotion

Finally, we can distinguish between PYD programmes that aim only to promote positive assets and those that also aim to address risks. Schwartz *et al.*¹²⁶ suggest that the prevention of problem behaviours and the promotion of thriving can occur together: 'prevention science and positive youth development may appear to represent incompatible approaches at worst . . . or complementary approaches at best'.

Catalano *et al.*¹¹⁸ and Schwartz *et al.*¹²⁶ argue that positive assets and protective factors for the prevention of risk are in practice often the same thing, with communication skills and self-awareness being both positive assets in themselves as well as assets that provide protection against risk behaviours such as violence and substance use. In such cases, PYD interventions might reduce risks without explicitly addressing them. In other cases, PYD programmes might go beyond the promotion of positive assets so that they also address protective factors which are not themselves positive assets, such as lack of access to tobacco, as described by Catalano *et al.*¹¹⁸

Causal theory of change for Positive Youth Development

We now explore the theme of how PYD is intended to impact, first, on positive assets and, second, on reducing risk behaviours. Here, we are thus considering causal theories, which describe how programmes work to deliver their goals.¹⁴³

Positive Youth Development promoting intentional self-regulation

A key subtheme across a number of studies^{86,113–117,120,123–126} is how PYD promotes 'developmental regulations'.¹²⁴ This notion of developmental regulations is said to be rooted in ecological systems theory, a widely used theory of how individual development is shaped by social contexts.¹⁴⁴ Lerner *et al.*¹²⁴ define developmental regulations as bi-directional interactions between the individual and their social environment (such as their peer group, family, school and neighbourhood):

These bidirectional relations may be represented as individual/context relations. These relations regulate (that is, govern) the course of development (i.e., its pace, direction, and outcomes). When these 'developmental regulations' involve individual/context relations benefiting both the person and his or her ecology, they may be termed 'adaptive'.

But how might PYD interventions enable developmental regulations to occur? Busseri *et al.*,¹¹⁷ Schwartz *et al.*¹²⁶ and Lerner *et al.*¹²⁴ all suggest that PYD improves young people's capacities for 'intentional self-regulation'.¹²⁴ According to Lerner *et al.*:¹²⁴

[I]ntentional self regulation may involve the selection of positive goals (e.g. choosing goals that reflect important life purposes), using cognitive and behavioral skills (such as executive functioning or resource recruitment) to optimize the chances of actualizing ones purposes and, when goals are blocked or when initial attempts at optimization fail, possessing the capacity to compensate effectively.

Benson¹¹³ similarly suggests that PYD can help adolescents to reflect on existing behaviour, select personal goals and activities through which to pursue these, and apply available resources to pursue these goals and activities. Lerner *et al.*¹²⁴ suggests that if PYD programmes promote young people's senses of hope and their ability to intentionally self-regulate and provide them with opportunities and resources such as relationships and education, then positive 'development regulations'¹²⁴ will occur. Lerner *et al.*,¹²⁴ as well as Benson,¹¹³ suggest that this will enable virtuous circles to occur which enable positive development in the individual as well as improvements in their social environments via the young person's positive contribution to these.

However, although these authors argue that such processes of intentional self-regulation are required in order for young people to accrue positive assets, they are not explicit about how PYD interventions actually enable young people to develop their capacity for intentional self-regulation. Benson¹¹³ suggests that PYD interventions can optimise these processes by:

[I]ncreasing the developmental attentiveness of contexts . . . to increase their capacity to nurture, support, and constructively challenge the developing person, . . . [E]nhancing the skills and competencies of youth (to further enable their 'natural' capacity to engage with, connect, change, and learn from their social contexts . . . [and [c]reating processes and opportunities to invite youth to actively exercise and utilize their capacity to engage with and change their social contexts

But this quotation, as well as the theoretical literature more generally, stops short of explaining exactly how PYD programmes might promote better intentional self-regulation either with regard to specific intervention activities or more generally. However, informed by the claims in reports synthesised within our normative theme above (e.g. by Benson *et al.*¹¹⁴), we might fill in the gaps, speculating that PYD could promote general improvements in intentional self-regulation by providing individuals with the resources, in the form of relationships and training in specific skills, that are the critical inputs upon which the use of intentional self-regulation within specific intervention-related activities depends and by enabling individuals to practise intentional self-regulation in the context of multiple activities and settings so that they improve their general ability to intentionally self-regulate.

Busseri *et al.*¹¹⁷ come closest to making this explicit. They argue that PYD can provide a range of 'affordances' (opportunities) to which young people may respond. Different individuals, with different needs and goals at different points in their maturation, will make use of different affordances. Busseri *et al.*¹¹⁷ argue that PYD should provide diverse opportunities for young people to explore different activities and then to commit to pursuing a narrower range of activities in depth. The extent to which a young person commits to an activity and to the development of specific assets may indicate growing stability in an identity.

A final way in which PYD interventions might help young people to use intentional self-regulation in order to develop positive assets is by refocusing existing capacities for intentional self-regulation away from antisocial goals and towards pro-social goals. This might occur via rewarding young people when they abandon antisocial activities and engage in specific pro-social activities. Kim *et al.*,¹²¹ for example, refer to social learning theory (Akers *et al.*¹⁴⁵) to suggest that, by providing positive examples and celebrating achievements in the realm of pro-social activities, PYD programmes can reinforce positive behaviours and bonds to conventional society.

Theories of risk reduction

Despite being the central question informing our synthesis, how interventions reduce risk behaviours such as substance use and violence was a relatively minor theme in the PYD literature, a point acknowledged by Kia-Keating *et al.*¹²⁰ and Lerner *et al.*¹²⁴ Lerner *et al.*¹²⁴ for example recognise that 'the PYD perspective needs to . . . focus more on what seem to be quite diverse and intricately linked pathways of thriving and of risk/problem behaviours'. The PYD literature offered a number of generic suggestions about how the development of positive assets might reduce risks, as well as a few isolated examples of how specific assets might protect against risk, but stopped short of offering a comprehensive theory of change for how PYD reduces risk.

Buffering of and compensating for risk

One subtheme was that positive assets might reduce risk by generic processes of 'buffering' (Catalano *et al.*¹¹⁸) or 'compensation' (Busseri *et al.*¹¹⁷). The use of the term 'compensation' is different here to that referred to above when describing 'compensation' as one element in intentional self-regulation. For example, Catalano *et al.*¹¹⁸ described buffering as a process whereby risk factors in the environment have less influence on the behaviour of those with positive assets than on those who lack these assets. This possibility is also raised using similar terminology by Kia-Keating *et al.*¹²⁰ As an example, Catalano *et al.*¹¹⁸

suggest that where individuals have developed a positive sense of identity as a positive asset, they may be better able to make choices based on 'internal values' and thereby be less prone to involvement in risk behaviours as a result of peer pressure. Now considering compensation, Catalano *et al.*¹¹⁸ also describe a generic process whereby those possessing positive assets may engage in risk but possession of these assets reduces the harmful consequences of the risks. They offer the example of a young person who engages in violence but possesses school engagement as a positive asset and, because of this, the harmful developmental effects of violence are offset by their educational attainment. Schwartz *et al.*¹²⁶ refer to similar processes using slightly different terminology.

Pile-up and molecular impacts on risk

Another subtheme in understanding how PYD might in general reduce risk centres on 'pile-up' and the 'molecular' impacts of positive assets.¹¹⁴ Benson *et al.*¹¹⁴ argue that assets can reduce risk behaviours non-specifically via the protective pile-up of multiple non-specific assets or by the molecular effects of specific assets on specific risk behaviours. Catalano *et al.*¹¹⁸ suggest that, because of their focus on the development of multiple positive assets, PYD interventions are likely to address multiple determinant of problem behaviours. However, this theoretical strand does not offer further detail about the specific pathways via which particular or multiple positive assets might protect against substance use and violence risks.

However, without explicitly engaging with the notion of molecular impacts, a few authors offer suggestions about how specific assets may offer protection. Kim *et al.*¹²¹ argue that engagement with pro-social institutions, as a positive asset, will lead to reductions in antisocial behaviours. Other authors simply report empirical evidence of correlations between certain specific assets and reductions in risk behaviours. Benson and Scales¹¹⁵ for example report empirical research suggesting that positive peer influence, peaceful conflict resolution and school engagement are correlated with reduced involvement in substance use and violence. They argue that the mechanisms involved might be direct (e.g. young people developing skills for resolving conflict) or indirect (e.g. young people developing connections to pro-social peers and other individuals who model responsible behaviours, and to pro-social institutions which militate against young people defying pro-social norms). Kia-Keating *et al.*¹²⁰ suggest that social integration and self-efficacy will offer protection against conduct problems; that pro-social behaviours will protect against substance use; that emotional self-regulation will protect against externalising problem behaviours; that adult supervision will protect against delinquency; and that self-efficacy can interact with parental monitoring to protect against alcohol use. However, the literature here is offering piecemeal suggestions rather than any comprehensive, systematic theory about how positive assets will lead via pile-up or molecular impacts to reductions in substance use and violence that are not already apparent within conventional prevention science.

The generalisability of Positive Youth Development theories of change

An additional theme discernible in the included reports concerns the extent to which normative and causal theories of PYD might be generalisable across cultures. This was raised in several papers. Benson,¹¹³ Lerner and Lerner,¹²³ Schwartz *et al.*¹²⁶ and Kia-Keating *et al.*¹²⁰ all suggest using similar terminology, that the precise form that specific development assets and specific forms of individual–environment interactions will take may vary between cultures but that the basic framework of PYD is potentially generalisable. Benson *et al.*,¹¹⁴ for example, suggest that what constitutes resilient, competent, protective or connected is geographically, temporally, culturally and socioeconomically bounded. For Schwartz *et al.*,¹²⁶ the main divide is that between cultures that value individualism and cultures that value more collectivist societies. They suggest that how thriving is defined may differ in cultures with different emphases on the individualism and collectivism.

Furthermore, Benson *et al.*¹¹⁶ contend, based on empirical research (albeit only that conducted in different contexts within the USA), that, although absolute levels of assets vary geographically, associations with positive developmental outcomes are similar across samples by sex, ethnicity, geography and SES. Benson and Scales¹¹⁵ and Schwartz *et al.*¹²⁶ both suggest that, although the overall PYD framework is generalisable, different specific strategies and tactics will be required for different groups and settings. For example, Benson *et al.*¹¹⁴ cite evidence that stricter parenting is protective for urban African Americans but not for white children in less stressed contexts.

However, Ginwright and Cammarota¹¹⁹ go against this body of opinion, arguing that mainstream PYD theories treat young people as homogeneous and resembling middle-class white US youth and fail to engage with how identity might vary among other groups.

Factors potentially moderating the impact of Positive Youth Development

Our synthesis also aimed to examine what included theory reports had to suggest about how the effectiveness of PYD interventions might be moderated by various factors. A number of disparate suggestions from the literature can be identified.

Age

Busseri *et al.*¹¹⁷ argue that the transitional status of the young person is an important moderator. During stages in the developmental cycle in which there is a high level of transition, the young person will benefit from a wide range of activities to explore their identity. When they reach a more steady state, the opportunity to participate more intensely in fewer activities will be of more value.

Duration of intervention

Lerner and Lerner¹²³ suggest that the duration of a relationship between the young person and adult providers in PYD interventions will influence their effects:

[W]hen young people are in relationships that last a year or longer they are most likely to experience improvements but when youth are in relationships that last for only between six and 12 months, fewer positive outcomes occur and when young people are in mentoring relationships that end relatively quickly, it appears that mentoring may actually be detrimental.

Programme features

Lerner and Lerner¹²³ suggest that PYD programme effectiveness will be influenced by clarity of goals; attention to the diversity of youth, families, communities and cultures; assurance that the programme represents a safe space; integration of community values into programmes; collaboration with other youth-serving organisations; engagement of providers with programme evaluation; and use of advocacy for youth. Benson,¹¹³ informed by the accumulation hypothesis, suggests that interventions that address more assets will have more impact.

Baseline risk

Finally, Schwartz *et al.*¹²⁶ discuss the role of baseline risk at both the individual and community levels as a determinant of programme effectiveness. They suggest that in some cases high baseline levels of individual risk may overwhelm the beneficial effect of positive developmental assets and that such assets may not bring the same benefits in poor compared with affluent neighbourhoods.

These potential moderators inform our hypotheses listed below.

Summary of theory synthesis

Taxonomy of Positive Youth Development

Positive Youth Development can be categorised according to:

- whether it aims to challenge or contribute to the existing social order
- whether it aims to promote the development of individuals or groups of individuals
- whether or not it aims to promote the development of the individual and/or modify the wider environment in which the individual develops
- the breadth, depth and duration of activities provided and
- whether or not it explicitly addresses the reduction of risk behaviours.

Our aim was to combine this taxonomy with categories arising from the synthesis of process evaluations and to use the combined categories to describe interventions included in the synthesis of process and outcome evaluations.

Theory of change for Positive Youth Development effects on violence and substance use

Our synthesis of how PYD is theorised to promote positive assets and reduce substance use is summarised in *Figure 6*. PYD interventions are intended to provide positive expectations; enduring and affective relationships with adults; diverse activities and settings; and active participation, whereby young people are empowered to choose activities and to take on responsibilities. Interventions vary according to whether or not they aim to enable young people to contribute to or challenge the existing social order; whether or not they focus on individual and/or collective development; whether or not they address the individual and/or the environment in which they live; the breadth, depth and duration of activities offered; and whether or not they address only positive development or also prevention of risk behaviours.

Positive Youth Development aims to offer ‘affordances’,¹¹⁷ that is, resources that individuals make use of in the course of their development (e.g. relationships, challenges, education). Different individuals will make use of different affordances at different points in their maturation. Young people will make use of these affordances in applying intentional self-regulation to specific intervention-related activities, such as sports, arts or outdoor activities. Intentional self-regulation comprises intentionality (assessing their current skills); selection (setting goals for what they want to achieve); optimisation (using their existing skills and the new resources that PYD provides to achieve these goals); and compensation (reviewing and if necessary redirecting actions to meet goals).¹¹⁷ PYD interventions aim to reward young people when they make progress with activities regarded as pro-social within that programme. Drawing on but filling in gaps in the PYD literature, we suggest that because PYD is thus enabled in a range of diverse but overlapping and mutually reinforcing intervention-related activities, young people develop a more generalised ability to engage in intentional self-regulation focused on pro-social goals.

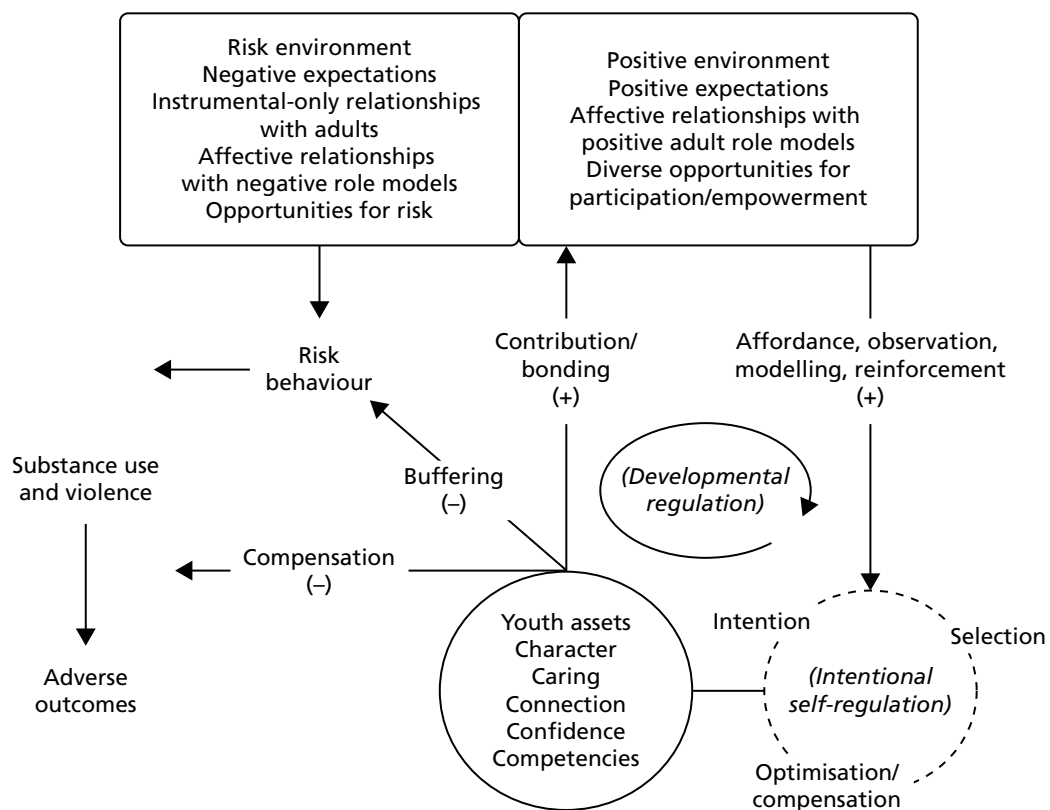


FIGURE 6 Synthesised theory of change for PYD effects on substance use and violence.

As a result of developing intentional self-regulation, young people are better able to accrue a range of 'positive assets' such as the '5 Cs': competence; confidence; connection; character; and caring.¹²⁴ As these accrue, young people can make better use of the opportunities available in their wider environment, which leads to positive 'developmental regulations'¹²⁴ (i.e. a positive feedback cycle, whereby the individuals gain more benefit from opportunities in their environment). This enables them in turn to make a positive contribution to their communities and societies, or, as a few authors would argue, enables them to contribute by critiquing and challenging inequities present in the existing social order.

These positive assets may then reduce risk behaviours via 'buffering',¹¹⁸ whereby risk factors in a young person's environment have less impact, or via 'compensation',¹¹⁷ whereby even if young people engage in a risk behaviour, their possession of positive assets ameliorates the impact of this on their overall development. (This use of the term 'compensation' is distinct from that used within the description of intentional self-regulation above.) Furthermore, positive assets may reduce risk via molecular mechanisms (whereby a specific asset exerts specific protection against a specific risk) and/or via pile-up (whereby an accumulation of multiple assets is protective regardless of the specific assets involved). However, the theoretical literature synthesised here offers only limited insights beyond these general ideas. Engagement with pro-social peer groups or institutions might lead to reductions in antisocial behaviours via greater exposure and adherence to pro-social norms. Improved emotional self-regulation, social skills and self-efficacy might contribute to better decision-making to avoid violence and substance use. However, this falls short of a systematic theory of how the accrual of assets in particular or in general contributes to reductions in substance use or violence.

Hypotheses arising from the theory synthesis

The discussion above suggests the following hypotheses:

- Interventions that offer a breadth of activities may be more effective for younger adolescents, whereas those that emphasise depth may be more effective for older adolescents.
- Interventions that combine prevention and positive development may be less effective than those that focus only on positive development.
- Interventions of more than 1 year's duration may be more effective than those of a shorter duration.
- Interventions may be more effective for participants with low or moderate levels of baseline risk, as there is more scope for stimulating 'intentional self-regulation'.¹²⁴

Our aim was to combine these hypotheses with those arising from the synthesis of process evaluations and assess these in the light of findings from outcome evaluations.

Feedback from consultation on the theory synthesis

The policy and practitioner advisory group and young people provided specific feedback on the theory synthesis and the hypothesis generated from this synthesis (see *Appendix 12*).

Policy stakeholders' feedback

All of the themes from the theory synthesis made sense to policy and practitioner stakeholders. They commented on its usefulness when thinking about current programmes in the UK, particularly the idea of positive assets 'piling up' to support the prevention of risky behaviours. The hypotheses generated from the theory syntheses also made sense and did not require further modifications or refinements.

Young people's feedback

Young people offered three key messages in response to the findings of the theory synthesis. First, to ensure a positive environment, PYD programmes should aim to provide a different experience from the one that young people receive in school. This means ensuring that the setting is 'relaxed' and 'fun', that young people can have relationships with youth workers who are non-judgemental and that opportunities for young people to lead are provided.

Second, the group did not feel that violence/reducing conflict was a current focus of youth work, and the potential of youth work to reduce this was not obvious. However, the opportunity to engage in positive behaviours and with positive adults acting as role models might, as PYD theory reports suggest, plausibly help to reduce the influence of peers in promoting substance use and violence.

Third, projects that do try to reduce risky behaviour and promote positive development may not always work. It might be better to focus on the positive development rather than explicitly addressing risk behaviours, with the hope that the former causes decreases in the latter because of factors such as diversion. There is a risk that too explicit a focus on risk reduction might cause youth work to seem too much like school.

The ALPHA group also strongly agreed with the hypothesis that breadth of activities was important for younger age groups (e.g. 10- to 14-year-olds) but that older teenagers may prefer more regular, 'in-depth' activities. They felt that the ideal length of a project would vary but that this might typically be 1 year.

The above comments from policy/practitioner stakeholders and young people are thus broadly in agreement with the assertions made in the theoretical literature and do not suggest any amendments to the hypotheses above.

Chapter 5 Synthesis of process evaluations

About this chapter

This chapter describes and reports on the quality of process evaluations of PYD interventions that met our inclusion criteria. It also reports our thematic synthesis of these studies and uses this synthesis to refine our taxonomy of PYD interventions and develop further hypotheses about what factors might moderate their effects.

Included studies

Studies and their quality are tabulated in *Appendix 14*. Of the 10 included studies,^{77,84–89,94,99,112} eight were conducted in the USA,^{77,84–88,90,99} one in Australia⁹⁴ and one in England.¹¹² A total of four studies targeted youth aged 14 years or older;^{77,86,88,99} three targeted those aged both above and below 14 years^{87,89,112} and three did not report the age range targeted.^{84,85,94} Four interventions targeted individuals on the basis of individual disadvantage; two targeted individuals on the basis of area or school disadvantage; one targeted individuals on both individual and area disadvantage; and three did not involve targeting on either basis. Three interventions were delivered to participants for ≥ 1 year; three were delivered for < 1 year and six studies did not report their duration. We aimed to describe the interventions using the taxonomy developed from our synthesis of theoretical studies and refined using our synthesis of process evaluations. However, studies did not report sufficient information for this to provide a useful categorisation (see *Appendix 15*).

Armstrong and Armstrong⁸⁴ reported on an intervention delivered in a city in the south-west of the USA involving supervised after-school provision to at-risk youth delivered by staff from the Parks, Recreation, and Libraries Department in after-school sites. The intervention, the duration of which is not reported, aimed to provide life skills, educational support, healthy living skills, social and peer interaction, physical activity, cultural awareness and fine arts and locally relevant programme activities, and was evaluated using in-depth interviews and non-participant observation.

Baker *et al.*⁸⁵ reported on an intervention delivered in Baltimore, MD, USA, involving prevention activities targeting youth at risk of delinquency, violence and substance abuse delivered by South Baltimore Youth Centre. The intervention, of unreported duration, involved provision of a safe space for youth to engage in positive social activities and job training and included case management, mentoring, tutoring and community involvement. The intervention also involved outreach and collaboration with other agencies. It was evaluated using unstructured interviews and focus groups.

Berg *et al.*⁸⁶ reported on an intervention delivered in Hartford, CT, USA, targeting youth aged 14–16 years. The Youth Action Research for Prevention (YARP) intervention, delivered by research educators within a community-based after-school and summer programme, involved participatory action research, that is, formative community ethnography whereby participants were trained to identify adolescent risk behaviours, develop a collective action plan and carry out activities as a group, including using research to understand their community better. The intervention, which lasted for approximately 8 months, was evaluated in a mixed-methods study examining outcomes and processes, with the latter being assessed using interviews with staff, ethnographic observation, youth focus groups and youth self-reflection.

Bloomberg *et al.*⁸⁷ reported on an intervention delivered in Minnesota, USA, targeting Mexican American youth aged 12–17 years. The intervention, of unstated duration, was delivered by a community-based provider and involved support for youth participation in planning and implementing a community service project. Participants engaged in leadership opportunities through conference presentations and specific youth service projects. Evaluation was undertaken through focus groups and youth reports.

Bulanda and McCrea⁸⁸ reported on the Chicago Stand Up Help Out, a leadership development ASP of unstated duration targeting African American youth residing in socioeconomically disadvantaged neighbourhoods and delivered by school social workers and/or social work graduate students in schools and community settings. Participants were offered training and money to participate in an apprenticeship in social work. Activities also included college tours, support for completing curricula vitae, learning about non-violent conflict resolution, mentoring elementary-school children and planning community health and safety fairs. Counselling was also available to young people. Evaluation was undertaken by means of youth reports, round-table discussions, interviews and participant observations.

Cross *et al.*⁸⁹ reported on the All Stars prevention curriculum, an enhanced ASP targeting pupils from under-performing middle schools aged 11–14 years in a city on the east coast of the USA. The intervention, delivered by a county-level government agency specialising in providing recreation and leisure activities for youths, was provided for 3 days per week for 3 hours after school for an unstated period of time. It involved fitness activities, board games, arts and crafts, field trips, computer projects or computer free time, service learning, workforce skills and holiday or other special event celebrations. The evaluation drew on routine documents and data, observations and a survey of participants.

Lee *et al.*⁹⁴ reported on a community-driven youth initiative delivered in an Australian rural community to all young people in the community by a combination of indigenous and other staff. Activities included technical studies, film-making, a driving licence programme, marine debris clean-up, youth leadership opportunities, recruitment, support and training for new employees in the local youth recreational sector, youth and community festivals, a mural painting programme, sporting carnivals, a cultural knowledge programme, discos, mental health promotion posters and health promotion. The intervention, the duration of which is not reported, was evaluated by drawing on interviews, document analysis and staff diaries.

Maxfield *et al.*⁹⁹ reported on the QOP delivered across the USA to educationally disadvantaged youth by schools and community-based organisations. This ASP involved intensive case management, mentoring and educational, developmental, cultural and recreational and community-based activities delivered for up to 5 years. Evaluation was via observational site visits.

Schwartz *et al.*⁷⁷ reported on the NGYCP Youth-Initiated Mentoring programme, which was delivered across the USA to youth aged 16–18 years who had dropped out or been expelled from school. Delivered in a military-style boot camp, the intervention involved an initial 2-week orientation and assessment followed by a 20-week residential phase involving classes on academic learning, life skills, health and job skills, as well as other activities such as physical training, sports, leadership and citizenship activities and community service. A post-residential phase involved participants setting and fulfilling their own development plan which involved education, vocational training or employment. The intervention lasted for over 1 year and was evaluated using a mixed process and outcome design, the former involving semistructured interviews.

Wiggins *et al.*¹¹² reported on the Young People's Development Programme (YPDP), an intervention delivered in disadvantaged areas of England and targeting young people aged 13–15 years at risk of teenage pregnancy, substance use or school drop-out. Delivered in youth centres after school for 1 year, the intervention included activities focusing on young people's health and education as well as on their broader social development with specific programme content determined by the individual projects delivering services, including education, training/employment opportunities, life skills, mentoring, volunteering, health education, arts and sports, and advice on accessing services. The evaluation assessed outcomes and processes, the latter via routine monitoring data, questionnaires and interviews with young people, staff questionnaires and focus groups, and observations.

Quality of process evaluation methods

Reliability of process evaluations

Sampling

There was a general lack of studies reporting information on sampling methods. Three studies were judged to have taken steps to ensure sampling rigour.^{77,84,112} These studies attempted to sample a diverse range of participants. However, most studies ($n = 7$) failed to provide sufficient detail about their recruitment strategies.^{85–89,94,99}

Data collection

Four studies were judged as rigorous in the collection of data.^{77,88,89,112} These studies described how they approached data collection, with aims including putting participants 'at ease', providing them with opportunities to ask questions⁷⁷ or employing youth as interviewers.⁸⁸ In some cases, studies attempted to increase the range of data collected by using more than one method of collection, for example, by conducting in-depth interviews as well as drawing on field notes. The remaining six studies did not report these methods clearly.^{84–87,94,99}

Data analysis

Half the studies ($n = 5$) were judged as making attempts to increase analytical rigour.^{77,86,88,89,112} The remaining studies failed to provide a description of the analytical approach undertaken.^{84,85,87,94,99} Detailed descriptions of methods of analysis were generally missing. Most studies were 'thin' in their interpretive scope. Rather than developing clear, second-order interpretations, studies often provided descriptive accounts of how contextual and participant characteristics could function as barriers to or facilitators of implementation or receipt of PYD programmes.

Grounded in/supported by data

In many cases, we found a lack of first-order constructs in the form of verbatim quotes, relying mostly on authors' observations and descriptions to support accounts of participants' experiences, views or beliefs. Seven studies^{77,84,86,88,89,99,112} were judged as being grounded and supported by the data. However, the three remaining studies did not include enough data in the form of participant quotes or author description to show how the authors arrived at their findings.^{85,87,94}

Usefulness of qualitative studies

Breadth and depth

The majority of studies were judged as providing both good breadth and depth in their findings.^{77,84,86,88,89,99,112} The remaining three studies were limited in the reporting of their findings, both overall and in relation to answering the RQ.^{85,87,94} During the coding process it also became apparent that, although there were some data on contextual factors, these were often quite 'thin'. There were even fewer data on how participant characteristics affected PYD implementation and receipt.

Perspectives

The last criterion assessed the extent to which young people's own perspectives had been considered in the process evaluation. Half the studies^{77,86-88,112} were judged as giving appropriate weight to young people's experiences of engaging in programmes alongside other informants (e.g. programme providers). Strikingly, for studies of an intervention that aims to place young people's empowerment at the heart of their programme model, very few studies engaged with young people as co-researchers. Reflective accounts of the power dynamics that might emerge during the research process were also scarce.

Overall weight of evidence

The final phase of quality assessment was for two reviewers to individually assign two types of 'weight of evidence' to studies (Table 3). First, a weight (high, medium or low) was assigned according to the reliability of the findings (i.e. the extent to which the methods used were rigorous and minimised bias in the findings). A second weight (high, medium, low) was assigned according to the usefulness of the findings in terms of how well the contexts and barriers to implementation were described. Overall, three studies were judged to be of high reliability and usefulness.^{77,89,112} One study judged to have high reliability was considered to be of medium usefulness⁸⁸ and one study was judged as being of medium reliability and usefulness.⁸⁶ Two studies were judged to be of low reliability but high usefulness,^{84,99} whereas the remaining three studies were judged to be of both low reliability and usefulness.^{85,87,94} Studies were not excluded from the synthesis based on judgements made about their low quality, but quality was used to determine the qualitative weight given to findings in our synthesis.

TABLE 3 Reliability and usefulness of findings

Study	Quality of evidence					
	Reliability of findings			Usefulness of findings		
	Low	Medium	High	Low	Medium	High
Armstrong and Armstrong ⁸⁴	✓					✓
Baker <i>et al.</i> ⁸⁵	✓			✓		
Berg <i>et al.</i> ⁸⁶		✓			✓	
Bloomberg <i>et al.</i> ⁸⁷	✓			✓		
Bulanda and McCrea ⁸⁸			✓		✓	
Cross <i>et al.</i> ⁸⁹			✓			✓
Lee <i>et al.</i> ⁹⁴	✓			✓		
Maxfield <i>et al.</i> ⁹⁹	✓					✓
Schwartz <i>et al.</i> ⁷⁷			✓			✓
Wiggins <i>et al.</i> ¹¹²			✓			✓

Thematic synthesis of process evaluations

A comparison of coding revealed no important differences, with reviewers developing very similar coding templates comprising similar themes and subthemes (see *Appendix 10*). The main differences related to slight variations in the length of findings extracted from the original paper.

Theme 1: collaboration with the community

Resonating with the suggestions from Lerner and Lerner¹²³ in our synthesis of theoretical literature, a major theme across a number of studies^{77,84,86,87,94,99,112} was the importance of collaborating with local communities in various ways to support implementation.

Cultural sensitivity, collaboration and integration with ethnic minority communities

The importance of cultural sensitivity and collaboration when implementing programmes in ethnic minority communities was a subtheme across three studies,^{84,87,94} all of which were judged to be of low reliability but varying degrees of usefulness. These reported that formal and informal community engagement was a key factor in ensuring that programmes were culturally sensitive, accessible and appealing to young people and their parents. This was particularly important when programmes were targeting, or situated within, marginalised ethnic minority populations. For example, Armstrong and Armstrong⁸⁴ (a study judged to be of high relevance but low reliability) reported from interviews with site coordinators delivering ASPs in a south-western US state that a programme's cultural relevance 'within the community was critical to successful implementation'. The authors acknowledged that being located within an 'ethnically diverse community' meant that it was 'important to have a strong cultural awareness'. They focused their efforts both on 'outreach projects with parents' and on schools in the local area, reporting on the importance of 'liaison . . . with a trusted member of the community who could communicate with the parents, often times in Spanish'.

The study by Lee *et al.*,⁹⁴ judged to be of both low reliability and relevance, supported this finding. In their evaluation of a youth development programme targeting the Aboriginal communities of the Northern Territory of Australia, the authors highlighted the importance of seeking and incorporating the views of the wider community as well as of young people or parents. This, it was suggested, could provide support through the generation of ideas and allay fears among minority ethnic groups that the programme was: 'a non-Indigenous solution so there is little ownership of it by Indigenous people'.⁹⁴ Over the course of the programme, consultation with the community was 'enhanced . . . with more programmes developed from community ideas and delivered with Indigenous people at the forefront'. This included delivering activities that were youth-oriented but culturally relevant, such as 'bush hunting excursions and using computers to record traditional music'. It was found that activities, such as watching organised sports events, could bring about community changes that led to reinstating 'understanding and respect between generations'. This study found that as the programme progressed, staff became more active in encouraging community members to get involved; this was made possible by staff working flexibly and 'engaging in regular formal meetings and informal discussion' with members of the community.

Lee *et al.*'s⁹⁴ study also highlighted the importance of increasing both the cultural relevance and participation of the local community by addressing potential language barriers through 'translating key proceedings' and communicating with indigenous members 'in their language'. The US study by Armstrong and Armstrong⁸⁴ also found instances in which young people were allowed to access and participate in the programme only because 'the parents were able to communicate with, and trusted the liaison' officer connected with the programme.

In some cases, programmes actively recruited community members as staff. For example, Lee *et al.*⁹⁴ found that employing a 'respected community member with a close collaborative working relationship with non-Indigenous staff' was perceived by many participants as 'pivotal to the initiative's success'.

Such actions could also be seen as providing the additional benefit of promoting local role models. For example, after identifying a 'lack of Chicano Latino adult role models' who could 'encourage, empower and develop leadership skills and qualities' of local Mexican American youth from the Minnesota area of the USA, programme providers in the low-quality study by Bloomberg *et al.*⁸⁷ trained local community members as 'facilitators' who could 'work closely with youth in the initial 2 day training', with the aim of establishing and maintaining a bond with them.

Challenges with community engagement and establishing trust

However, challenges in relying on volunteer community engagement and establishing trust of parents was a subtheme across two studies^{77,99} of differing quality. These were studies of interventions that did not specifically target diverse ethnic populations but that attempted to involve parents and local members of the community. Difficulties were reported in establishing good working relationships. For example, as reported by a study of high reliability and usefulness by Schwartz *et al.*,⁷⁷ ensuring the successful implementation of an intervention component implemented with the support of a volunteer mentor was challenging when they were not always reliable in maintaining contact, leaving participants feeling 'disappointed'. One young person from this study talked about her mentor, stating that she hoped that 'she would be there more than she was. . .and, and she wasn't'. She believed that the mentor's lack of presence was not only because 'she was always workin', but also because her mother and her mentor had worked together', and 'whenever my mom had quit that job or whatever, I guess she just went her separate way, and didn't come around any more'.⁷⁷ In some cases, programmes attempted to turn to their 'back-up' mentors, but these were still a 'small pool' of volunteers who were often outside the mentees' immediate community. There was also a reluctance to draw on them too heavily, as they could not necessarily deliver a programme with good fidelity in the absence of sufficient training (see Bloom *et al.*,¹⁰⁶ a study linked to that of Schwartz *et al.*⁷⁷).

Building trusting and openly communicative relationships with parents could also be challenging. Regarding the QOP, an initiative consisting of intensive case management, mentoring and positive youth activities, the evaluation by Maxfield *et al.*⁹⁹ judged as providing highly useful findings but having low methodological reliability, found that trust and open communication were seen as important means of maintaining contact with young people and encouraging uptake of intervention activities. The case managers in this programme reported that parents who appeared 'anxious to limit the exposure of family problems', who seemed to experience case managers as 'intrusive', or who may have 'felt threatened' by the mentoring relationships that case managers established with their children were subsequently the 'most difficult to reach' compared with parents who actively supported case managers to 'locate' young people and 'reinforced the value of attending program activities'.⁹⁹

Collaborating with and utilising local community resources

Another subtheme concerned with collaborations with others in the local community was the importance of collaboration with other community agencies to enable programme implementation, again echoing the suggestions from Lerner and Lerner¹²³ cited in our synthesis of theoretical literature. This was apparent in three studies^{84,86,99} of variable reliability and usefulness. In Armstrong and Armstrong's⁸⁴ evaluation of ASPs in a south-western US state, it was found to be 'extremely important for the site to utilise community resources from a programmatic standpoint'. These 'additional resources', made possible through links with the community, enabled providers to expand the range of activities offered, a critical element of PYD. For example, local libraries proved to be an 'unplanned benefit' that could help deliver a reading programme. Programme providers cited being able to host 'occasional large-scale events' by 'taking advantage' of a nearby Boys and Girls Club. Local funding bodies were considered another important local resource to support positive youth activities. This was the case in the study of medium reliability and relevance by Berg *et al.*,⁸⁶ in which the programme received a grant that 'enabled the visual research group [of young participants] to receive training in photography and show their work at a photography exhibit'.

This subtheme was also apparent in the study of the QOP by Maxfield *et al.*,⁹⁹ who reported providers maximising the range of opportunities provided in the community by 'forming partnerships with agencies that specialised [in a range of life skills training topics] such as substance abuse prevention, conflict resolution training, date rape, and sexual abuse'. Like Armstrong and Armstrong,⁸⁴ they also collaborated with the local Boys and Girls Club, which enabled them to offer young people workshops on a range of topics, 'such as anger management, self-esteem, planned parenthood, family planning and other life skills'. The importance of being able to make use of other local services to maximise breadth of opportunities was regarded as particularly important when there were gaps in programme providers' expertise, such as when drawing on 'student volunteers from the local university'⁹⁹ to offer tutoring to support sites at which case managers felt that they lacked the skills to provide such services.

Collaboration with schools

A final subtheme regarding community collaboration was that collaboration with schools was critical to implementation but could be time-consuming and challenging. Three studies,^{84,99,112} two of which were based in the USA and one of which was undertaken in the UK, examined the importance of liaising with schools to support the successful implementation of programmes. All three studies were judged to be of high usefulness but were variable in terms of methodological reliability. Site co-ordinators in Armstrong and Armstrong's⁸⁴ study (judged to be of low reliability) of ASPs in south-western USA indicated that communicating with other community stakeholders to support the development of youth 'such as schools, had an important impact on program implementation'. This was particularly relevant in their study because they had a number of ASPs located off schools' sites. One way of dealing with barriers to communication arising from this was to designate a school liaison, who could work across programme sites, but who was an employee of a single school. The schools then also acted as a channel to disseminate information about programme events to young people and their families in order to reach a wider audience and increase programme reach.

A study by Wiggins *et al.*¹¹² of after-school youth development targeting at-risk young people across England and judged to be of high reliability and usefulness also found that 'working with schools was crucial' for recruiting young people to programmes. The authors noted that liaising with staff 'with appropriate authority within schools was critical' and that negotiating 'access and referral routes' was time-consuming. However, in a context of providers aiming to meet challenging recruitment targets, some sites reformatted their programme so that young people attended as an alternative rather than a supplement to their normal schooling, a major distortion of the intended intervention model. In an evaluation of the QOP, Maxfield *et al.*⁹⁹ also report that collaborations with schools could raise challenges. Although providers encouraged those 'most in need of academic remediation' to take up professional tutoring, these young people had to travel across the city to a different school because this is where tutoring was provided. Case managers transported young people to sessions to ensure their attendance but soon found that this 'proved too burdensome' for them and that 'youth ceased to attend tutoring sessions with any regularity'.

Theme 2: young people's relationship with programme providers and peers

The second major theme that was apparent across a number of studies^{88,89,99} was the importance of young people's relationships with programme providers and peers in maximising the acceptability and potential impact of interventions.

Calm and authoritative programme providers

Within this theme, one key subtheme was the importance of programme providers attending to young people in a calm and nurturing yet authoritative way, including in response to any challenging behaviour exhibited by participants. Three studies of varying reliability and usefulness described provider attitudes and responses to young people in this context. The 'Stand Up Help Out' programme, which aimed to develop youths' capacity for constructive relationships with adults and peers, was evaluated by Bulanda and McCrea⁸⁸ and judged to be highly reliable and of medium usefulness. This reported that successful implementation was associated with staff signalling their continued commitment to providing

'unconditional positive regard' when faced with challenging behaviour from young participants. It was reported that this response style was acceptable to the young people, who did not feel that they were treated 'negatively'. One participant commented:

That's what I like about [the instructor]. Cuz, even when he don't get all the respect he should get out of the kids in the program, he still be [himself]. You know calm and collective [sic].⁸⁸

This was in contrast to the All Stars prevention curriculum, an enhanced ASP offering a range of social and educational activities evaluated by Cross *et al.*⁸⁹ and judged to be a highly reliable and relevant study. This reported that staff struggled to respond to young people's challenging behaviour. Drawing on data from different programme sites, the authors described sites as 'rife with behaviour problems', whereby young people received 'very little redirection from staff members' when they 'acted out', leading the authors to be concerned about participants' physical safety. The authors also noted that when staff from this site did interact with young people, the disciplinarian approach taken 'appeared capricious and confusing to youth'.⁸⁹ For example, when providers applied the use of 'quiet time', they did so to *all* young people rather than merely to those engaged in misconduct. In another site, the same evaluators found staff to be 'irritated and apathetic', appearing to engage more with each other than with young people and failing to address young people's challenging behaviour.

The evaluation of the QOP by Maxfield *et al.*⁹⁹ found evidence supporting the need for case managers to engage with young people as individuals rather than as part of a group. Maxfield *et al.*⁹⁹ found that 'the most successful mentors used a balance between nurturing and discipline' when interacting with young people. In contrast, if staff took what was described as a 'parental or authoritarian' approach, young people often responded with resistance; if they adopted a merely 'friendly approach', case managers could be taken 'advantage of' or simply 'dismissed' by some young people. Case managers in this study were also said to identify a point at which young people needed less nurturing 'in order to encourage them to become more self-reliant'.

Positive peer relations

A further subtheme was the importance of positive and supportive peer relations, underpinned by staff and by programme structure. Three studies, of differing reliability and usefulness, examined the contextual factors that appeared to contribute to young people being potentially able to overcome difficulties or tensions between them. For example, a high-quality study by Bulanda and McCrea⁸⁸ described a US ASP providing young people with an apprenticeship in social work. They reported that social differences such as membership of different 'street alliances', which could be a cause of conflict outside the programme, did not necessarily prevent mutual collaboration and support within the programme, as long as participants were able to 'prioritize their connectedness over the potential discord created by differences'. The authors suggested that this connectedness was achieved by youth being able to 'recognize relationship problems and focus on relationship strengths'. However, reporting on another US-based ASP providing access to social and educational opportunities, Cross *et al.*⁸⁹ argued that tensions among participants or between participants and staff could be overcome only in sites that were well managed. A lack of organisation and high turnover of staff at one site within their study was a key factor in young people not seeming 'to enjoy each other's company', whereas the positive outcomes observed in another site may have been attributable to 'the friendships among students, which were in part facilitated by stable site management'.

A third study, also judged to be of high reliability and usefulness and conducted in the USA, reported that potential discord among young people owing to gang affiliation was not considered by providers to be a reason for excluding young people from participating. Instead, Bloom *et al.*¹⁰⁶ described how the NGYCP separated participants who belonged to different gangs into different 'squads' and removed or obscured gang symbols such as tattoos that could function as 'physical reminders of past affiliations'. However, problems associated with gang membership were not always easily overcome, with one site in Mississippi reporting higher incidences of 'fights and intimidation' and 'gambling and extortion' and noting that there were 'higher levels of attrition among this subset of the population'. However, staff members did report

that external problems were less likely to intrude during the residential phase of this programme, where they 'have them 24/7' and can instil values that young people can then take 'home with them'.¹⁰⁶

Theme 3: staff retention

Staff retention was a key theme evident across three studies^{84,89,99} of differing reliability and usefulness. These studies reported on the importance of staff continuity to ensure that programmes were implemented fully and appropriately.

Staffing continuity essential to successful implementation

Within this theme, a key subtheme was the importance of staffing continuity to intervention delivery. Site co-ordinators in the study of ASPs in south-western USA by Armstrong and Armstrong,⁸⁴ which was judged to be of low reliability, felt that effective implementation and sustainability of ASPs relied on ensuring that staff turnover was kept to a minimum. This was a challenge for some programmes. For example, Cross *et al.*⁸⁹ report on a high-quality evaluation of a social and educational activity-based after-school initiative, finding that 'six of the original fourteen staff members quit or were fired before the end of the year'. Similarly, Lee *et al.*⁹⁴ reported that turnover of staff 'impacted significantly on program continuity and workloads'. Maxfield *et al.*,⁹⁹ evaluating a community-based case-management and mentoring programme in the USA, noted that they were 'fortunate [that] turnover [in certain sites] was relatively low'. However, two instances in which staff turnover was 'most damaging' and led to a failure in sustaining mentoring relationships was 'when a position remained vacant', resulting in participants not having a 'primary mentor for as long as two or three months', or when participants had 'more than two case managers' and rarely developed 'strong relationships with their third case manager'.

Difficulty offering full-time posts in the field of youth work

It is clear across a number of studies that the limited number of hours involved in delivering youth programmes could sometimes mean that it was not possible to offer full-time positions, thereby increasing the difficulty in securing and retaining qualified staff. To overcome this challenge, Armstrong and Armstrong⁸⁴ reported how one programme aimed to recruit ASP staff who were not looking for full-time work, such as college students interested in gaining experience of youth work or other professionals who were interested in working with young people, in addition to their primary employment. Difficulties with retaining trained employees could also mean that replacement staff were not well trained. Cross *et al.*⁸⁹ report that youth workers who had been retained since programme initiation 'received more than 40 h[ours] of training on average', compared with 'less than 6 h[ours]' for those who had replaced them, and that sites with high employee turnover were less likely to have staff who were highly trained. They found that in some cases, high staff turnover meant that it was not possible to employ staff who had participated in the original 3-day training that was offered before the start of the programme and that in some cases it was 'necessary for untrained replacement staff to deliver the All Stars' programme (see Gottfredson *et al.*,⁹¹ a study linked to Cross *et al.*⁸⁹).

Creative attempts to compensate for lack of trained staff included drawing on existing skills that happened to be held by staff members and incorporating these opportunistically into programme activities. For example, Armstrong and Armstrong⁸⁴ observed that at one site an employee 'with extensive orienteering skills' was encouraged 'to organize camping trips and day hikes for youth' and that at another site, a staff member 'who enjoyed jazz dancing started a dance program'.

However, two studies^{77,99} reported that it was difficult to overcome limitations in skills owing to a lack of training and that this could mean that programmes might not be able to provide the range of activities normally expected of a PYD programme. For example, Bloom *et al.*¹⁰⁶ (a report linked to the study by Schwartz *et al.*⁷⁷) found that provision of 'a one-on-one paired tutoring model to be provided an hour and a half per week' was impossible to implement because of a lack of tutor capacity and had to be 'abandoned midway through the year'. The authors felt that despite providing an alternative academic activity, the lack of 1 : 1 tutoring may have 'contributed to withdrawal of youths whose parents viewed tutoring as the main draw of the program'.¹⁰⁶ In addition, Maxfield *et al.*⁹⁹ reported that programmes

found it difficult to secure staff with expertise across the range of PYD domains. For example, when programmes were expanded to include an educational component, their delivery could be a challenge for staff who had been 'hired on the basis they could be case managers not tutors or teachers'. Such staff lacked the skills needed to provide educational activities, and 'required extensive training and technical assistance'. This proved difficult for some sites, with providers reporting that when case managers were acting as tutors, 'more expert tutoring would have greatly improved services'. Other sites that did not provide 'extensive in-service training to improve case managers' tutoring skills' relied on volunteer tutors instead. This was also associated with retention problems because these volunteers tended to work for the programme only for 'one or two semesters', thereby causing difficulties with 'forging long-lasting relationships' with staff on this basis.

Theme 4: youth-led empowerment

Our final theme, drawn from five studies,^{77,85,86,89,99} concerns the importance of, and potential contradictions and challenges inherent in, ensuring that young people are empowered to make decisions about their engagement in programme activities.

Young people determining their own engagement in activities

One subtheme within this relates simply to the extent to which young people were empowered to determine in which PYD activities to participate. This was described in three studies of variable reliability and usefulness. A study by Berg *et al.*,⁸⁶ judged to be of medium reliability and usefulness, described 'youth empowerment' as a key component in the YARP programme and suggested that staff needed to be trained in 'facilitation techniques' to halt the tendency for staff to determine how community engagement projects are undertaken. Young people's decision-making processes were considered more important than their final choice of activity in the study by Baker *et al.*⁸⁵ of the South Baltimore Youth Centre. This programme, which aimed to prevent violence, delinquency and substance abuse, attempted to provide access to a range of educational and social activities. The evaluation, judged to be of low reliability and relevance, reported that when activities were 'imposed [in a] top down [manner, they] failed and were abandoned'. Providers instead aimed to give young people authority so that all activities provided by the centre were 'planned and carried out by the youth' themselves. Schwartz *et al.*'s⁷⁷ study of youth-initiated mentoring, which was judged to be highly reliable and useful, found that when young people were able to choose their mentors without input from family members or programme staff, there was a greater likelihood that the mentoring relationship would 'endure'.

Limitations to choice provided

However, two studies, judged to be highly reliable, reported that young people in some particular interventions had very limited empowerment to shape and determine their involvement in positive youth activities. For example, empowerment in the 'All Stars' curriculum, evaluated by Cross *et al.*,⁸⁹ was highly restricted. In this study, young people could choose activities but were restricted to a list predetermined by the site director and programme assistant at the start of each day. They also found that 'despite intentions to provide a wide variety of activities to youth, it is clear that only a few activities were offered consistently during leisure activity time' (see Gottfredson *et al.*,⁹¹ a study linked to Cross *et al.*⁸⁹). Empowerment was also restricted in the programme evaluated by Schwartz *et al.*⁷⁷

Tensions arising from choice

Another subtheme in terms of youth empowerment was the tensions that could arise when empowering young people to choose which activities in which to engage. Four studies, judged to be of high relevance with variable reliability, provided data on young people's choice of activities, showing that some programme components were often rejected by young people on the basis that they were unappealing. Sometimes these were activities with an educational component. These included 'computer-assisted instruction' and 'community service', which were not received with 'enthusiasm', as well as standardised tests, which participants showed 'little patience for', particularly those young people 'who had difficulty reading'.⁹⁹ This was also the case for 'computerised job training' which was 'ignored'⁸⁵ and academic assistance.⁸⁹ Wiggins *et al.*¹¹² argue that taking an academic 'school like approach' could alienate young

people, particularly those whose lives are 'chaotic and hard' and suggest that young people need to be able to get involved in activities at a level that is 'most appropriate for them at any given time'. This might suggest the importance of a diversity of provision, not only to enable choice but also to take account of the fact that different young people will have different preferences and developmental needs, a point resonating with suggestions from Busseri *et al.*¹¹⁷ cited in our synthesis of theories of change.

However, some process evaluations, as well as much of the theoretical literature, suggested that young people's empowerment to choose activities was central to PYD provision. The study of the QOP by Maxfield *et al.*⁹⁹ suggested that facilitating choice may in some cases deter engagement in the broad range of activities which is commonly regarded as a central feature of PYD. This study, judged to be of low quality but high relevance, reported that some sites offered more recreational activities (e.g. outings to the cinema, ice-skating, swimming, etc.) because they attracted 'more enrolees than did other activities'. However, as young people got older, they resisted staff's promotion of 'activities with learning content' and continued to favour the more recreational activities that providers had originally used 'to attract youth to the program'. Participants reported that they missed doing 'fun things' and that 'museums and other cultural activities were boring'. Similarly, when there was a scheduling clash between attending summer school and taking up summer employment, case managers were more likely to recommend summer school but they were 'not able to prevent an enrolee from choosing [paid work]'. This contrasted with programme sites that provided a balanced combination but offered participants little choice, which appeared to have 'less difficulty in maintaining interest' among young people. In the highly reliable and useful US-based study by Schwartz *et al.*⁷⁷ and the linked study by Bloom *et al.*,¹⁰⁶ evaluators found that a compulsory academic component was viewed much more positively. They report that participants of the National Youth Guard mentoring programme 'welcomed the small class size, tailored instruction, and self-paced approach'¹⁰⁶ of the graduate education programme. One young person was reported as stating that they liked the General Educational Development (GED) programme because:

*[t]hey do it, one course at a time. That's just awesome. And also the way they treat you is like people. I thought it was going to be like Juvie, you know, where they treat you like crap all the time. It's not. They treat you like people. That's what I like.*¹⁰⁶

The authors noted that a key element of the success of their educational component was that it was noticeably different from what young people were used to experiencing in school, because it combined both structure and individual support. One participant felt that he was 'less distracted by girls' and could focus on his goal 'to get the work done'.

The study by Maxfield *et al.*⁹⁹ reported that some sites provided financial incentives to increase engagement in specific educational activities, such as computer-assisted instruction and assessment tests. However, the two sites that took these approaches found that it was 'effective for only short periods of time and only for students already inclined to spend time on the computer' and did not prove effective in motivating already resistant young people. Later on, stipends came to be used only for time 'spent engaged in education, developmental, and community service activities'. The use of incentives was also reportedly problematic in the US-based study by Cross *et al.*⁸⁹ (cited in the linked study by Gottfredson *et al.*⁹¹). To increase engagement in programme activities, young people were randomly assigned to groups which would accrue points for attendance. However, programme staff thought the system unfair and decided to place high-attending youth together 'to ensure the attending students would receive the maximum point', thus undermining the intended system. This 'probably did not encourage attendance among the lower attending youth because they were placed in groups with very low probabilities for receiving points'. In both of these programmes that used incentives, there was a tension in providers' attitudes to empowerment. Although programme providers appeared to want to enable choice, they also sometimes wanted to constrain and incentivise choices to ensure that young people engaged in the programme overall or in specific activities that staff regarded as important.

Summary of synthesis of process evaluations

A number of themes emerged from our synthesis. Formal and informal community engagement was a key factor in ensuring that programmes were culturally sensitive, accessible and appealing to young people and their parents, as well as to the wider community. Employing community members could be pivotal to successful implementation and providing role models. However, volunteers could also be unreliable, for example in acting as mentors. Collaboration with other community agencies could also be important, particularly in expanding the range of activities offered. Another theme was the importance of young people's relationships with providers and peers. Providers should ideally relate to young people in a calm, nurturing yet authoritative way. Peer support was also important, sometimes in the face of challenges with social differences among peers such as in membership of different gangs. Skilled providers could achieve this by facilitating participants to recognise common concerns, but this was difficult when staff were poorly trained or retained.

More generally, staff continuity was reported to be crucial for PYD, because the approach requires staff with a diversity of skills and experiences who can offer participants a range of activities as well as durable relationships. Retention was challenging where programmes, particularly those operating after school or at weekends, could not offer full-time positions. A final theme concerns the importance of, and challenges to, ensuring that young people are empowered to make decisions about programme activities. This required that staff were trained in facilitation rather than merely being directive. Tensions could arise between PYD's aims of empowering young people to choose and requiring young people to engage in different activities to develop specific assets, such as vocational or academic skills.

Hypotheses arising from the synthesis of process evaluations

The above discussion suggests the following hypotheses:

- Interventions that have specific methods to engage communities will be more effective.
- Projects that engage with schools will achieve better recruitment.
- Interventions that are delivered by well-trained staff will be more effective.
- Interventions that have better staff retention will be more effective.
- Interventions that offer some choices but require some engagement with educational components will be more effective.

Feedback from consultation on the synthesis of process evaluations

The policy and practitioner advisory group and young people provided specific feedback on the synthesis of process evaluations and the hypothesis generated from this synthesis (see *Appendix 12*).

Feedback from policy and practitioner stakeholders

The themes from the process synthesis made sense to policy and practitioner stakeholders and resonated with their personal experiences, particularly the relevance of working with the community and the importance of staff retention in building relationships with young people as key factors in promoting programme success.

Many of the stakeholders raised the importance of drawing out the extent to which the findings can be transferable to the UK, considering that the majority of studies were conducted in the USA. Stakeholders appreciated that this is a problem for many evidence syntheses and that we might not be able to address this issue fully. They suggested that, as we move forward in finalising the draft report, we consider the following participant characteristics and contextual issues: the extent to which programmes are accessible

and acceptable to 'hard to reach' young people in the UK who might not necessarily participate in universal non-targeted PYD programmes; the possible significance of the need to provide different activity programme components depending on age group and the implications that this has for programme service delivery; the demands placed on young people's 'spare time' in the UK, particularly now that 16- to 18-year-olds are expected to be in some form of education or training; the different roles that local volunteers might have in delivering PYD programmes in the UK compared with the USA; and the importance of sex differences and the rates at which girls develop compared with boys (highlighting if there is a lack of process and outcome data on these differences).

Feedback from consultation with young people

The ALPHA group agreed that it was important to have one worker who young people could get to know and become comfortable with. Young people suggested that staff should include those with a mix of qualifications, ideally with good skills to work with young people. The group argued that there should ideally be no engagement with schools except as a site for recruitment, but that it was important and acceptable to engage with other services, such as health services and the police, so that young people might develop a better understanding of these agencies. Participants felt that PYD agencies need to find the right balance between offering young people choices and providing educational elements, with the risk that agencies that felt too much like school would be unappealing.

The above comments from policy/practitioner stakeholders and young people are thus broadly in line with the assertions made in the theoretical literature and do not suggest any amendments to the hypotheses raised above.

Chapter 6 Synthesis of outcome evaluations

About this chapter

This chapter describes and reports on the quality of outcome evaluations of PYD interventions that met our inclusion criteria. It also reports our narrative and meta-analytic synthesis of these studies.

Included studies

Studies and their quality are tabulated in *Appendix 16*. We included 13 study reports^{79,86,92,93,95,98,103–105,107,108,111,112} of 10 distinct studies in our narrative synthesis of intervention evaluations and 12 study reports^{79,86,92,93,95,98,103–105,107,108,112} of nine distinct studies in our quantitative synthesis. All but one study (from the UK) was conducted in the USA. A total of three studies involved those aged 11–14 years; two involved 14- to 16-year-olds; one study comprised 16- to 18-year-olds; and four studies involved participants across these categories. Five studies assessed targeted interventions and five examined universal interventions. Five studies assessed interventions lasting 1 year or less and five examined interventions of more than 1 year's duration. We aimed to develop a taxonomy for the interventions subject to outcome evaluation using categories which emerged from our synthesis of theory and process evaluations. Although we produced this (see *Appendix 14*), it did not prove instructive because the outcome evaluations did not describe interventions in the detail required for our taxonomy. No economic evaluations were found, so we were unable to synthesise evidence on the cost-effectiveness of PYD interventions.

The YARP intervention reported by Berg *et al.*⁸⁶ was developed to involve young people in understanding the challenges facing their communities and in engaging in action research to address these problems. The programme began in the summer and continued throughout the school year in after-school meetings, although the programme was not tied to a specific school. Staff were educators trained in the specific programme model, including the use of action research methods and group facilitation. Participants were trained in action research methods over the summer and then, over the school year, researched a variety of challenges, including sexual risk behaviour in adolescence, disengagement from school and illegal trading (called 'teen hustling' in the evaluation). Research culminated in a variety of presentations, exhibitions and dissemination strategies. Rather than present didactic curricula about substance use and violence, the programme largely focused on the promotion of positive assets by encouraging young people to become community agents of change through developing their own self-efficacy. The intervention's theory of change was built on a variety of theories, including ecological systems theory, theories of identity and learning, and critical sociological approaches to social change. Although methods of recruitment were unclear, the programme appeared to be targeted to low-income minority ethnic young people living in urban areas.

Another study examined an intervention providing a combination of academic support, leisure activities and 'All Stars', a branded prevention curriculum.^{90–92} This was delivered over the course of one academic year to students in Baltimore public schools using staff from a government agency charged with management and delivery of local parks and recreation programmes. Staff were trained in the delivery of general ASP activities as well as in the All Stars⁹⁰ curriculum itself. Although the All Stars⁹⁰ curriculum

appeared to be primarily oriented towards teaching participants skills and normative beliefs for the avoidance of substance use, some components promoted positive assets by encouraging participants to develop links with community groups, which was consistent with the programme's goals and aimed to improve students' communication with their families. The core of the theory of change of the prevention curriculum was that change in beliefs and commitment to abstain from substance use, as well as school bonding, improved family communication and improved decision-making and goal-setting skills, would reduce the risk of substance-use and conduct problems. In the context of this trial, the ASP would additionally work to promote school bonding, reduce unsupervised socialising and create positive peer influences, all of which would further decrease the risk of conduct problems. Although the intervention was not targeted per se, low-performing schools without an ASP were chosen for participation, based on expressed need.

Gottfredson *et al.*⁹³ report on a US state-wide demonstration project of Maryland ASP (MAPs). Sites differed somewhat in terms of what was implemented but all offered academic support and social skills training, as well as athletic and arts activities, over the course of a school year. Programmes were located both in schools and in community centres and were offered by a diversity of groups, including schools and local authorities. Because interventions were diverse, no one theory of change unified all of them, although the tendering brief specified a focus on the above activities and stated implementation standards to match. A general principle underlying the evaluation was that reductions in unsupervised socialising could lead to decreases in problem behaviours, including substance use. The intervention was targeted not at an individual level but at areas experiencing high crime rates.

Cool Girls, Inc.⁹⁵ was delivered mainly as an ASP with weekly meetings, although it included additional weekend activities. Cool Girls, Inc. consisted of a life skills curriculum with academic support and specific programme-relevant academic and social activities, and was delivered solely to girls at local grammar schools. Programme participants were further eligible for mentoring after 1 year in the programme. Parental involvement was encouraged through specific activities. The programme was run by a community-based organisation and balanced prevention through didactic curricula with the promotion of positive assets by encouraging engagement with 'supportive social networks', cultural programming and the development of self-esteem. The intervention's theory of change was focused on developing individual skills and access to 'external resources' to promote healthy lifestyles. Although recruitment did not appear to be selective, the programme was targeted at girls living in economically disadvantaged communities.

The QOP reported by Rodriguez-Planas,⁷⁸ which was delivered by staff from community-based organisations in after-school settings, included a substantial case-management component tied in with mentoring provided by programme staff. It also included academic assessment, planning and tutoring, community service and leisure activities and, when necessary, support over the summer vacations. Staff members were youth workers who assumed 'round-the-clock' on-call responsibilities for participants assigned to them as part of the case-management model. The intervention's theory of change was not explicit but appeared to be premised on completion of secondary school education as a way to prevent antisocial behaviours and to attain employment. There appeared to be little specific prevention education. Promotion of positive assets was achieved both by mentoring and by life skills training, cultural awareness programming and community service. Students were enrolled upon entry to the first year of secondary school through to graduation and were eligible for inclusion in the programme based on being in the bottom two-thirds of the grade distribution in the entering class of their secondary school and not having special educational or disability needs that would prevent participation. Schools with dropout rates of 40% or more were targeted in this intervention. Implementation differed across the seven programme sites, with some sites funded by the Ford Foundation and one site (Washington, DC) implementing the programme 1 year later than the others.

The NGYCP reported by Schwartz *et al.*⁷⁷ was delivered to adolescents between 16 and 18 years of age who had either left school or been excluded, who were unemployed and who were not involved in the correctional system. It was run as a 5-month military-style 'boot camp' including a 'pre-ChalleNGe' and a residential component that included life skills education, work preparation and completion of the secondary school diploma. After the military boot camp, participants completed job placements and structured mentoring. Mentoring was provided both by programme staff and by mentors from the community nominated by participants. In an unusual feature, the intervention was primarily delivered by the National Guard, a branch of the US military run at the state level. Although the programme did not appear to include a large amount of prevention education, promotion of positive assets was achieved through community service, job skills and life skills training and development of leadership and citizenship skills, as well as subsequent job placement. The military boot camp was designed to incorporate what evaluators called 'positive youth development' activities to increase self-efficacy and self-esteem. Although the intervention did not set out an explicit theory of change, the key principle was that a 'wraparound' approach that addressed underlying issues in youth achievement and exposed them to the structure of a military context would be more effective than other less intensive approaches.

Stay SMART (St Pierre and Kaltreider¹¹¹) was delivered through selected Boys and Girls Clubs of America as an expanded substance-use and sexual risk prevention curriculum that also aimed to promote the development of a number of positive social skills. Programme group leaders were staff of the Boys and Girls Clubs who received training in programme delivery. This study only narrowly met inclusion criteria in that its primary focus was traditional prevention education with only a marginal focus on the promotion of positive assets. The theory of change underlying this programme was based on the development of skills and knowledge that would help in avoiding deviant peer influences. The 'basic' programme was delivered over 3 months, with booster sessions over the subsequent 2 years.

The Positive Youth Development Collaborative (PYDC) reported by Tebes *et al.*⁷⁹ was an after-school substance-use prevention programme that aimed to promote decision-making and learning about cultural heritage in minority ethnic adolescents. Community leaders, who received bi-weekly supervision from the research group, delivered the intervention over the course of a school year. Unlike other included evaluations, the comparator in this evaluation was an ASP without the substance-use prevention content, which appeared to constitute 'usual treatment'. The intervention included a substantially prevention-focused component that was delivered as a curriculum to teach methods for coping with stress and for improved decision-making about substance use. However, the promotion of positive assets was present in the curriculum through sessions on identifying positive personal goals and characteristics, and the intervention was described as being 'strengths-based'. This focus was also apparent in the additional activities offered, which included cultural heritage programming, community engagement and participant-organised activities, as well as mentoring from community groups. The intervention was targeted to ethnic minority adolescents living in urban areas, but other young people were not excluded from recruitment.

As described in the evaluation, Big Brothers Big Sisters of America (BBBS), reported by Tierney⁹⁸ and linked to the study by Rhodes *et al.*,⁹⁶ was a mentoring programme in which potential mentors were evaluated by programme staff and then matched with a young person for regular (generally several times a month) meetings. Although processes varied between sites, eligible youth generally lived in single-parent households and, along with their parents, agreed to the match. BBBS⁹⁸ was evaluated as a specific model of mentoring that focused on lay volunteers, long-term relationships and frequent meetings. Programme staff were often professional social workers, although the mentors actually matched with children (who thus provided the intervention) were adults drawn from the community. These adults were trained in

recognising and reporting abuse and, although this was not required, also often received training in communicating with youth. Volunteers receive monthly supervision for the first year of the match and quarterly supervision thereafter. The intervention included no formal education but rather the ongoing relationship with a trusted adult was intended to develop assets such as academic performance and social relationships with family and friends.

The YPDP, reported by Wiggins *et al.*^{46,112} and loosely informed by the Children's Aid Society–Carrera programme developed in the USA, aimed to include tutoring, sports activities, referral to health services and work preparation delivered in ASPs by staff members who provided support to youth at risk of a range of behaviours. The intervention was delivered over the course of a school year by youth service organisations in the voluntary or statutory sectors. Youth workers providing the programme received training in the programme model, which was focused on the promotion of positive assets in youth. This occurred through a wide range of activities, including mentoring, volunteering and opportunities for training and employment, although 'standard' prevention education was also delivered through health education. It should be noted that each site was allowed to set the specific mix of services delivered to participants across eight pre-specified components, described in the original programme briefing as including: education, training and employment activities; life skills development; health activities and education; mentoring by programme staff; sports and athletic activities; creative arts; volunteering; and access and referral to health and social services. The intervention was targeted to youth who were between the ages of 13 and 15 years and who lived in areas of high deprivation with high rates of teenage pregnancy. Teachers and social workers nominated participants based on perceived risk for substance use, teenage pregnancy or exclusion from school.

Evaluation design

Of the 10 included studies, four^{77,78,90,98} were RCTs. Five studies^{79,86,95,111,112} were prospective studies with non-random matched control groups. The evaluation of MAP⁹³ included both a randomised trial component and a non-randomised matched comparison component, with data from both being analysed together. All four randomised trials involved allocation by individual rather than site. Three non-randomised evaluations^{79,111,112} recruited comparison sites and then sought comparison group members within those sites. Two non-randomised trials^{86,95} recruited individuals, rather than sites, to the comparison group. The evaluation of MAP⁹³ used a combination of strategies.

All included evaluations collected substance-use outcomes, although only four study reports^{98,103,107,108} from three studies^{77,78,98} collected violence outcomes. Four evaluations reported outcome data near the end of a school year-based intervention schedule.^{86,90,93,95} BBBS⁹⁸ reported one post-test taken 18 months after baseline, at which point 77.6% of treatment group youths had been matched to a mentor. Stay SMART¹¹¹ captured intermediate and post-intervention outcomes during a 27-month intervention schedule. An additional three evaluations^{77,79,112} captured outcomes both post intervention and at one follow-up (4, 9 and 18 months, respectively). Finally, the evaluation of the QOP captured outcomes over several time points, from post intervention to as far as 6 years post intervention.

In all four randomised trials^{78,90,98,112} and three non-randomised trials,^{86,93,95} comparison groups did not receive alternative programming, although control group participants in All Stars⁹⁰ were offered monthly 'fun' activities with generally low uptake. Control groups for evaluations of Stay SMART,¹¹¹ YPDP¹¹² and the PYDC⁷⁹ were recruited from similar settings and programmes, which generally involved some baseline level of intervention, which did not offer the additional PYD component.

Risk of bias and quality of evidence

Table 4 and Figure 7 summarise risk of bias respectively for each study and overall.

Of the four randomised trials included here, only one⁹⁰ stated how the randomisation sequence was generated. This was done using a random number table. Each of the other three randomised trials^{77,78,96} included here did not present sufficient information on sequence generation to judge risk of bias.

Allocation concealment

Only one study (Tierney,⁹⁸ linked to the study by Rhodes *et al.*⁹⁶) noted how allocation was concealed from investigators. They did this by using an external survey contractor. Each of the other three randomised trials^{77,78,90} did not report sufficient information to judge risk of bias on allocation concealment.

Blinding

Although often difficult in trials of interventions, none of the included evaluations explicitly reported blinding with respect to intervention allocation. The six non-randomised evaluations^{79,86,93,95,111,112} included in this project all recruited comparison groups, making blinding difficult, if not impossible. Of the four randomised trials included, two randomised students within schools^{78,90} and two randomised at the point of sign-up for the intervention.^{77,96} Both strategies meant that participants were not blinded to intervention allocation

TABLE 4 Risk-of-bias summary review authors' judgements about each risk-of-bias item for each included study

Study	Sequence generation	Allocation concealment	Blinding	Incomplete outcome data	Selective outcome reporting	Key confounders	Clustering	Other source of bias
Berg <i>et al.</i> ⁸⁶	N/A	N/A	–	–	–	+	–	+
Cross <i>et al.</i> ⁹⁰	+	?	?	–	?	N/A	+	–
Gottfredson <i>et al.</i> ⁹³	N/A	N/A	–	–	?	+	–	+
Kuperminc <i>et al.</i> ⁹⁵	N/A	N/A	–	+	?	+	+	+
Rhodes <i>et al.</i> ⁹⁶	?	+	–	–	?	N/A	–	+
Rodriguez-Planas ⁷⁸	?	?	–	+	?	N/A	+	+
Schwartz <i>et al.</i> ⁷⁷	?	?	?	+	?	N/A	+	+
St Pierre and Kaltreider ¹¹¹	N/A	N/A	–	–	–	?	–	–
Tebes <i>et al.</i> ⁷⁹	N/A	N/A	–	–	–	+	–	+
Wiggins <i>et al.</i> ¹¹²	N/A	N/A	–	+	–	+	+	+

?, not reported; +, high risk of bias; –, low risk of bias; N/A, not applicable.

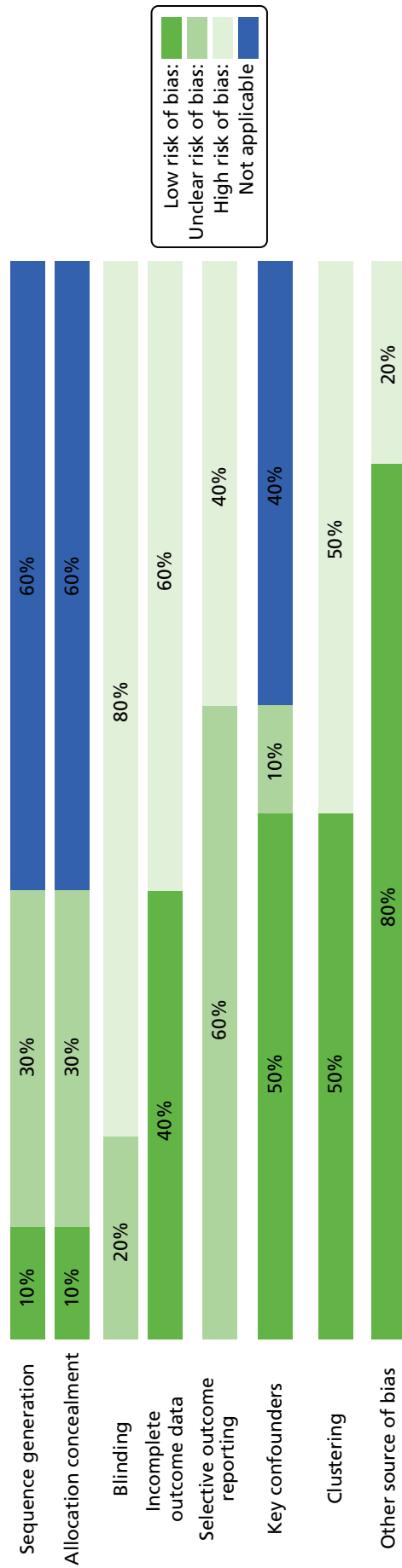


FIGURE 7 Review authors' judgements about each risk-of-bias item presented as percentages across all included studies. Sequence generation.

Incomplete outcome data

Incomplete outcome data presented a major problem across many evaluations. On the positive side, evaluations of the QOP,⁷⁸ the NGYCP⁷⁷ and the YDPD¹¹² used weighting methods to account for and examine incomplete outcome data. The evaluation of Cool Girls, Inc.⁹⁵ examined incompleteness and carried out multiple imputation. In less adequate approaches, several evaluations (BBBS,⁹⁸ All Stars⁹⁰ and Stay SMART¹¹¹) used data only from participants who completed all measurement occasions. Three evaluations^{79,86,93} experienced significant attrition but did not present any information about how incomplete outcome data were handled for included outcomes.

Selective reporting

We were unable to assess selective outcome reporting for 6 of the 10 included studies^{77,78,90,93,95,96} owing to inadequate information on pre-planned outcomes or the absence of a study protocol. The remaining four studies^{79,86,111,112} had a high risk of bias on selective reporting. Wiggins *et al.*¹¹² reported a subset of outcomes at second follow-up that were reported at first follow-up. Berg *et al.*⁸⁶ described collecting a variety of substance-use measures but reported sufficient information to calculate an effect size only for marijuana use. Tebes *et al.*⁷⁹ did not report outcomes for tobacco use. St Pierre and Kaltreider¹¹¹ reported tests for outcome differences inconsistently across time points and in ways that were difficult to interpret.

Type of comparator

As noted above (see *Evaluation design*), included evaluations generally compared a PYD intervention with minimal or no intervention. The one key exception is Tebes *et al.*,⁷⁹ in which the comparator was an 'active' usual care intervention that included a full after-school component. This may put the evaluation at a higher risk of bias than others, as the intervention provided to the specifically chosen comparator groups may have been harmful or otherwise biased the comparison against the PYD intervention offered to the treatment group. We ultimately decided to sensitivity analyse our findings by excluding this study, although this was based on the method of analysis rather than on type of comparator.

Accounting for clustering

Clustering was not appropriately addressed in many included evaluations. Evaluations of Stay SMART,¹¹¹ YARP,⁸⁶ BBBS,⁹⁸ PYDC⁷⁹ and MAP⁹³ either did not include clustering in analyses or did not specifically attend to how clustering was modelled in estimating intervention effects. Evaluations of All Stars,⁹⁰ QOP,⁷⁸ NGYCP⁷⁷ and YDPD¹¹² included clustering generally by including site as a covariate (a 'fixed effects' model, which is appropriate when few clusters are included). The evaluation of Cool Girls, Inc.,⁹⁵ reported a test for clustering that was not significant; intraclass correlation coefficients were reported as being between 0% and 6% and thus clustering effects were not included in analyses.

Other sources of bias

Almost all included non-randomised evaluations, with the exception of Stay SMART, used and reported appropriate methods to control for confounding variables in analyses. Moreover, three of the four randomised trials (Tierney,⁹⁸ linked to the study by Rhodes *et al.*;⁹⁶ Millenky *et al.*,¹⁰⁷ linked to the study by Schwartz *et al.*;⁷⁷ and Cross *et al.*⁹⁰) used extensive sets of participant-level covariates in analyses.

Effects of interventions on substance use

We first present a narrative synthesis of findings from included evaluations before presenting the results of meta-analyses. We conclude by narratively synthesising subgroup analyses of intervention effects as presented in included evaluations.

Narrative synthesis of effects on substance use

Interventions used a variety of substance-use measures (*Table 5*). Reported effects on substance use behaviours were, in general, mixed.

TABLE 5 Measures used in studies reporting substance-use outcomes

Programme	Outcomes as reported	Remarks
All Stars ⁹⁰	Drinking initiation	
	Smoking initiation	
	Inhalant initiation	
	Marijuana initiation	
	Other drug initiation	Unclear: understood to cover additional illicit drug use
MAP ⁹³	Past month drug use (binary)	Includes all substance use
	Past year variety of drug use (continuous)	Unclear: understood to include all substance use
BBBS ⁹⁶	Likelihood of smoking	
	Likelihood of initiating alcohol use	
	Likelihood of initiating drug abuse	Includes all illicit drug use
Cool Girls, Inc. ⁹⁵	Drug use avoidance in past 6 months	Understood to include all substance use
NGYCP ⁷⁷	Binge drinking (five or more drinks in a row) in past 14 days	
	Charged with a drug crime in past 12 months	
	Convicted of a drug crime in past 12 months	
	Ever used other illegal drugs	Includes illicit drug use besides marijuana
	Frequent illegal drug use (six or more occasions) in past 12 months	Includes illicit drug use besides marijuana
PYDC ⁷⁹	Frequent marijuana use (10 or more occasions) in past 12 months	
	Alcohol use in past 30 days	
	Marijuana use in past 30 days	
	Other drug use in past 30 days	Includes illicit drug use besides marijuana
QOP ⁷⁸	Any type of drug use in past 30 days	Includes all substance use
	Drinking in past 30 days	
YARP ⁸⁶	Drinking on 8 or more days in past 30 days	
	Drunk or high at school in past 12 months	
	Binge drinking in past 30 days (past month)	Slightly different wording between reports
	Binge drinking on eight or more days in past 30 days (past month)	Slightly different wording between reports
	Smoked cigarettes or used tobacco in past month	
	Smoked cigarettes or used tobacco daily in past month	
	Used any illegal drug in past 30 days (past month)	Slightly different wording between reports
YARP ⁸⁶	Frequency of alcohol use in past 30 days	Data were unavailable for this outcome
	Frequency of cannabis use in past 30 days	

TABLE 5 Measures used in studies reporting substance-use outcomes (*continued*)

Programme	Outcomes as reported	Remarks
YPDP ⁷	Drunk monthly in past 6 months	Reported at first follow-up
	Drunk monthly in past 3 months	Reported at second follow-up
	Monthly use of cannabis in past 6 months	Reported at first follow-up
	Weekly use of cannabis in past 6 months	Reported at first follow-up
	Weekly use of cannabis in past 3 months	Reported at second follow-up

The evaluation of MAP,⁹³ which combined randomised and non-randomised evaluations in a sample of approximately 440 middle-school students, showed that the variety of drug use over the past year (i.e. the average number of different drugs or substances used, although differences between illicit drugs and other substances were not made clear) was significantly less at post intervention (p -value < 0.01) in intervention participants (0.038) than in control participants (0.086), although it does not appear that this finding was adjusted for baseline differences (0.036 in intervention vs. 0.053 in control). Although timing of post-intervention measurement was not made explicit, it appears to have been conducted at the end of the school year. Intervention participants were compared against those not receiving the programme, a group comprising those recruited to a no-treatment comparison group as well as those assigned to a waitlist or to some sort of alternative service.

In the randomised All Stars trial, Gottfredson *et al.*⁹² found no statistically significant differences between All Stars and a control group offered minimal recreational activities on any past month substance use or on initiation of any of smoking, drinking, marijuana, inhalant or other (unspecified) illicit drug. The included sample consisted of 416 students completing the post-test survey near the end of the school year. Probabilities of substance use were extremely similar across intervention and control groups, with the exception of marijuana initiation, in which the intervention group reported a covariate-adjusted probability of 1% compared with one of 11% for the control group, although this difference was no greater than would be expected by chance.

Tierney⁹⁸ found that the probability of initiating any drug abuse was 45.8% less (p -value < 0.05) in those randomised to receive the BBBS intervention than in the young people randomised to the control group, who did not receive a mentor match. BBBS was also associated with a 27.4% decrease (p -value < 0.10) in the probability of initiating alcohol use in youth who did not report alcohol use at baseline (adjusted probability of 11.5% in the control group vs. 6.2% in the intervention group), although effects for smoking (a 19.7% reduction in the intervention group from a control group probability of 17.2%) were not statistically significant. The sample size included here was 959 young people at both baseline and post-test surveys, taken 18 months apart.

Kuperminc *et al.*,⁹⁵ in the evaluation of Cool Girls, Inc., found that girls enrolled in the intervention tended towards less 'drug use avoidance' overall than girls not in the programme (OR 0.62), although girls who were engaged in the mentoring component of the programme did tend towards more drug-use avoidance than all other girls (OR 1.53). At post test, unadjusted probabilities of past drug use were 41.9% in the intervention group and 36.0% in the control group. Neither of these findings was statistically significant. This was a non-randomised evaluation including 175 girls, with pre-testing near the start of the school year and post-testing near the end.

Findings for the randomised trial of the NGYCP were described in two study reports^{107,108} covering post-intervention ($n = 1196$) and 18-month ($n = 1173$) follow-up. At post intervention, in covariate-adjusted linear probability models, intervention participants were less likely to report binge drinking in the past two weeks (intervention group 2.8% vs. control group 4.7%; p -value = 0.076) than control group members,

who did not receive the intervention. Similar models did not yield significant results (all p -values > 0.10) for frequent past-year marijuana use (22.5% vs. 25.2%), any use of illicit non-marijuana drugs (24.1% vs. 23.1%) or frequent past-year illicit drug use (5.6% vs. 4.4%). Intervention group participants were less likely to be charged with a drug crime (2.9% vs. 5.3%; p -value = 0.030), but were neither more nor less likely to be convicted of a drug crime (1.4% vs. 1.9%; p -value > 0.10). At 18-month follow-up, intervention participants were not significantly more or less likely (all p -values > 0.10) to report binge drinking (26.1% vs. 30.2%), frequent past-year marijuana use (26.0% vs. 24.4%), or frequent past-year illegal drug use (4.7% vs. 4.2%). Intervention group participants were, however, more likely to report any use of illicit non-marijuana drugs (28.2% vs. 23.2%; p -value = 0.044). Intervention participants were not more or less likely to have been convicted of a drug crime at this follow-up (8.1% vs. 5.9%; p -value > 0.10). Outcomes for being charged with a drug crime were not reported at follow-up.

Tebes *et al.*⁷⁹ measured outcomes at post-intervention and 4-month follow-up in their non-randomised evaluation ($n = 304$ at baseline) of the PYDC. They examined separately use of alcohol, marijuana, tobacco, illicit drugs and any type of drug. Results for tobacco were not presented. At post-intervention follow-up, participants in intervention and control groups were not significantly different in the degree to which their substance use changed from baseline across all outcomes: use of alcohol (OR 1.179, 95% CI 0.49 to 2.87; predicted probabilities 12.2% intervention vs. 5.7% control), use of marijuana (OR 1.759, 95% CI 0.66 to 4.68; predicted probabilities 7.6% vs. 3.7%), use of illicit drugs (OR 1.266, 95% CI 0.52 to 3.10; predicted probabilities 8.7% vs. 4.8%) or use of any drug (OR 1.694, 95% CI 0.74 to 3.90; predicted probabilities 19.3% vs. 10.3%). In contrast, at 4-month follow-up, intervention participants' substance use had significantly decreased from baseline compared with control group participants across all outcomes: alcohol (OR 0.365, 95% CI 0.15 to 0.90; predicted probabilities 7.2% vs. 5.8%), marijuana (OR 0.178, 95% CI 0.08 to 0.42; predicted probabilities 12.1% vs. 17.4%), illicit drugs (OR 0.188, 95% CI 0.08 to 0.44; predicted probabilities 12.8% vs. 16.8%) and any drug use (OR 0.289, 95% CI 0.13 to 0.67; predicted probabilities 18.4% vs. 19.9%). Control groups received an ASP without the PYD curriculum. Note that predicted probabilities by arm as reported do not account for change from baseline, which was the outcome metric used in this evaluation.

Outcomes for the randomised trial of QOP⁷⁸ were reported in several forms across different study reports using the evaluation sample of 1069 students. We refer here to three study reports¹⁰³⁻¹⁰⁵ presenting unadjusted effect estimates. Other evaluations^{78,100-102} reported effect estimates with additional covariates included. Although timings for outcome evaluations were not exact, the first post-intervention measurements were taken near the end of the fourth year of the programme, around which time participants should have been preparing to complete secondary school. These measurements were supplemented with a telephone survey completed about 7 months after this first survey. Both the post-intervention and 7-month follow-ups are reported in Schirm *et al.*¹⁰³ Findings from linear probability models weighting for programme site demonstrated that, at first measurement, intervention participants were *more* likely than control participants (i.e. those not in the programme) to report any drinking in the past 30 days (intervention 40% vs. control 33%; p -value < 0.05), although not more frequent drinking (i.e. on eight or more days in the past month) (11% vs. 11%; p -value > 0.10), binge drinking (24% vs. 20%; p -value > 0.10) or frequent binge drinking (7% vs. 5%; p -value > 0.10). Intervention participants were also neither more nor less likely to be drunk or high at school in the past 12 months (20% vs. 20%; p -value < 0.10). Intervention participants were, however, *more* likely to report any illegal drug use in the past 30 days (34% vs. 28%; p -value < 0.05). In the 7-month telephone survey, past-month binge drinking (19% vs. 23%), past-month frequent binge drinking (5% vs. 4%) and any past-month illegal drug use (16% vs. 19%) were not significantly different between groups (all p -values > 0.10). A subsequent set of measurements on binge drinking and illegal drug use was taken between 3 and 4 years following the anticipated graduation date and is reported in Schirm and Rodriguez-Planas.¹⁰⁴ Neither past-month binge drinking (25% vs. 31%) nor frequent past-month binge drinking (7% vs. 5%) were significantly different between groups, but in contrast to the initial set of measurements, intervention group participants were now *less* likely to have reported any illicit drug use in the past month (12% vs. 18%; p -value < 0.05). Finally, measurements were taken around 6 years after anticipated graduation.¹⁰⁵ Past-month binge

drinking (31% vs. 31%), frequent binge drinking (8% vs. 6%) and illicit drug use (12% vs. 13%) were not significantly different between groups (all p -values > 0.10). Past-month tobacco use (34% vs. 34%) and daily past-month tobacco use (22% vs. 24%), which were also reported for the first time in this evaluation, were not significantly different between groups (both p -values > 0.10).

At post-intervention follow-up, Berg *et al.*⁸⁶ reported a marginally significant effect of YARP in reducing marijuana use (standardised path -0.12 ; p -value = 0.053). Effects of the intervention on reductions in alcohol use were not statistically significant and not reported. Arm-level descriptive statistics for included outcome measures were not reported. This non-randomised evaluation used a recruited comparison sample of youth not receiving the intervention for a total of 316 in the evaluation sample.

Finally, the non-randomised evaluation of YDP, reported by Wiggins *et al.*,^{46,112} collected outcomes on alcohol consumption and cannabis use at post-intervention and at 9-month follow-up. At post-intervention follow-up, programme participants were more likely than control group participants to use cannabis monthly or more often (OR 1.56, 95% CI 0.93 to 2.63; unadjusted probabilities 23% intervention vs. 21% control), or more than once a week (OR 1.41, 95% CI 0.75 to 2.68; unadjusted probabilities 15% vs. 13%), although these differences did not rise to statistical significance. Intervention and control group participants were equally likely to be drunk monthly or more often (OR 0.98, 95% CI 0.65 to 1.47; unadjusted probability 37% in both groups). At 9-month follow-up, intervention groups continued to be more likely to use cannabis more than once a week (OR 1.97, 95% CI 0.93 to 4.17; unadjusted probabilities 16% vs. 11%) and appeared to be somewhat more likely to be drunk monthly or more often (OR 1.20, 95% CI 0.78 to 1.84; unadjusted probabilities 39% vs. 32%), although again these differences were not statistically significant. Control group participants were similar in vulnerability to those enrolled in the intervention and may have been receiving some sort of youth service. A sample of 2724 at baseline was reduced to 904 by the second follow-up by attrition.

Meta-analyses of effects on substance use

Overview of included data

After data transformation and preparation (see *Appendix 11*), we included 54 effect sizes addressing substance-use outcomes from 12 reports of nine studies. We did not include effect sizes from the evaluation of Stay SMART,¹¹¹ as these were not clear from study reports and we obtained no further information from study authors. We were also unable to include an effect estimate for alcohol use from the evaluation of the YARP intervention,⁸⁶ despite multiple attempts to contact the study's surviving authors. The evaluation of the PYDC⁷⁹ measured the difference between intervention and control groups in a change model from baseline to follow-up of risk for substance use. When intervention effects are compared across studies as standardised mean differences, outcomes comparing change scores and those comparing 'final values' between intervention and control groups cannot be analysed in the same model, because the standard deviations are incommensurate.⁶⁷ We did not have sufficient information to convert these outcomes into a metric completely consistent with the other studies. We decided to include findings from Tebes *et al.*⁷⁹ in the main analyses but also to re-estimate all relevant multilevel meta-analysis models without these findings.

All substance-use outcomes

Findings from 12 reports of nine studies (54 effect sizes) informed the analyses of all substance-use outcomes. Included PYD interventions did not have a statistically significant effect on substance-use outcomes, either across all time points ($d = 0.079$, 95% CI -0.025 to 0.183) or in the short term ($d = 0.086$, 95% CI -0.025 to 0.197) (*Table 6* and *Figures 8* and *9*). From a perspective of public health significance, pooled effect sizes would have been considered very small. Outcomes across studies were not characterised by a large degree of heterogeneity (programme-level $I^2 = 27\%$ in both analyses). Both the all-points and short-term meta-analyses were robust to sensitivity analyses.

TABLE 6 Substance-use outcome meta-analyses

Outcomes	Analysis	Main analysis				Sensitivity analysis			
		ES (95% CI)	k	n	P (%), programme level	P (%), outcome level	Cochran's Q (df; p-value)	High variance, ES (95% CI)	Low variance, ES (95% CI)
All time points	All substance-use outcomes	0.079 (-0.025 to 0.183)	9	54	27	20	132.92 (53; <0.0001)	0.078 (-0.029 to 0.185)	0.080 (-0.021 to 0.181)
	Omnibus drug-use outcomes	0.127 (-0.035 to 0.290)	7	11	0	45	25.45 (10 to 0.005)	0.133 (-0.040 to 0.306)	0.121 (-0.036 to 0.277)
	Illicit drug use	0.047 (-0.117 to 0.212)	6	21	34	26	63.17 (20; <0.0001)		
	Alcohol	0.050 (-0.063 to 0.163)	6	18	21	24	37.39 (17 to 0.003)		
	Smoking	0.053 (-0.038 to 0.143)	3	4	0	0	1.42 (3; 0.71)		
	All substance-use outcomes	0.086 (-0.025 to 0.197)	9	36	27	14	100.08 (35; <0.0001)	0.0852 (-0.0294 to 0.1998)	0.0862 (-0.0219 to 0.1943)
Short-term time points	Omnibus drug-use outcomes	0.169 (0.012 to 0.326)	7	10	0	36	19.65 (9; 0.02)	0.177 (0.008 to 0.347)	0.158 (0.006 to 0.310)
	Illicit drug use	0.050 (-0.141 to 0.242)	6	14	8	60	46.54 (13; <0.0001)		
	Alcohol	0.070 (-0.084 to 0.224)	6	10	46	0	22.64 (9; 0.01)		

df, degrees of freedom; ES, effect size.

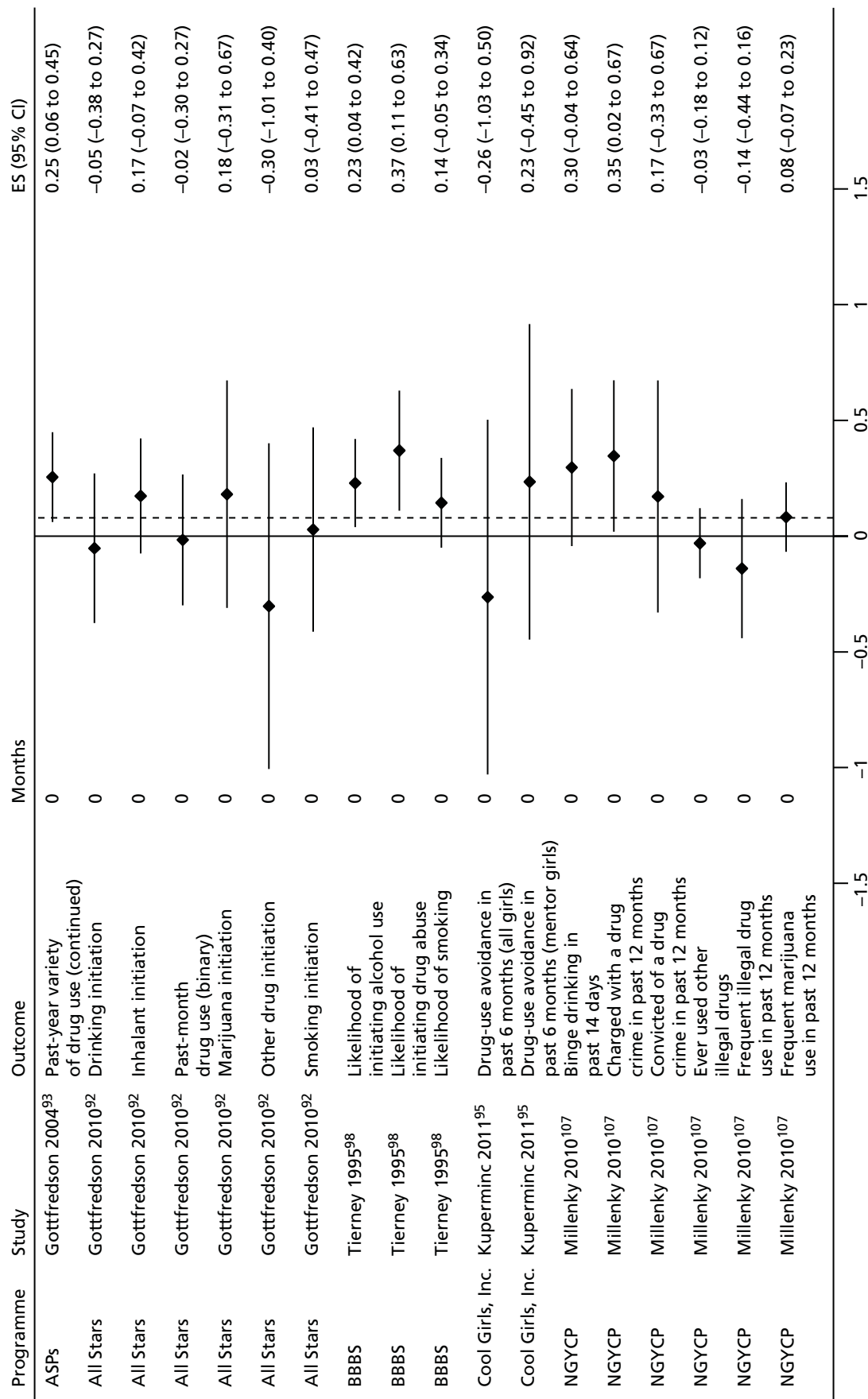


FIGURE 8 All substance-use outcomes. ES, effect size. (continued)

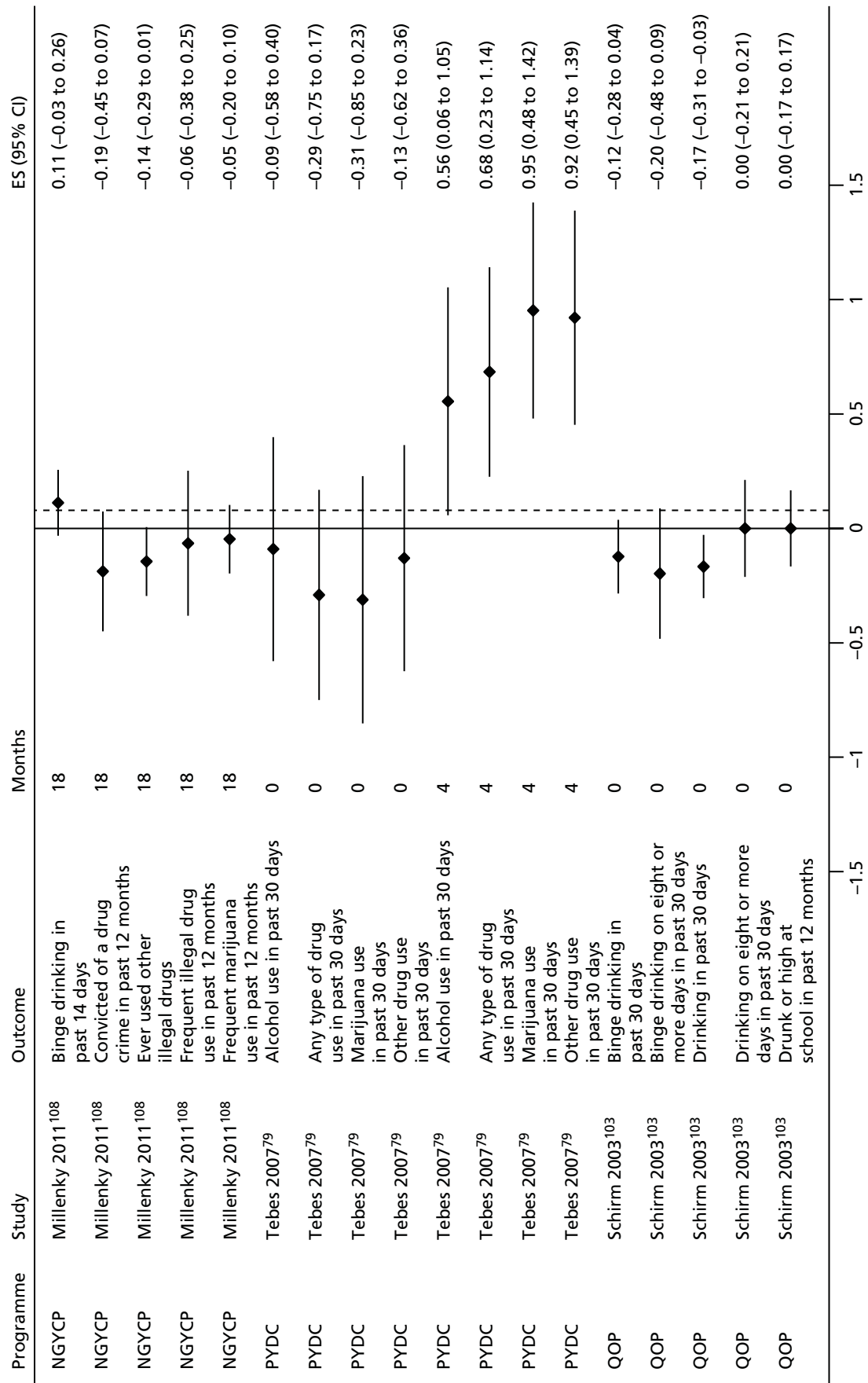


FIGURE 8 All substance-use outcomes. ES, effect size. (continued)

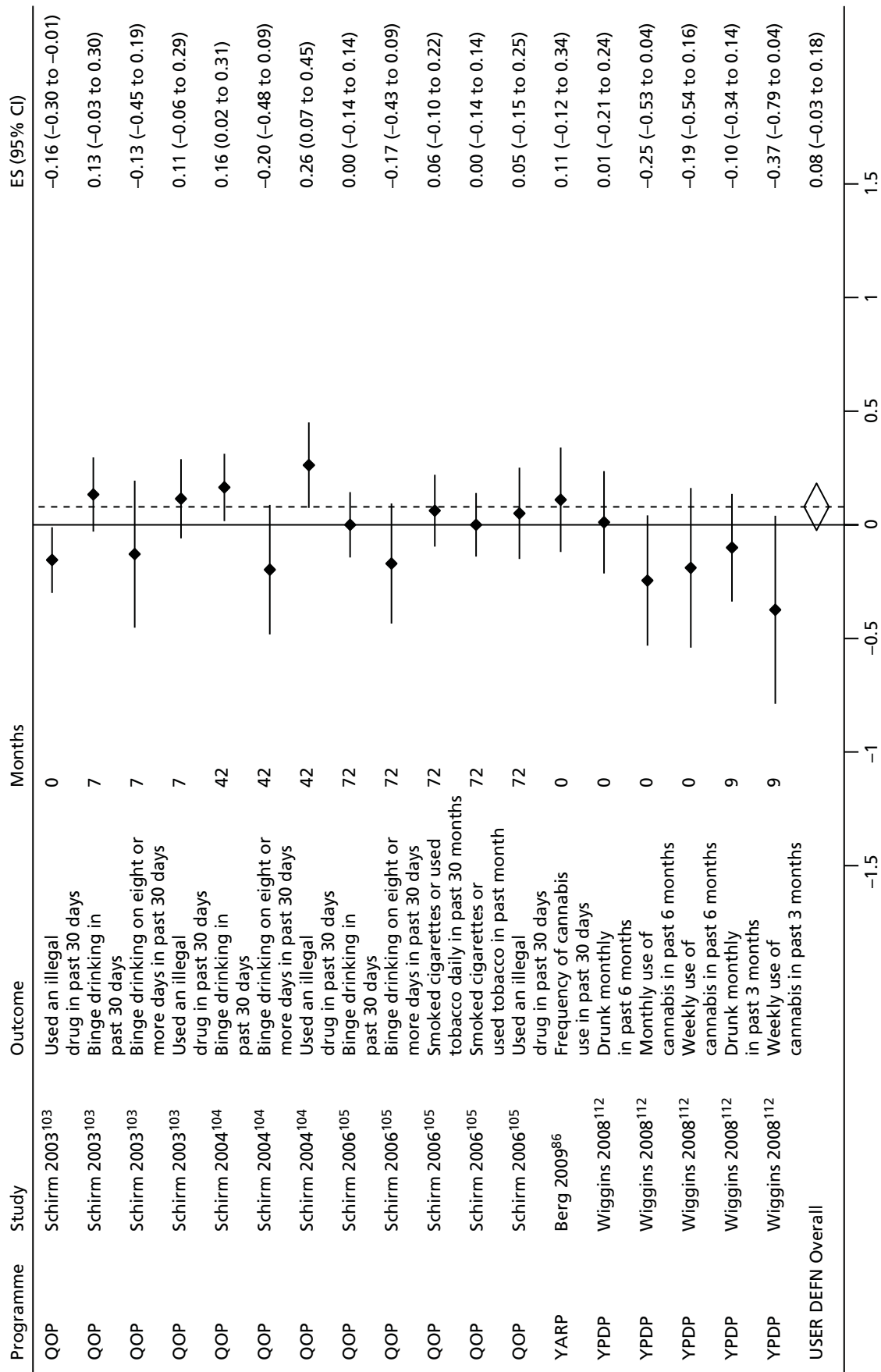


FIGURE 8 All substance-use outcomes. ES, effect size.

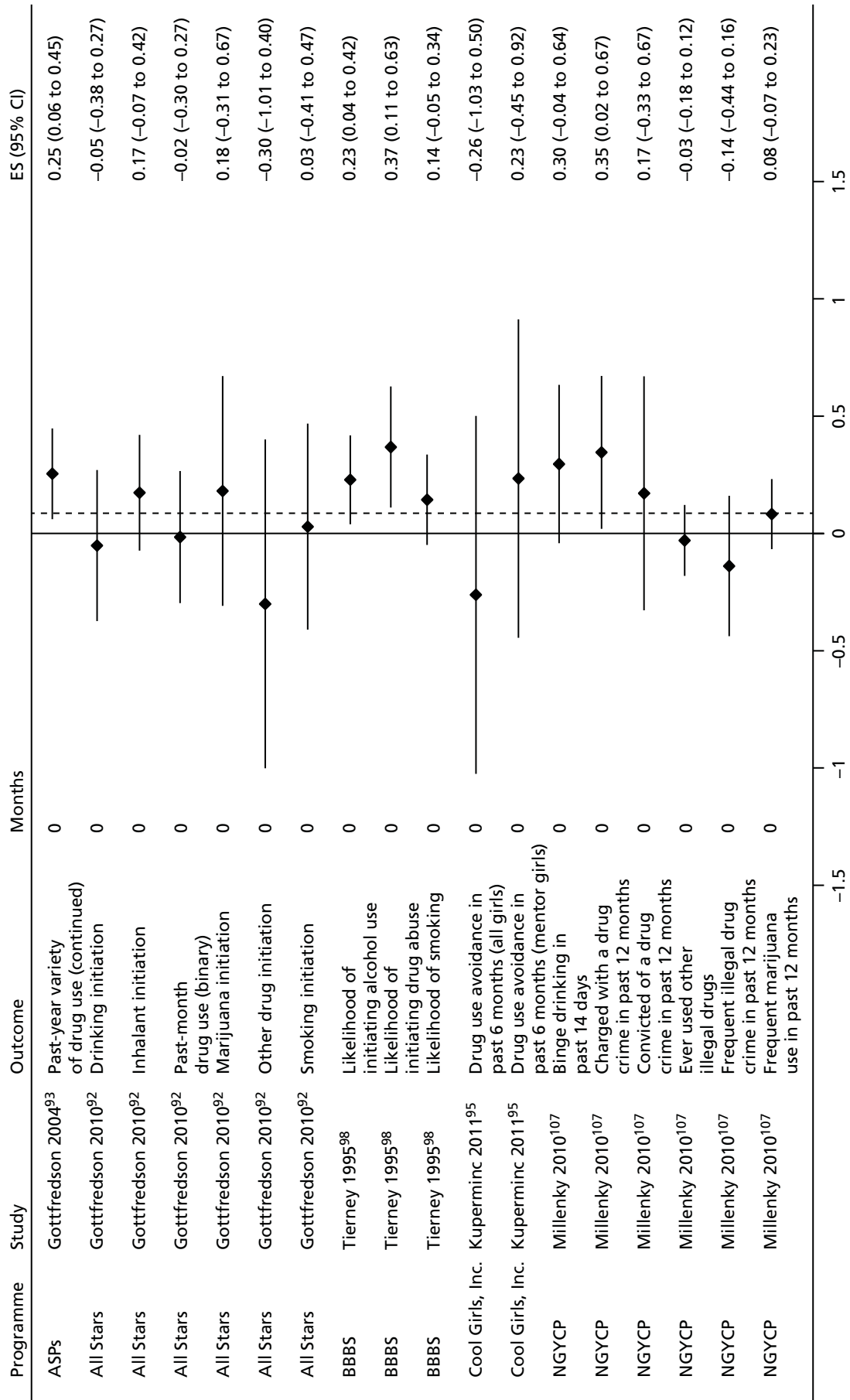


FIGURE 9 Short-term substance-use outcomes. ES, effect size. (continued)

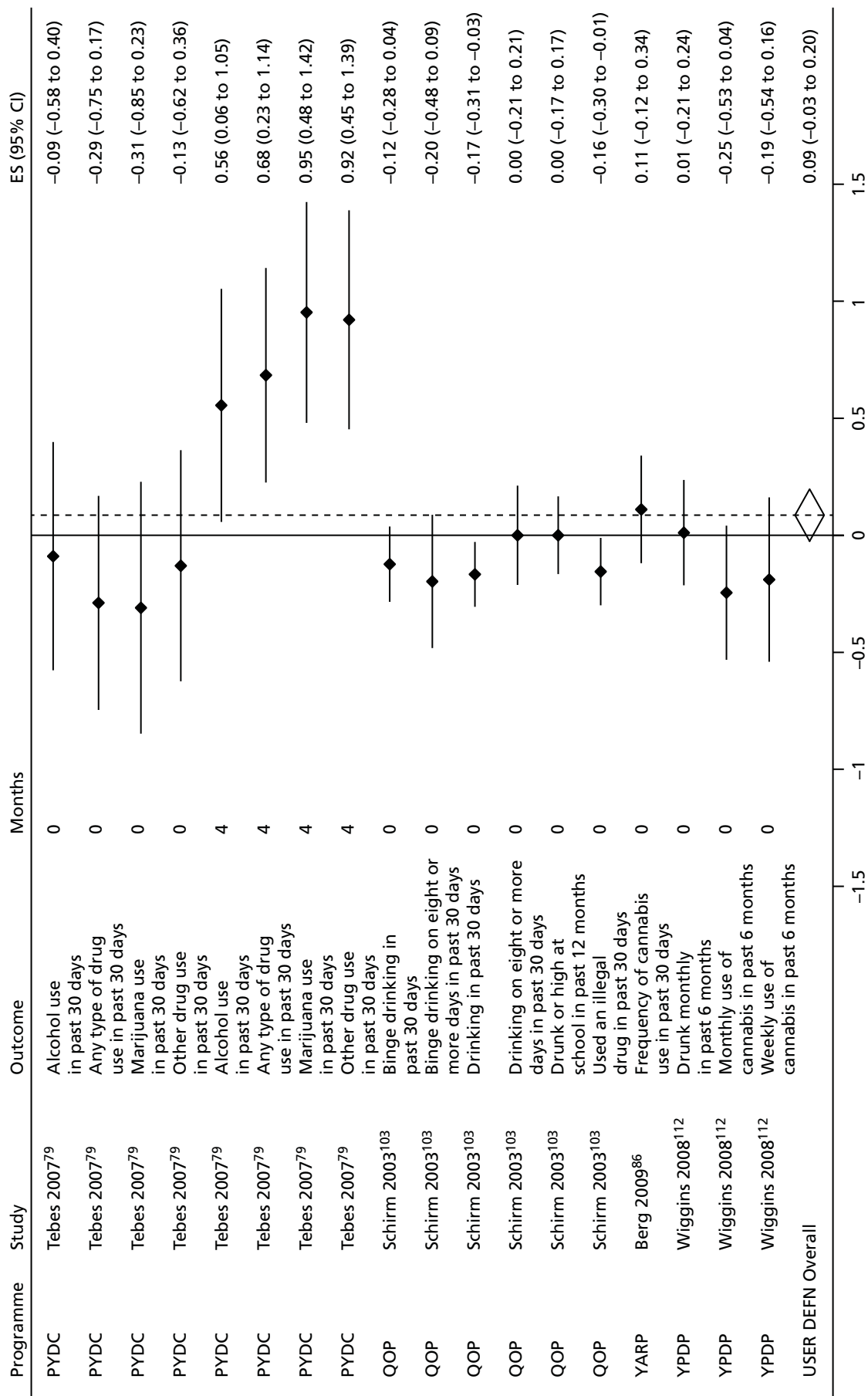


FIGURE 9 Short-term substance-use outcomes. ES, effect size.

Omnibus substance-use outcomes

Findings from seven studies (11 effect sizes) informed the analyses of omnibus substance-use outcomes (i.e. outcomes where the measure of substance use reported encompassed both illicit drug use and other substances). Although there was no statistically significant effect across all time points ($d = 0.127$, 95% CI -0.035 to 0.290), the meta-analysis of short-term outcomes yielded a small but statistically significant effect ($d = 0.169$, 95% CI 0.012 to 0.326) (see *Table 6* and *Figures 10* and *11*). Both findings were robust to all sensitivity analyses and demonstrated little heterogeneity at the programme level ($I^2 = 0\%$ for both analyses). However, it is worth pointing out that the difference between the analyses reflects the inclusion in the former but not the latter analysis of one effect size: the odds of being convicted of a drug offence measured at 18 months, reported in Millenky *et al.*¹⁰⁸ as part of the evaluation of the NGYCP. Thus, this statistically significant finding should be interpreted with caution.

Illicit drug-use outcomes

Findings from six studies informed the analyses of illicit drug-use outcomes. PYD interventions did not have a statistically significant effect on illicit drug-use outcomes either across all time points ($d = 0.047$, 95% CI -0.117 to 0.212) or in the short term ($d = 0.050$, 95% CI -0.141 to 0.242) (see *Table 6* and *Figures 12* and *13*). Both meta-analyses were robust to sensitivity analyses, although removal of findings from Tebes *et al.*⁷⁹ yielded negative pooled effect sizes in both meta-analyses. These meta-analyses included 21 and 14 effect sizes, respectively. Programme-level heterogeneity was small to moderate in the first analysis ($I^2 = 34\%$) and nearly negligible in the short-term meta-analysis ($I^2 = 8\%$).

Alcohol-use outcomes

Meta-analyses included findings from six studies, with 10 of 18 effect sizes included in the short-term meta-analysis. PYD interventions had small and statistically non-significant effects reducing alcohol use both across all time points ($d = 0.050$, 95% CI -0.063 to 0.163) and in the short term ($d = 0.070$, 95% CI -0.084 to 0.224) (see *Table 6* and *Figures 14* and *15*). Findings were robust across sensitivity analyses. Some programme-level heterogeneity was present in the meta-analysis with all time points ($I^2 = 21\%$), but heterogeneity was large in the short-term outcomes meta-analysis ($I^2 = 46\%$). The small number of interventions included in the short-term outcomes meta-analysis precluded further exploration of heterogeneity.

Smoking outcomes

Analyses for smoking outcomes included only four effect sizes from three studies and, thus, are presented with caution. PYD interventions had a small, statistically non-significant effect reducing smoking outcomes ($d = 0.053$, 95% CI -0.038 to 0.143) with no programme-level heterogeneity ($I^2 = 0\%$) (see *Table 6* and *Figure 16*). We do not present meta-analyses for short-term smoking outcomes, as only two effect sizes would have been included.

Substance-use subgroup analyses by sex

The use of covariate-adjusted models in almost all studies precluded the use of metaregression to explore group-level moderators of intervention effect. This is because covariate-adjusted models present 'conditional' estimates of intervention effect that apply in magnitude only to the specific population defined by the regression models used in adjustment. Moreover, the quality and number of included studies and the lack of substantial study-level heterogeneity would have made metaregression specious. As a rule of thumb, metaregression generally requires at least 10 studies for each categorical moderator tested.⁶⁷ Our metaregression models would also have been confounded by study quality and risk of bias between levels of our covariate. Instead, we present narratively the findings of studies in which subgroup analyses were undertaken for sex. Evaluations of three interventions^{78,96,98,103-105,112} undertook subgroup analyses by sex. Findings were mixed and do not present a clear picture.

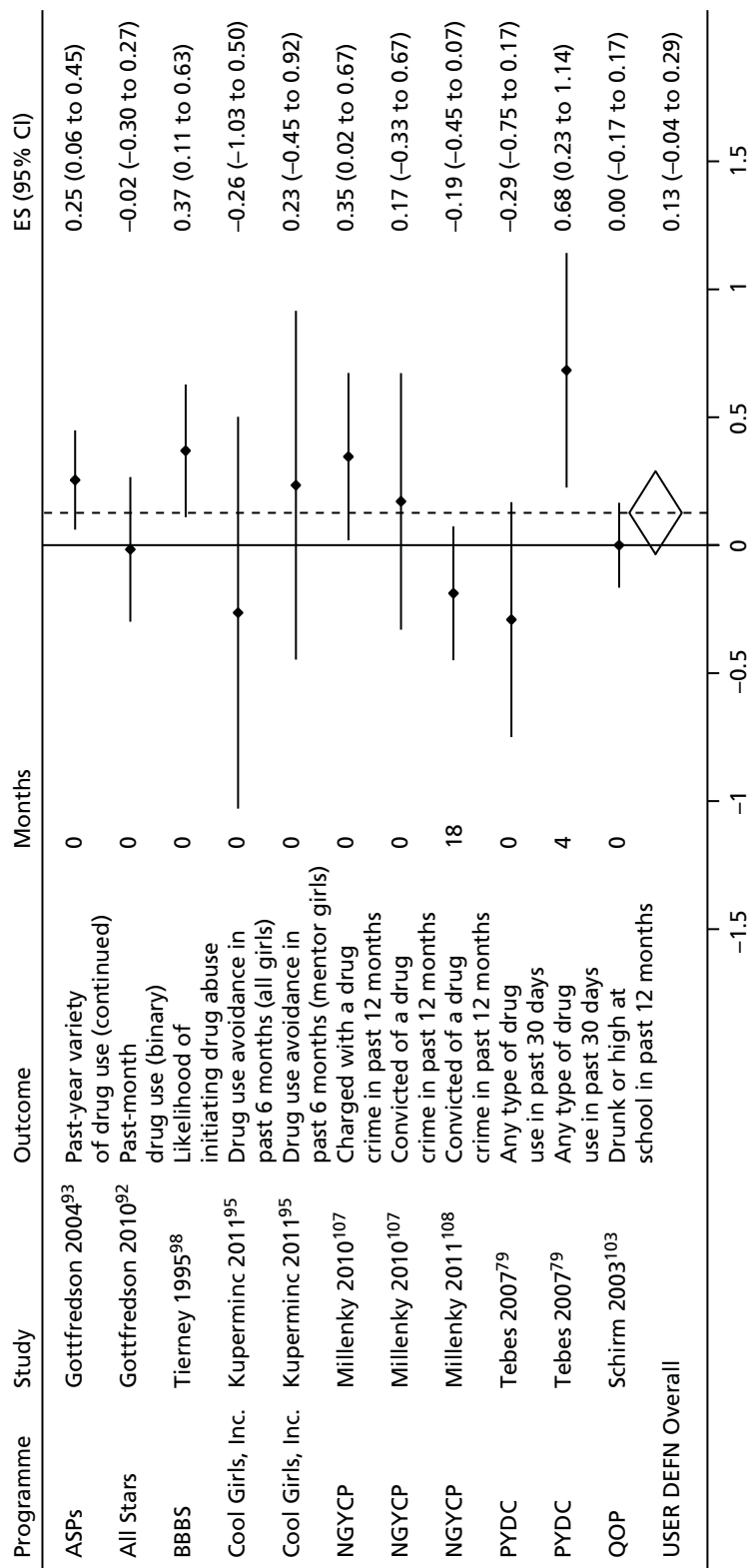


FIGURE 10 Omnibus substance-use outcomes. ES, effect size.

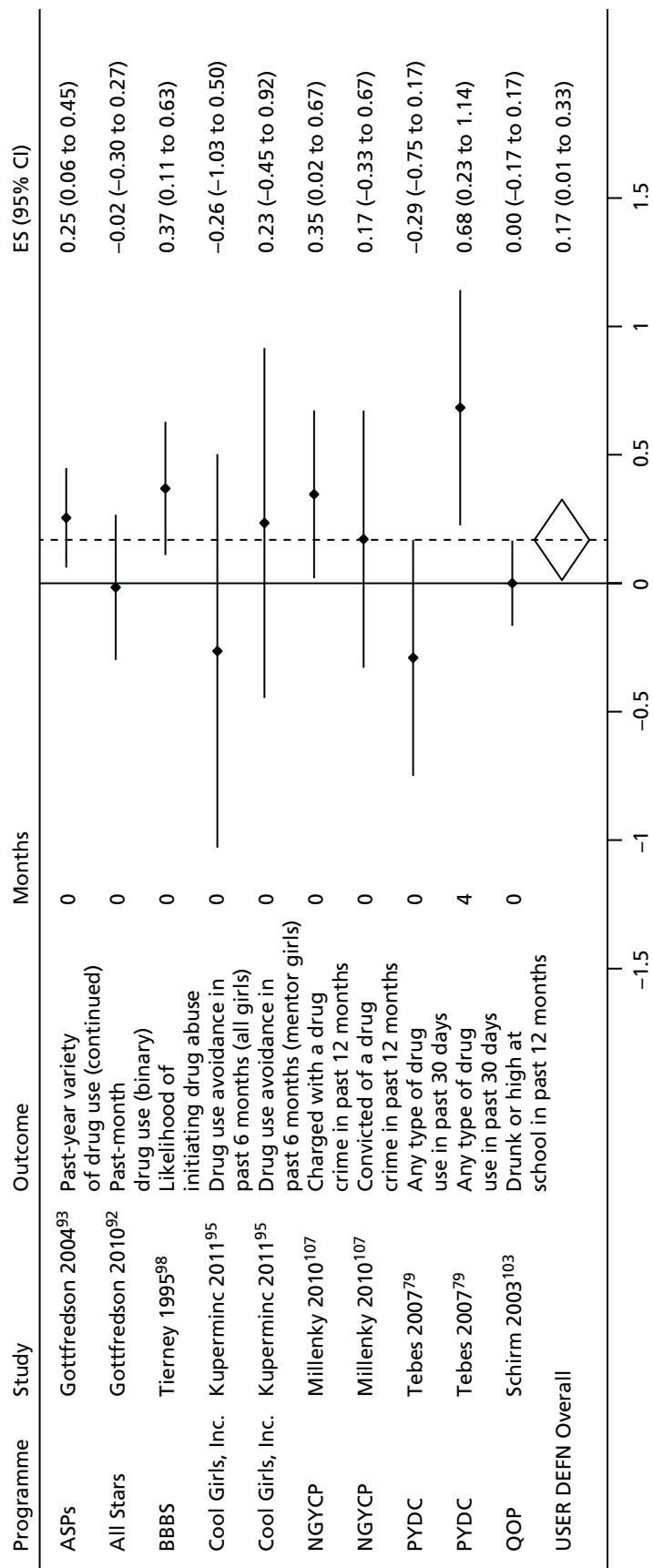


FIGURE 11 Short-term omnibus substance-use outcomes. ES, effect size.

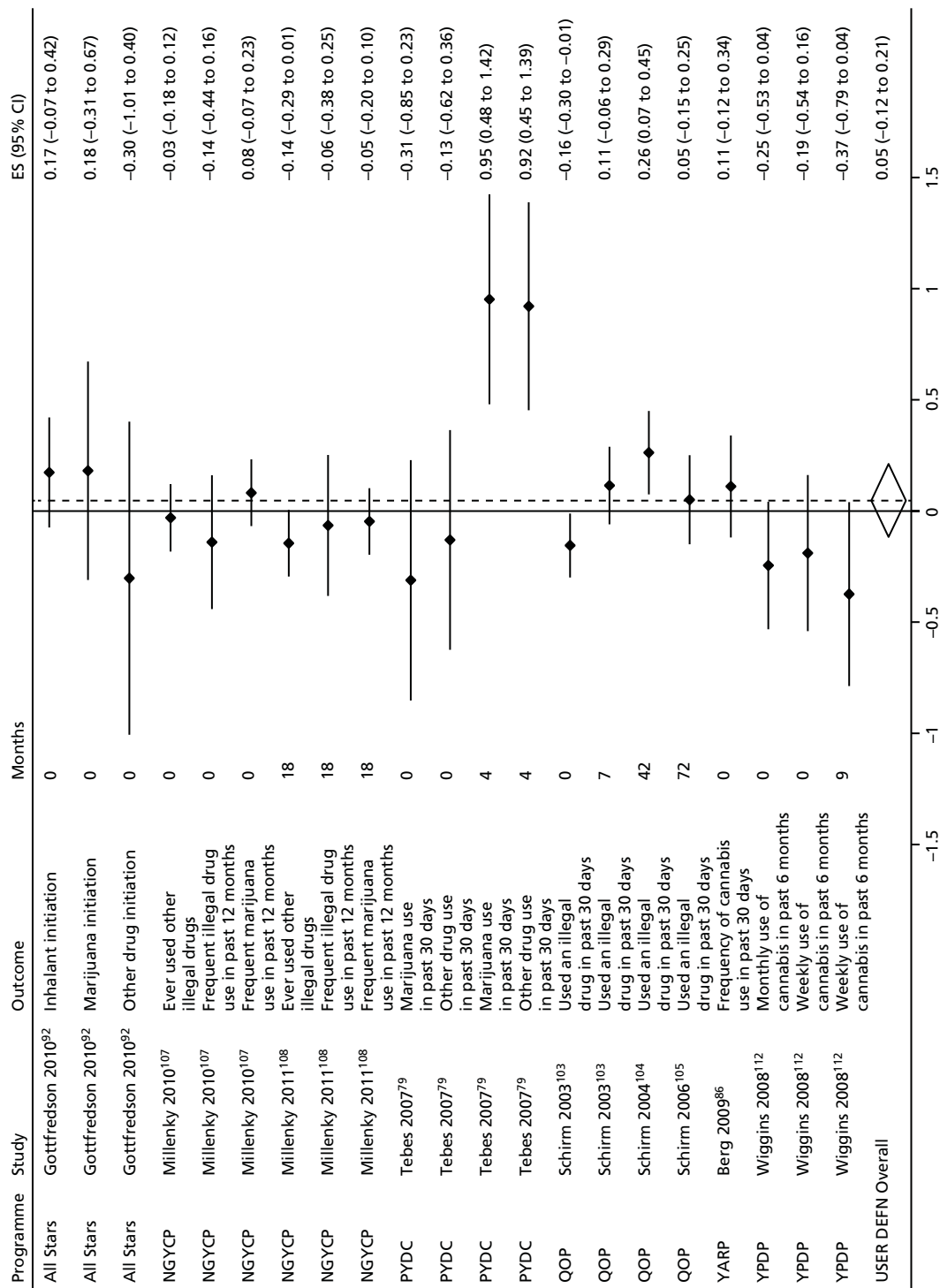


FIGURE 12 Illicit drug-use outcomes. ES, effect size.

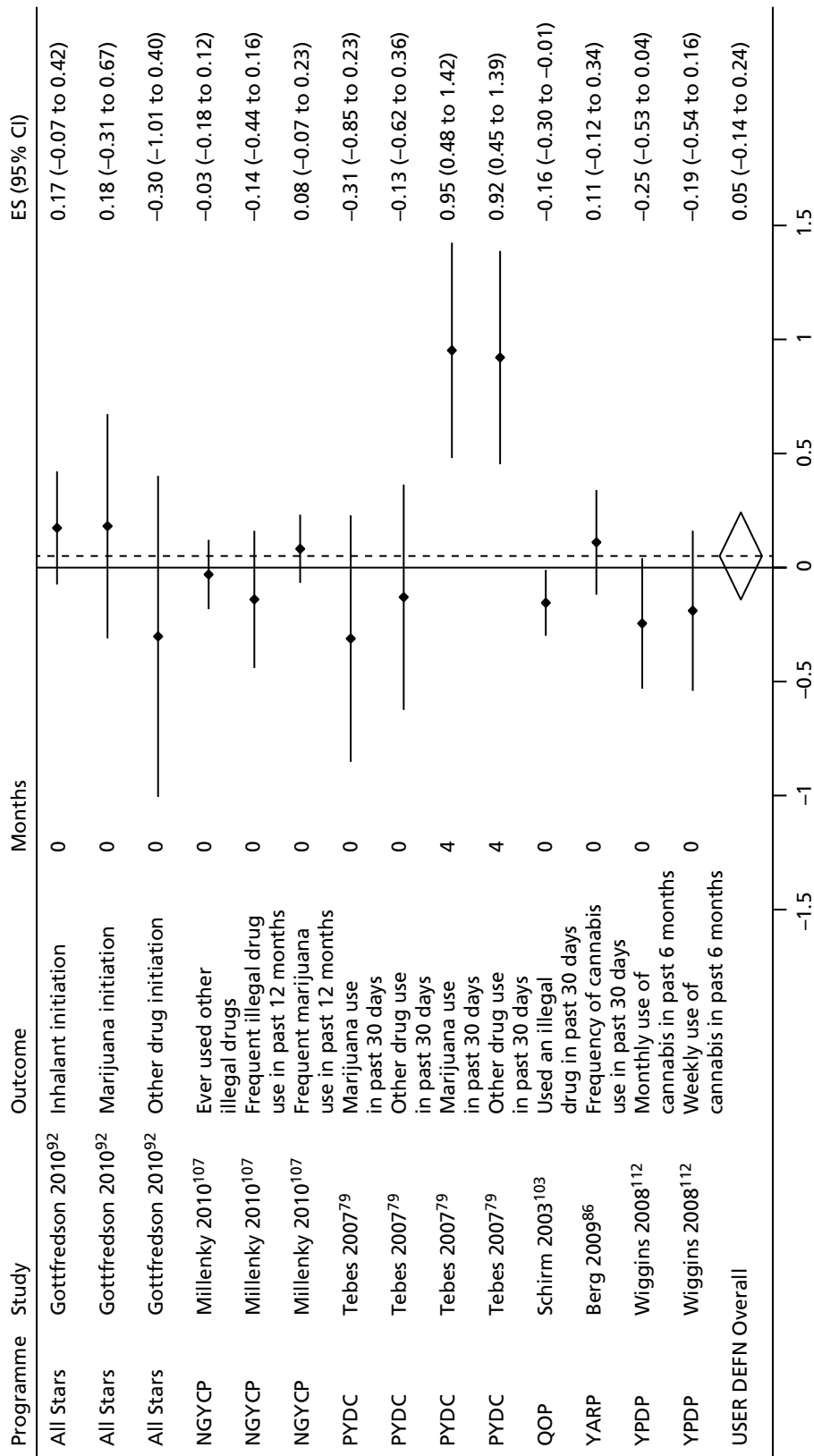


FIGURE 13 Short-term illicit drug-use outcomes. ES, effect size.

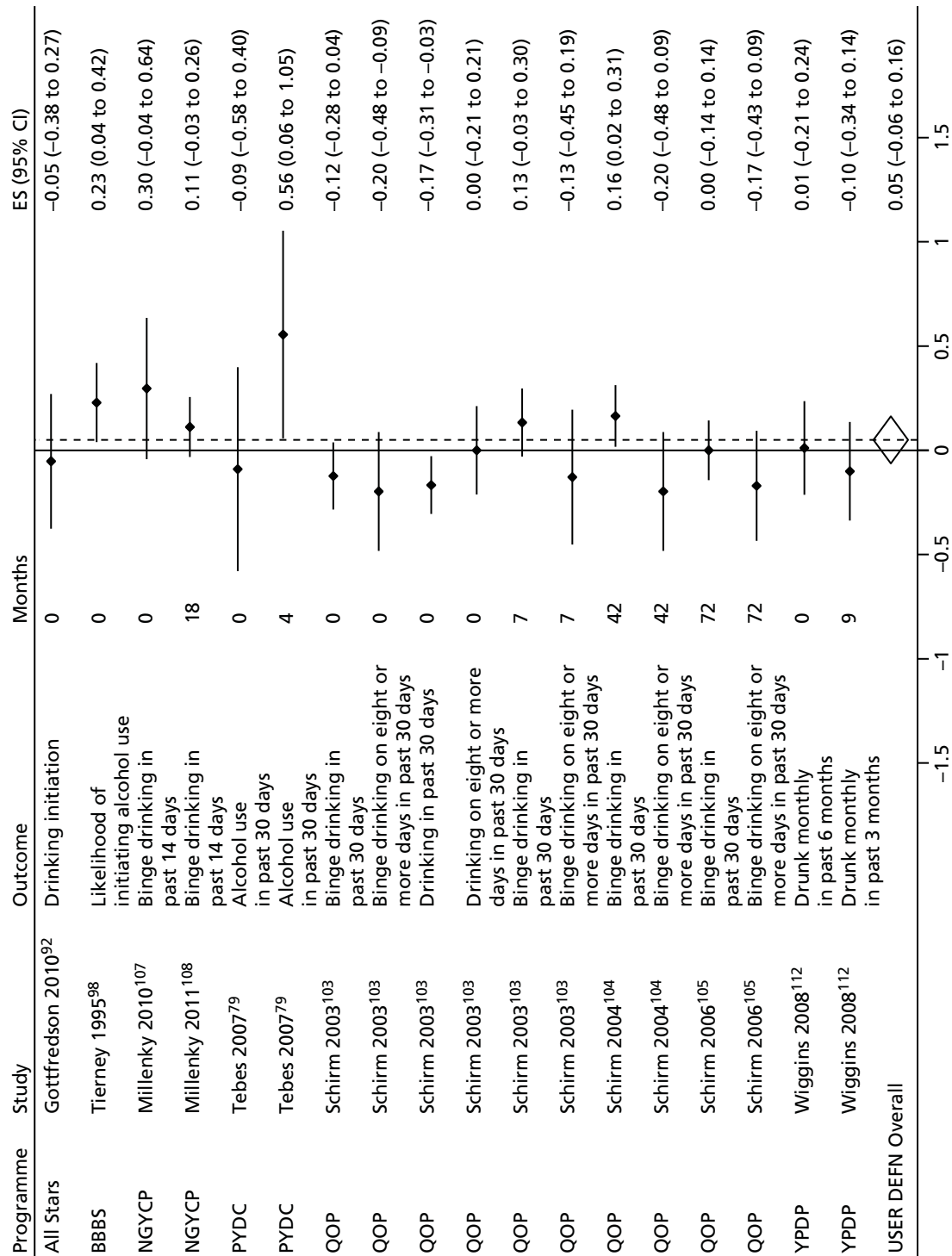


FIGURE 14 Alcohol outcomes. ES, effect size.

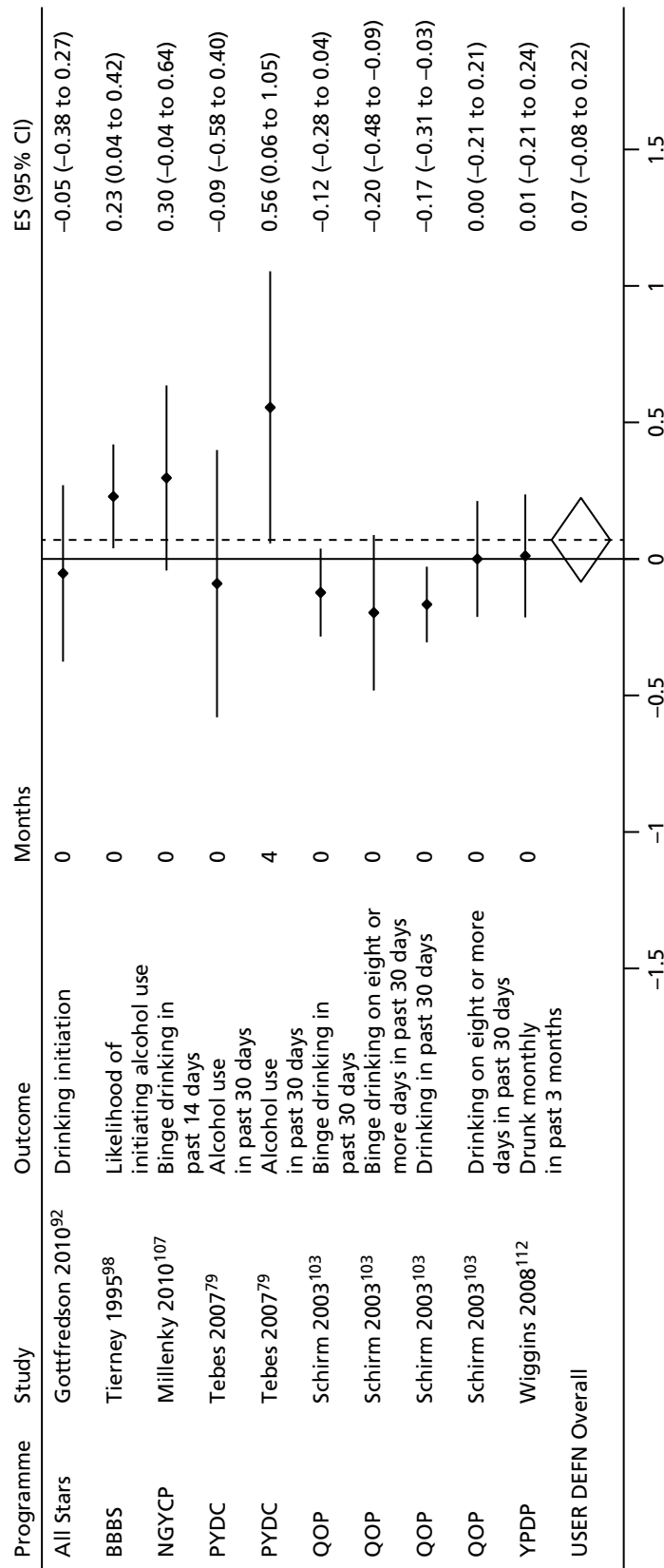


FIGURE 15 Short-term alcohol outcomes. ES, effect size.

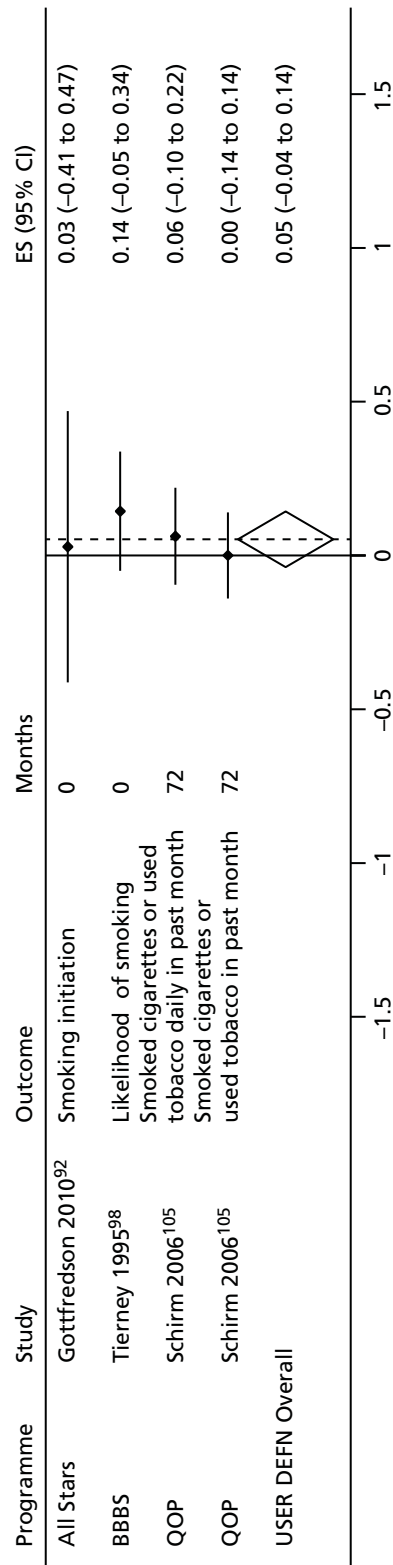


FIGURE 16 Smoking outcomes. ES, effect size.

In the evaluation of BBBS,^{96,98} the decrease in the likelihood of initiation of drug abuse experienced by youth in the intervention group was felt more strongly in young men (55.0% decrease in risk; p -value < 0.05) than in young women, for whom the decrease in risk (26.6%) was not statistically significant. Findings were reversed for the initiation of alcohol use, with young women experiencing twice as much decrease in risk (38.8%) than young men (19.2%), although the differences were not statistically significant for either subgroup. Although the overall decrease in risk of smoking (19.7%) was not statistically significant, young men demonstrated a greater decrease in smoking risk (24.5%) than young women (9.9%). Tests of moderation did not reveal any significant differences between subgroups in any of these analyses.

Short-term impacts reported in Schirm *et al.*¹⁰³ for the QOP demonstrated that, although the overall risk difference in probabilities for any past-month binge drinking between intervention and control group was a non-significant 4% (i.e. 24% in the intervention group vs. 20% in the control group, as reported above), this was driven by a marginally significant (p -value < 0.10) risk difference of 7% in males favouring the control group, compared with a risk difference of 0% in women. Risk differences for any past-month illegal drug use were the same in men and women. Tests of moderation were not significant at the p -value < 0.10 level for any of these analyses.

Medium-term impacts¹⁰⁴ for the QOP revealed a different picture. The risk difference of 6% in past-month binge drinking, which favoured the intervention group, was driven by a risk difference of 12% in young men favouring the intervention group (p -value < 0.05), compared with a non-significant risk difference of 2% in women favouring the intervention group. The risk difference for frequent past-month binge drinking was 0% in men, but 4% favouring the control group (p -value < 0.05) in women. Finally, the 6% risk difference favouring the intervention group in past-month illicit drug use was decomposed into a marginally significant (p -value < 0.10) risk difference of 8% in men and a non-significant risk difference of 3% in women, both favouring the intervention group. Tests of moderation were not significant at the p -value < 0.10 level for any of these analyses.

The distribution of risk differences in the late-term evaluation¹⁰⁵ of the QOP showed that risk differences for past-month tobacco use, past-month daily tobacco use and any past-month binge drinking favoured the intervention in men but favoured the control in women, although none of the effect sizes in these subgroup analyses rose to significance. The overall risk difference for past-month frequent binge drinking favoured the control group, as did the risk differences for men and women separately. As in the prior evaluations, tests of moderation were not significant at the p -value < 0.10 level for any of these analyses.

Finally for YPDP, Wiggins *et al.*¹¹² reported stratified analyses for frequency of cannabis consumption more than once a week at the 9-month follow-up. Young men (OR 2.50, 95% CI 0.86 to 7.28) and young women (OR 2.27, 95% CI 0.71 to 7.33) were not substantially different on this outcome. Moderation was not tested in these analyses.

We further intended to examine the moderation of effects by economic status, ethnicity or area deprivation, although no intervention evaluations presented analyses in these subgroups. For BBBS, Tierney⁹⁸ presented subgroup analyses by combinations of sex and ethnicity. None of the moderation analyses for substance-use outcomes appeared to demonstrate that minority ethnic status was an effect modifier.

Effects of interventions on violence

We first present a narrative synthesis of findings from included evaluations before presenting the results of meta-analyses. We conclude by narratively synthesising subgroup analyses of intervention effects as presented in included evaluations.

Narrative synthesis of effects on violence

The three programmes reporting on violence (BBBS,^{96,98} NGYCP¹¹² and QOP⁷⁸) used a wide variety of outcome measures (*Table 7*). Findings were mixed, but tended towards the null.

In the RCT of BBBS, Tierney⁹⁸ found that intervention youth reported 0.85 fewer occasions of hitting someone in the past year than control group youth (adjusted mean of 1.83 in the intervention group vs. 2.68 in the control group; p -value < 0.05). Differences between groups in the average number of times being involved in a fight were not statistically significant (adjusted means of 1.52 vs. 1.54; p -value > 0.10). The sample size included here was of the 959 youth with both baseline and post-test surveys.

Drawing on post-intervention data ($n = 1196$) from the RCT of the NGYCP,¹⁰⁷ covariate-adjusted linear probability models demonstrated no significant differences in the probability of conviction for a violent crime (intervention 1.4% vs. control 1.2%; p -value = 0.748), in being charged with a violent crime (3.4% vs. 3.6%; p -value = 0.842) or in reporting any violent incidents (54.0% vs. 57.3%; p -value = 0.263). Covariate-adjusted regression models demonstrated fewer violent incidents in the intervention group than in the control group (2.0 vs. 2.3; p -value = 0.035). At the 18-month follow-up ($n = 1173$),¹⁰⁸ there were no significant differences in conviction for a violent crime (2.1% vs. 2.3%; p -value = 0.208), in the probability of reporting any violent incidents (48.7% vs. 44.5%; p -value = 0.157) or in number of violent incidents reported (0.9 vs. 0.8; p -value = 0.388).

Finally, post-intervention data (assessed near the end of the fourth year of the programme) comprising an analytic sample of 1069 students from the randomised trial of the QOP¹⁰³ showed that the probabilities of being involved in a gang fight in the past 12 months were not significantly different between intervention (16%) and control (14%) groups (p -value > 0.10).

TABLE 7 Measures used in studies reporting violence outcomes

Programme	Outcomes as reported
BBBS ⁹⁸	Number of times hit someone Number of times involved in a fight
NGYCP ⁷⁷	Any violent incidents in past 12 months Charged with a violent crime in past 12 months Convicted of a violent crime in past 12 months Number of violent incidents in past 12 months
QOP ⁷⁸	Involved in gang fight in past 12 months

Meta-analysis of effects on violence

Overview of included data

We included 10 effect sizes addressing violence outcomes from four reports of three studies.

Effects on violence

Positive Youth Development interventions did not have a statistically significant effect on violence outcomes across all time points ($d = 0.021$, 95% CI -0.050 to 0.093) (Table 8 and Figure 17). There was no meaningful programme-level heterogeneity in this finding ($I^2 = 0\%$), although this model drew on only 3 interventions and 10 effect sizes. Short-term outcomes did yield a statistically significant effect ($d = 0.076$, 95% CI 0.013 to 0.139), although this finding was marginally significant (p -value < 0.10) in sensitivity analysis and should thus be regarded with caution (Figure 18). Again, there was little meaningful programme-level heterogeneity ($I^2 = 0\%$).

Subgroup effects on violence

We aimed to examine how intervention effects were moderated by participants' sex, SES and ethnicity or by area deprivation in order to assess the potential impacts of PYD interventions on health inequalities. We did not undertake these analyses for reasons similar to those presented above in our description of subgroup effects on substance use. Regardless, the inclusion of only three studies would have made any such test grossly underpowered. Instead, we present narratively the findings of BBBS, the evaluation of which⁹⁸ was the only one to include subgroup analyses for violence outcomes by sex.

As described above, intervention youth hit someone else 0.85 times less frequently than did control group youth. This effect was larger in young women (1.17 times less; p -value < 0.10) than in young men (0.67 times less; p -value < 0.10), although subgroups were not significantly different from each other. In comparisons between intervention and control groups on the number of times youth were involved in a fight, young men were not different from young women (0.03 times less vs. 0.01 times less). Tests of moderation did not appear to be statistically significant.

No outcome evaluations presented analyses in terms of participants' SES, ethnicity or area deprivation. The evaluation of BBBS⁹⁸ presented subgroup analyses by categories defined by sex and ethnicity. Tests of moderation did not appear to be statistically significant.

TABLE 8 Violence outcome meta-analyses

Outcomes	Main analysis				Sensitivity analysis		
	ES (95% CI)	k	n	I^2 (%), programme level	I^2 (%), outcome level	Cochran's Q (df; p -value)	Low variance: ES (95% CI)
Violence, all time points	0.021 (-0.050 to 0.093)	3	10	0%	18%	12.27 (9; 0.20)	0.020 (-0.046 to 0.089)
Violence, short-term outcomes	0.076 (0.013 to 0.140)	3	7	0%	0%	4.94 (5; 0.55)	0.062 (-0.001 to 0.125)

df, degrees of freedom; ES, effect size.

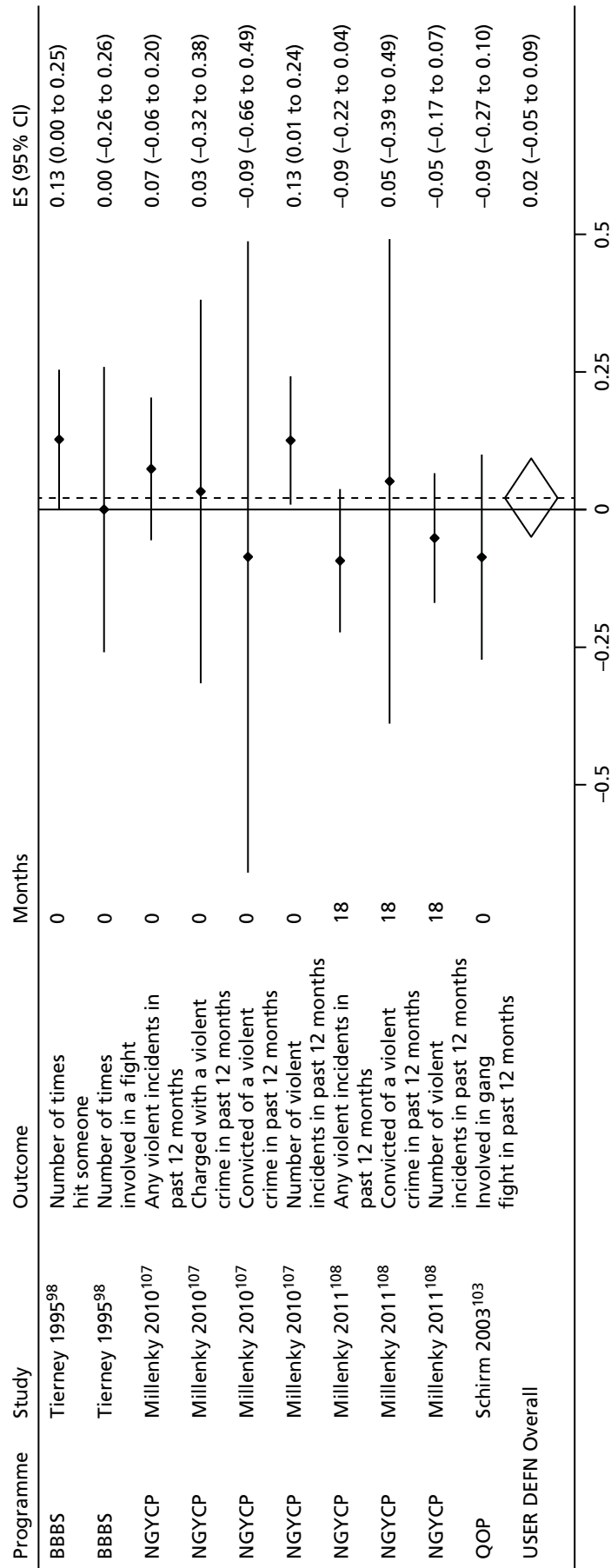


FIGURE 17 Violence outcomes. ES, effect size.

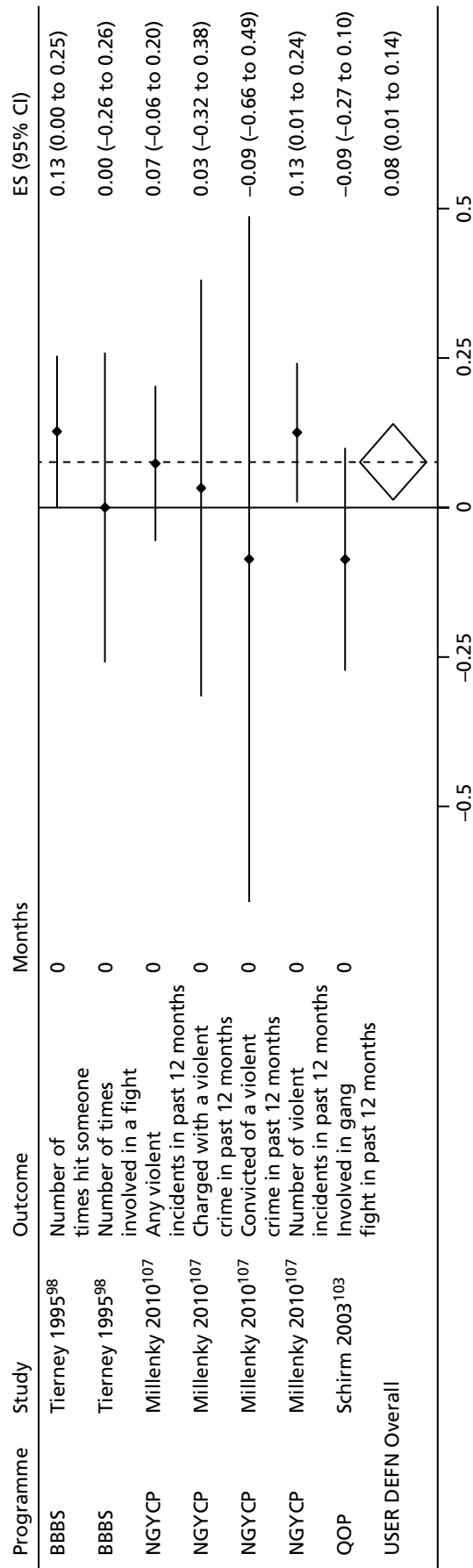


FIGURE 18 Short-term violence outcomes. ES, effect size.

Characteristics of participants and contexts appearing to moderate or to be necessary and sufficient for Positive Youth Development effectiveness

We aimed to examine what characteristics of participants and contexts appear to moderate, or are necessary and sufficient for, PYD effectiveness by undertaking metaregression or qualitative comparative analysis. However, the paucity and limitations in study designs, as well as the lack of statistical heterogeneity or qualitative differences in study effect sizes, precluded any such analyses.

Overall quality of evidence and Grading of Recommendations, Assessment, Development and Evaluations assessment

For both outcomes, evidence was rated 'very low' (Table 9).

Substance-use outcomes received ratings of 'very serious' for risk of bias owing to the inclusion of multiple non-randomised trials of variable quality, inappropriate analysis (e.g. ignoring clustering) and missing outcome data. Low levels of heterogeneity for substance-use analyses and 'direct' comparisons in this review led to ratings of 'not serious' for both inconsistency and indirectness. Wide, statistically non-significant CIs for most comparisons led to a rating of 'very serious' for imprecision. Finally, although publication bias could not be assessed directly, the combination of no statistically significant effect in most meta-analyses with the potential for residual confounding in non-randomised trials also downgraded the evidence.

TABLE 9 Summary of findings for PYD for substance use and violence

Positive Youth Development for substance use and violence outcomes							Summary of findings
Quality assessment						Overall quality of evidence	
Number of participants ^a (studies), follow-up	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias		
<i>Substance-use outcomes, all time points (follow-up: range 72 months)</i>							
7576 (9 RCTs), 72 months	Very serious ^b	Not serious	Not serious	Very serious ^c	All plausible residual confounding would suggest spurious effect, although no effect was observed	⊕○○○ Very low	All substance-use outcomes, all time points: ($d = 0.079$, 95% CI -0.025 to 0.183)
<i>Violence outcomes, all time points (follow-up: range 18 months)</i>							
3201 (3 RCTs), 18 months	Serious ^d	Not serious	Not serious	Very serious ^c	None	⊕○○○ Very low	All violence outcomes, all time points ($d = 0.021$, 95% CI -0.050 to 0.093)
<p>a Number of participants reflects the number for which data were reported to be analysed and is approximate.</p> <p>b Risk of bias was generated through inadequate analysis methods, incomplete outcome data and inclusion of non-randomised trials.</p> <p>c CIs were wide in almost all substance-use analyses.</p> <p>d Risk of bias was generated through inadequate analysis and incomplete outcome data.</p>							

Violence outcomes received ratings of 'serious' for risk of bias owing to the inclusion of only randomised trials in this comparison. As above, inconsistency and indirectness were rated as 'not serious'. Wide, statistically non-significant CIs for the overall analysis led to a 'very serious' rating for imprecision. Publication bias could not be detected, and there was little risk of residual confounding.

Feedback from consultation on the synthesis of outcome evaluations

The policy and practitioner advisory group and young people provided feedback specifically on the synthesis of outcomes on completion of the final draft, including the discussion and implications. Briefs reporting the results of our synthesis were produced for policy-makers and young people (see *Appendix 17*).

Policy stakeholders' feedback

Policy and practitioner stakeholders found the findings of the outcome evaluation, in context with the previous synthesis on theory and process, useful and comprehensive. They commented on the relevance of the findings in relation to their current programme developments in youth work in the UK. They welcomed the recommendations (see *Chapter 7, Implications for research*) on the importance of continued evaluations of PYD to reduce substance use and violence concurrent to its implementation.

Young people's feedback

Young people agreed that PYD programmes can have an effect on their drug use but not on their alcohol consumption unless programmes provide an alternative to drinking in terms of social venue. They were in both agreement and disagreement that PYD programmes can influence smoking or have a beneficial influence on violent behaviour.

Chapter 7 Discussion and conclusions

Review limitations

Deviations from protocol

Table 10 presents deviations both from the original funding proposal when finalising our review protocol at the start of the project, as well as deviations from this protocol in the course of completing the review.

Other limitations

Assessing the quality of theory reports

It is rare for systematic reviews of theory to assess their quality; this is exemplified by those reviews reported by Campbell *et al.*,¹⁴⁶ Harris *et al.*¹⁴⁷ and Kreiger.¹⁴⁸ Our quality criteria were informed by previous work by Michie *et al.*⁶⁴ and Bonell *et al.*,¹³¹ and were accompanied by guidance which was applied by two researchers who then met to discuss their scores. However, as shown in *Appendix 13*, reviewers found it difficult to make decisions in applying the criteria and there was very little inter-reviewer agreement on the scores.

There were a number of challenges. Few studies presented a single theory of change; many instead drew on a range of existing theories. This meant that reviewers were uncertain about whether or not they should assess each contributing theory (which would be difficult without references to primary texts) or to give an overall score for each report (which would be limited in terms of transparency). Reviewers also found it difficult to assess criteria of testability and parsimony because, although guidance was provided, this did not provide a clear benchmark against which to make such judgements, and reviewers felt that their judgements often shifted in the course of reviewing the literature. Application of the criterion of generalisability was, to some extent, easier to assess and more likely to result in agreement between reviewers simply because this was a more absolute judgement; most reports had been included because they offered an abstracted and therefore potentially generalisable framework for understanding PYD rather than, for example, reporting local data. However, we were unsure of the extent to which the theories would actually be relevant to varying structural or cultural contexts in practice. Finally, our criteria were intended to assess the quality of causal theories of change and thus were not useful in assessing the quality of the normative elements of theories.

TABLE 10 Deviations from proposal and protocol

From original proposal or published protocol	Deviation	Reason for deviation
Proposal	Scope: language – any language. Our original proposal did not exclude studies published in other languages; we added this exclusion criterion at an early stage in the review, including this in our registered protocol ⁵¹	This was decided because PYD interventions have been overwhelmingly developed in English-speaking countries
Proposal	Scope: date. Our original proposal did not restrict studies by date. We decided early on the review process to search for and include studies published from 1985 onwards. This exclusion criterion is reported in our registered protocol ⁵¹	PYD interventions were developed from 1985 onwards

continued

TABLE 10 Deviations from proposal and protocol (continued)

From original proposal or published protocol	Deviation	Reason for deviation
Proposal	Search: databases and websites. The list of databases included in the proposal was reviewed at the protocol ⁵¹ stage. At this early stage in the review, three databases were not included in the search (CAB Health, EMBASE, PAIS). The Health Technology Assessment Database was added and further efforts were placed in searching topic specific websites relevant to the intervention and health outcomes of interest. The revised list was published in our registered protocol ⁵¹	On the advice of our search specialist
Proposal	Search: journals. Our original proposal stated that we would hand-search the five journals that yield the highest numbers of studies that meet inclusion criteria. However, we amended this aspect of our search, including it in our registered protocol ⁵¹	We decided very early in the review that the original approach was not a good use of resources, because it would not be specific to reports missed by other approaches
Proposal	Pilots of screening at title and abstract. We piloted 100 rather than the original 50 references at title and abstract	This provided more opportunities to discuss potential variations in applying and to aid discuss of the exclusion criteria
Protocol	Scope: not in school time. We included one programme (YPPD) in which PYD was in practice delivered in a few sites in school hours, deviating from the intended model of delivery	This delivery was an unintended deviation from YPDP theory of change and occurred only in a minority of sites
Protocol	Synthesis of economic evaluations. We did not include any economic evaluations in our synthesis	Our searches yielded no relevant economic evaluations
Protocol	Meta-analysis: multivariate meta-analysis. As indicated in the protocol, ⁵¹ we intended to use multivariate meta-analysis or another method to synthesise effect size. Instead, we used multilevel meta-analysis with random effects at both the outcome and study level	It was not possible to use multivariate meta-analysis or another method to synthesise effect sizes because of the heterogeneity of outcomes and lack of availability of variance-covariance matrix for reported outcomes
Protocol	Synthesis: metaregression and qualitative comparative analysis. We stated in the protocol ⁵¹ that we would use a combination of metaregression and qualitative comparative analysis to test hypotheses generated from the theory and process synthesis, as well as funnel plots to examine potential publication bias	We were unable to conduct metaregression models to examine subgroup effects because of inconsistent subgroup reporting. We were unable to test hypotheses on other moderators of effects because of insufficient heterogeneity. We were unable to conduct qualitative comparative analysis because of insufficient qualitative variation in effectiveness to examine conditions predicting effectiveness. We were unable to conduct funnel plots because of insufficient studies per outcome
Protocol	Risk of bias. In addition to allocating a score of 'high risk', 'low risk' or 'unclear risk' within each critical appraisal domain, we also applied the code 'not applicable' to studies where codes were not suitable (e.g. methods of sequence generation and allocation concealment for controlled trials and whether or not studies controlled for key confounders in RCTs)	This allowed more transparent reporting of risk of bias
Protocol	Partner collaboration. We consulted with a slightly different array of policy stakeholders	This was due to people's availability for consultation

PAIS, Public Affairs Information Services.

Limitations in included theory reports

The theoretical literature did not in general focus on descriptions of causal theory of change for how PYD interventions might reduce substance use or violence among young people. Much of the literature instead aimed to assert the normative value of PYD as an approach to youth provision. This normative theory was, however, useful in understanding the goals and assumptions of PYD programming. Causal theorising was a minor and generally unsystematic element of most theory reports, with a few exceptions.

Although our synthesis of theories of change was hampered by a lack of clarity within included reports about how PYD might optimise young people's capacity for 'intentional self-regulation' and by a lack of systematic consideration of how promoting positive assets might lead to reductions in risk behaviours, we nonetheless developed a synthesis of causal theory which described the mechanisms by which PYD interventions might reduce violence and substance use. In the case of theorising how PYD interventions affect intentional self-regulation, it went beyond synthesis to fill in some gaps. The synthesis was successful despite the lack of success of our quality assessment of the theoretical literature. We included reports in our synthesis regardless of their quality. Our synthesis involved bringing together theoretical fragments (which specified only certain parts of the pathway from PYD intervention to substance use or violence reduction) and would have been less comprehensive had we synthesised only theories of change that themselves set out a comprehensive and clear path from intervention to risk reduction. Because of this, quality criteria focused on, for example, parsimony would not have proven very useful even if they had been easier to apply.

However, the resulting synthesis was quite 'thin', particularly in terms of how positive assets might enable reductions in risk behaviours. As well as not constituting a comprehensive theory of how assets reduce risk, the suggested pathways offer little that is distinctive to that provided in more traditional psychological theories used in prevention science such as the social learning model¹²⁸ and the social control theory,¹³⁵ both of which PYD theorists cite.

Limitations in process and outcome evaluations

Process evaluations overall were generally of low or medium quality. Sampling and analysis methods were poorly reported. Analyses were generally descriptive and did not develop clear, second-order interpretations. Few quotes were used to substantiate the analysis. Nonetheless, we were able to develop a synthesis that provided some useful answers to our RQ concerning the characteristics of contexts and participants that could influence the implementation and receipt of PYD interventions. Only one process evaluation from the UK was included, but this did include similar themes to those reporting from the USA and Australia.

The lack of studies from the UK was also apparent in the case of outcome evaluations, as was the lack of studies of cost-effectiveness. There were also various methodological problems with outcome evaluations. Authors rarely presented data in a format that was readily analysable, which meant that our analyses required both extensive transformation of effect sizes and sensitivity analyses. We decided to exclude one study both because of the quality of the evaluation and because of the uninterpretable effect sizes it reported. Although our rationale was transparent, it is possible that another meta-analyst may have taken a slightly different approach or made somewhat different transformation decisions. Moreover, we performed a sensitivity analysis in which we excluded one study⁷⁹ that did not report findings in the same standardised metric as other studies (i.e. as change from baseline rather than adjusted by baseline). The force of the conclusions did not change as a result of excluding this study.

The variable quality of evidence (including inadequate adjustment for clustering in several included studies) and the need for extensive data transformation and for sensitivity analysis for key statistical decisions must qualify the interpretation of our statistical results. Although, on balance, it was an appropriate decision to meta-analyse the included studies, challenges we faced with the data may suggest an interpretation of the pooled effect size that relies more on its general magnitude and precision than on statistical significance per se. We were also unable to test the key hypotheses derived from our consultations owing to the nature of the included evidence.

Moreover, although multilevel meta-analysis is perhaps a more robust method than those using one effect size per programme, it was not as robust as the multivariate meta-analysis originally proposed in the protocol. This is because multivariate meta-analysis uses the known variance–covariance matrix between included outcomes to account for dependencies between outcomes within interventions or studies, whereas multilevel meta-analysis achieves this by partitioning the variance between outcomes into that attributable to variation within interventions and variation across interventions.

Finally, there were problems associated with including evaluations of two interventions (PYDC⁷⁹ and Stay SMART¹¹¹) in the meta-analysis, for different reasons discussed above. Although we ultimately sensitivity-analysed findings with Tebes *et al.*,⁷⁹ our findings still may not present the most complete picture of effects on substance use.

Finally, we did not perform funnel plots because these would not have been a good guide to publication bias given the small number of studies included.¹⁴⁹ Our very comprehensive search methods, although not precluding the possibility of publication bias, mean that we took all reasonable steps to prevent this arising from an insensitive search.

Key results

Included studies

We identified 32,394 unique references from searches. Of these, 31,634 were excluded by screening on title and abstract. Of the 760 remaining references, we could obtain full reports of 689. Screening these led to a further 641 studies being excluded. The remaining 48 included reports arose from a total of 30 distinct studies (i.e. a distinct description of theory of change or empirical evaluation). Sixteen reports described theories of change, 12 reports (from 10 distinct studies) evaluated processes and 26 reports (from 10 distinct studies) evaluated outcomes. Five reports combined theories of change, process evaluation and/or outcome evaluation.

Taxonomy of Positive Youth Development

Positive Youth Development can be categorised according to:

- whether it aims to challenge or contribute to the existing social order
- whether it aims to promote the development of individuals or groups of individuals
- whether or not it aims to promote the development of individuals and/or modify the wider environments in which they develop
- the breadth, depth and duration of activities provided and
- whether or not it explicitly addresses the reduction of risk behaviours.

Theories of change for Positive Youth Development effects on substance use and violence

Sixteen reports were included. We aimed to assess the quality of these theories by drawing on criteria used previously but we found that these were challenging to apply consistently to the PYD theoretical literature.

The included literature did not provide sufficient information to develop a comprehensive theory of change for the effects of PYD interventions on substance use and violence. Nonetheless, by filling in some of the gaps in the literature (in a transparent manner), we succeeded in generating a theory of change.

Positive Youth Development interventions are intended to provide a positive environment for youth in terms of positive expectations, enduring and affective relationships with adults, diverse activities and settings and active participation, whereby young people are empowered to choose activities and to take on responsibilities. Interventions vary according to whether or not they aim to enable young people to contribute to or challenge the existing social order, focus on individual or collective development, aim to

transform individuals only or also the environments in which they live, provide breadth, depth and duration of activities, and address only positive assets or also risk behaviours.

In providing positive expectations and active participation in diverse activities, PYD aims to offer 'affordances', that is, resources individuals use in the course of their development (e.g. relationships, challenges, education; see Busseri *et al.*¹¹⁷). We interpret that young people can make use of these affordances in learning to apply 'intentional self-regulation' to specific intervention activities such as sports, arts or outdoor activities. Intentional self-regulation comprises 'intentionality' (assessing their current skills); 'selection' (setting goals for what they want to achieve); 'optimisation' (using their existing skills and the new affordances that PYD provides to achieve these goals); and 'compensation' (reviewing and if necessary redirecting actions to meet goals (see Busseri *et al.*¹¹⁷). PYD interventions aim to reward young people when they make progress with activities regarded as pro-social within that programme. Drawing on, but filling in gaps in, the PYD literature, we interpret that PYD interventions enable young people to engage in and learn from diverse, mutually reinforcing intervention activities, so that they develop a generalised intentional self-regulation focused on pro-social goals.

As a result of developing intentional self-regulation, young people are better able to develop various 'positive assets' such as the '5 Cs': competence, confidence, connection, character and caring (Lerner *et al.*¹²⁴). As these accrue, young people can make better use of the opportunities available in their wider environments, which leads to positive 'developmental regulations': a positive feedback cycle whereby the individuals gain more benefit from opportunities in their environment (Lerner *et al.*¹²⁴). This enables them in turn to make a positive contribution to their communities and societies, or, as a few authors would argue, enables them to contribute by critiquing and challenging inequities present in the existing social order.

These positive assets may then reduce risk behaviours via 'buffering' (see Catalano *et al.*,¹¹⁸ who state that risk factors in a young person's environment have less impact on those with positive assets) or 'compensation' (see Busseri *et al.*,¹¹⁷ who state that even if a young person engages in a risk behaviour, his or her possession of positive assets ameliorates the impacts of this on their overall health and development. (Note that this use of compensation is quite distinct from that within the description of intentional self-regulation above.) It is also argued that positive assets may reduce risk via 'molecular' mechanisms, whereby a specific asset exerts specific protection against a specific risk, or via 'pile-up', whereby accumulation of multiple assets is protective regardless of the specific assets involved. However, the theoretical literature synthesised here offers only limited insights beyond these general ideas. It suggests that engagement with pro-social peer groups or institutions might reduce antisocial behaviours via exposure and adherence to pro-social norms. It also suggests that improved emotional self-regulation, social skills and self-efficacy might contribute to better decision-making to avoid violence and substance use. But these theories fall short of a systematic theory of how the accrual of assets in particular or in general contributes to reductions in substance use or violence.

Characteristics of participants and contexts that affect implementation and receipt of PYD

Of the 10 included studies, eight were conducted in the USA, one in Australia and one in England. Study quality ranged from high reliability and usefulness^{77,89,112} to low reliability and usefulness.^{85,87,94}

A number of themes emerged from the synthesis. Community engagement was key to ensuring that programmes were culturally sensitive, accessible and appealing to young people and their parents and the wider community. Employing community members could be pivotal to successful implementation and the provision of role models. However, volunteers could be unreliable, for example in acting as mentors. Collaboration with other community agencies could be important, particularly in expanding the range of activities offered, but could also lead to a move away from original approaches.

Another theme was young people's relationships with providers and peers. Providers should relate to young people in a calm and nurturing yet authoritative way. Skilled providers could bridge social differences between participants, but this could be undermined by poor training or retention. Retention was challenging when programmes could not offer full-time positions. A final theme concerns challenges to ensuring that young people are empowered to make decisions about programme activities while also requiring them to engage in diverse activities, including vocational or academic activities. Although studies were drawn from the USA, the themes identified were found in US, UK and Australian reports and are likely to be pertinent across different geographical settings.

Effectiveness and cost-effectiveness in reducing substance use and violence

We found 13 study reports of 10 distinct outcome evaluations and included 12 study reports of nine distinct outcome evaluations in our meta-analyses. All but one study was conducted in the USA; one study was conducted in the UK. We could not categorise interventions as per the taxonomy derived from our theory synthesis because of the lack of detail reported. We found no economic evaluations and thus cannot assess cost-effectiveness. Four studies were RCTs, five were non-randomised trials with prospectively matched control groups and one included both randomised trial and non-randomised components which were analysed together. Overall, quality of evidence for our analyses of substance use and violence outcomes was rated 'very low'. Only one study described a theory of change for how its PYD-type intervention components might lead to reductions in substance use or violence.

Comparing PYD with usual or no treatment, our meta-analyses found a small effect for substance-use outcomes overall, which was significant neither statistically nor in terms of public health relevance. Meta-analyses found no effect for an omnibus measure of substance use across all time points but did find a small, statistically significant effect for this outcome in the short term (0–4 months post intervention). Meta-analyses of illicit drug-use and alcohol outcomes found no significant effects either across all time points or in the short term. Our meta-analysis for smoking at all time points included only four studies and so its finding of no significant effects should be treated with caution. A meta-analysis of short-term smoking was not undertaken, because only two studies could have been included.

Concerning the overall effects of PYD interventions on substance use and violence, two interpretations are possible. The first is that studies pointed to a small effect for substance-use outcomes that was significant neither in terms of statistics nor public health (i.e. that there is evidence of no effect). Another interpretation is that there is no evidence of the effectiveness of PYD interventions on substance use and violence outcomes. This interpretation is more conservative and better reflects the variable quality of included studies, the fact that some interventions, although meeting our inclusion criteria, may not have reflected a purist PYD approach as identified in our synthesis of theories of change, and the paucity of outcome evaluations included in the case of smoking and violence.

Our meta-analysis for smoking at all time points included only four studies and so its finding of no significant effects should be treated with caution. Meta-analysis of short-term smoking was not undertaken, because only two studies could have been included. Meta-analyses of illicit drug-use and alcohol outcomes found no significant effects either across all time points or in the short term. However, although our meta-analyses found no effect for an omnibus measure of substance use across all time points, we did find a small, statistically significant effect for this outcome in the short term (i.e. 0–4 months post intervention).

The data in *Appendix 15* are intended to indicate the degree to which the programmes evaluated in included studies actually embodied PYD principles but are limited in their usefulness because of limitations in study reporting. Nonetheless, it is clear that some of the studies included appear to be interventions that, despite meeting our inclusion criteria, may not be exemplars of the PYD approach as suggested by our synthesis of PYD theory of change. Some involved an array of sites with intervention activities varying between them, for example, the MAP,⁹³ the YPDP,^{46,112} the QOP⁷⁸ and BBBS.^{96,98} Furthermore, a number of programmes appeared to focus on explicit risk reduction as much as, if not more than, on positive

development. This was particularly the case with the 'All Stars' intervention⁹⁰⁻⁹² and to a lesser extent the PYDC,⁷⁹ Cool Girls, Inc.⁹⁵ and Stay SMART.¹¹¹ The NGYCP,⁷⁷ although focused on positive development rather than explicit risk prevention, could also be regarded as atypical of PYD interventions in adopting a quasi-military 'boot-camp' style of delivery.

It is of use to consider the two key deviant cases with positive effects that we included in our meta-analyses. Both the evaluation of the PYDC⁷⁹ and the evaluation of BBBS^{96,98} reported significant effects on reducing substance use, in the case of the first evaluation, at 4-month follow-up but not at post-intervention follow-up. The authors of the first evaluation conclude that these differences are attributable to the effectiveness of the programme, but this conclusion is attenuated by the non-randomised evaluation design. Moreover, although attrition was roughly balanced by arm, attrition was approximately 40% overall at the second follow-up. This is a serious limitation that was not addressed using methods for missing data. It is likely that the positive effects at second follow-up, especially when compared with the null effects from post-intervention follow-up, are optimistic. A limited description of the intervention as implemented makes it difficult to connect the apparent successes of this intervention with the programme model, although the authors do note that the intervention combined the promotion of positive assets with 'traditional' risk-based prevention education.

In contrast, the evaluation of BBBS⁹⁶ was randomised, although their use of complete case data analysis is an important caveat. This design, together with extensive programme description, provides some confidence about the significant intervention effects that the authors claim to demonstrate. The authors describe that what makes BBBS⁹⁸ different from other mentoring-based interventions is the focus on long-term relationships between a young person, often from a disadvantaged social or economic background, and an adult who may have 'aspirational' characteristics (e.g. higher education). The focus on long-term, carefully selected matches also set this mentoring intervention apart from the other interventions included in this review. But, above all, those randomised were drawn from a pool of families who had approached BBBS⁹⁸ for a match – that is to say, programme participants already came from families who were at least somewhat interested in participation. This is in contrast, for example, to the QOP⁷⁸ or All Stars,⁹⁰ where participation was 'opt-out' or otherwise brought to students and their families in the schools that young people were already attending. Together, all of these factors may have accounted for the success of BBBS.⁹⁸

Caution is also needed in determining the extent to which our results are transferable to settings outside the USA given that most studies were conducted there. The one UK study also reported null results but was non-random in design.^{46,112}

In terms of subgroup effects, it was not possible to undertake metaregressions looking at effects by sex. Narrative synthesis of three studies that examined subgroup effects found that these were mixed, with no clear pattern. We aimed to examine effects by SES, ethnicity or area deprivation but no studies reported these. One evaluation⁹⁸ presented a subgroup analyses by combinations of sex and ethnicity, reporting that minority ethnic status was not a moderator of effects.

Meta-analyses suggested that PYD interventions did not have a statistically significant effect on violence outcomes across all time points but there was a beneficial effect on short-term outcomes. However, this finding was only marginally significant in sensitivity analysis and should be treated with caution.

Metaregression to look at subgroup effects was not possible. Only one evaluation⁹⁸ reported effects by sex and found that this did not moderate effects. It also examined presented subgroup analyses by categories defined by sex and ethnicity, reporting that these did not moderate effects.

Characteristics of participants and contexts that determine effectiveness

We aimed to examine what characteristics of participants and contexts appear to moderate, or are necessary and sufficient for, PYD effectiveness. A synthesis of PYD theories of change and process evaluations suggested several hypotheses:

- Interventions that offer a breadth of activities may be more effective for younger adolescents, whereas those that emphasise depth may be more effective for older adolescents.
- Interventions that combine prevention and positive development may be less effective than those that focus only on positive development.
- Interventions of more than 1 year's duration may be more effective than those of shorter duration.
- Interventions may be more effective for participants with low or moderate levels of baseline risk, because there is more scope for stimulating 'intentional self-regulation.'¹²⁴
- Interventions that have specific methods to engage communities will be more effective.
- Projects that engage with schools will achieve better recruitment.
- Interventions that are delivered by well-trained staff will be more effective.
- Interventions that have better staff retention will be more effective.
- Interventions that offer some choices but require some engagement with educational components will be more effective.

However, the limited number of studies and very low level of statistical heterogeneity or qualitative differences in the effects reported by these studies precluded our innovative use of meta-regression or qualitative comparative analyses to test these hypotheses.

Conclusions

The way in which PYD interventions are theorised as having impacts on substance use and violence is currently inadequate. It is not sufficiently clear how interventions enable participants to develop 'intentional self-regulation'. How the development of positive assets is meant to enable young people to decrease their risk of engaging in substance use and violence has also not yet been comprehensively theorised.

A number of tensions can arise in how PYD is implemented in practice. Some programmes aim both to empower young people to choose which activities they engage in (and such voluntarism is often a hallmark of youth work in the UK) and to ensure that young people engage in a range of activities in order to develop multiple assets. Tensions can arise in practice in aiming to implement these two aims. Most PYD programmes aim to deploy skilled providers who can engage participants in long-term affective relationships and diverse activities, but this can be undercut by agencies not being able to offer full-time jobs because of the hours of work or by agencies lacking the resources to train and retain effective staff. Use of volunteers in programmes can help to engender community support and offer positive role models, but volunteers may also be hard to retain.

Our meta-analyses do not offer evidence that PYD interventions delivered outside school evaluated to date and compared with usual or no treatment in general have effects of significance to public health in reducing substance use and violence among young people. Our conclusions with regard to smoking and violence in particular should be treated with caution because of the paucity of included studies. We found no economic evaluations of PYD cost-effectiveness.

Noting that the studies included in our review were of interventions that were often implemented variably between evaluation sites or that included elements addressing risk reduction as much as, if not more prominently than, positive development suggests that these may not be exemplars of the PYD model. Therefore, our review should not be taken as evidence that the PYD model in principle cannot be effective in reducing substance use and violence among young people. Better theorisation of interventions and efforts to overcome problems with provider capacity may produce more effective interventions in future.

This lack of evidence for PYD effectiveness in reducing substance use or violence appears to contrast with reviews that have concluded that PYD is effective in promoting sexual health and preventing teenage pregnancy. However, it should be noted that these conclusions are not themselves based on very sound evidence. The Gavin *et al.*⁴⁴ review involved a narrative review focusing on a diversity of outcomes and reporting mixed results, and the Harden *et al.*⁴⁵ review involved a meta-analysis of effects on teenage pregnancy which drew on only two studies. But if it is, in fact, the case that PYD interventions are effective in reducing teenage pregnancy but not substance use or violence, this might plausibly be because there are stronger peer influences on substance use and violence than on teenage pregnancy,^{150–152} and because PYD interventions that target individuals rather than all-neighbourhood or all-school populations are less able to address such peer effects.

Implications

Implications for research

In terms of research on PYD interventions, more research is needed on the effects and cost-effectiveness of PYD interventions delivered outside school on violence and substance use. However, prior to any such studies, greater thought must be given to clarifying the theory of change of such interventions, particularly in terms of how PYD interventions facilitate the development of intentional self-regulation and how the accrual of positive assets reduces risk behaviours such as violence and substance use. More work is also needed to resolve tensions between enabling participants to choose which activities to participate in and ensuring that participants engage in a sufficient breadth of intervention activities, as well as to ensure that PYD providers have sufficient capacity for implementation. Subsequent evaluations should describe interventions more clearly both in terms of their theory of change and intervention characteristics, perhaps using a taxonomy such as that developed in this review and using terminology consistent with existing behaviour change technique classifications where possible.

In terms of review methods, our review innovatively applied methods of qualitative metasynthesis to theories of change. To facilitate such work in the future, more work is needed to produce practical criteria for assessing the quality of theoretical literature. Our own criteria, despite being informed by previous methodological and review literature, could not be applied in a consistent manner to assess included reports. Although criteria such as parsimony and generalisability appear to be sensible criteria by which to judge theory, how to apply them needs more thought. Further guidance is required for considering how reviews distinguish between theoretical papers that aim to present a single theory in detail versus those that refer, perhaps briefly, to one or more existing theories and what might be the benchmark in decision-making about concepts such as parsimony and generalisability that can otherwise appear subjective or relative.

Furthermore, there is a need to develop tools for reviewers to assess the quality not only of causal but also of normative aspects of theories. Two tests of normative theory suggested by Sayer *et al.*,¹³² namely desirability and feasibility, might represent a starting point for such developments. Criteria might assess the clarity of description and evidence for the problem being described; the assumptions that lie behind the alternatives being proposed and whether or not these are stated or implied; and whether or not the advocates of the alternatives acknowledge the disadvantages of their proposed solutions as well as promoting the advantages. However, we recognise that judging what might be desirable as well as feasible might be extremely challenging, as any judgement will be rooted in different moral or political perspectives.

We undertook a post hoc analysis of outcome evaluations to assess whether or not problems with the fidelity of delivery and the extent to which the interventions appeared to embody key principles of PYD suggested by our theory synthesis might explain the null effects. This was intended to assess whether or not our largely null findings concerning intervention effects were likely to reflect the ineffectiveness of the PYD approach or merely the limitations of the interventions evaluated to embody and implement this approach. We recommend that other systematic reviews should include such considerations as part of their planned protocols.

Implications for public health

Given the current lack of evidence for the effectiveness of PYD and of any economic evaluations of its cost-effectiveness, any investment in PYD delivered outside school as a strategy to reduce violence and substance-use outcomes should occur only within the context of evaluation studies. Our review suggests that existing PYD interventions that are subject to evaluation do not appear to have produced reductions in violence or substance use of public health significance. However, these interventions may not be the best examples of a PYD approach, as explained above. Therefore, our findings should not be taken as evidence for the ineffectiveness of PYD as a theory of change for reducing violence and substance use among young people. However, better evaluations are required before such interventions are considered for scale-up.

It may be that with better theorisation of how it can impact on risk behaviours, PYD could still be a promising approach. PYD does have two advantages over traditional prevention. First, as Catalano *et al.*³⁰ point out, whereas traditional prevention approaches often target only one or two risk factors, PYD, because of its holistic normative theory, is more likely to address multiple determinants of problem behaviours. Furthermore, as Kim *et al.*¹²¹ point out, PYD might be less prone to the inadvertent tendency within some traditional prevention interventions to set up negative expectations which condition young people's self-concepts, thereby inadvertently encouraging deviant behaviour.⁴⁸ However, if PYD interventions are to capitalise on these potential strengths to reduce risk behaviours, they must have clearer and more comprehensive theories of change that set out the mechanisms by which interventions promote positive assets which are evidenced as being protective factors against substance use or violence. These theories of change should explicate how PYD interventions are intended to promote intentional self-regulation and how particular assets then cause reductions in risk. This situation would certainly stand in contrast to the interventions that were included in our synthesis of outcome evaluations, which generally lacked clear theories of change and appeared to have little or no effects on substance use and violence.

Existing PYD interventions are currently also limited by their individualistic focus. Despite recognition in the theoretical literature reviewed here that PYD can involve interventions to compensate young people for the limitations in the developmental potential of other environments or interventions to transform these other environments, the former receives most attention. This might in part be an artefact of the inclusion criteria for this review, which for example excluded interventions delivered in normal school hours. However, the lack of evidence from our previous reviews of school-based interventions^{56,153} that PYD approaches are being used to modify school environments suggests that this is not merely an artefact. Although they generally lacked theories of change, the 'mainstream' interventions included in our review of outcome evaluations appeared to focus on addressing deficits in young people's positive assets to optimise their contribution to conventional society, whereas one example of a 'radical' approach to PYD⁸⁶ aimed to develop not only young people's individual assets but also their potential to be change agents within their local environments. Although they generally lacked theories of change, the 'mainstream' interventions included in our review of outcome evaluations appeared to be focused on addressing deficits in young people's positive assets to optimise their contribution to conventional society. One example of a 'radical' approach to PYD⁸⁶ aimed to develop young people's individual assets as well as their potential to be change agents within their local environments. None of the studies that we reviewed evaluated interventions that aimed to target the broader public or private sectors to render young people's environments less damaging to their physical and mental health.

Acknowledgements

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the National Institute for Health Research Public Health Research programme or the Department of Health.

This project is funded by a grant from the National Institute for Health Research Public Health Research programme (grant number PHR 12/153/19).

The work was undertaken with the support of DECIPHer, a UK Clinical Research Collaboration Public Health Research Centre of Excellence. Joint funding (MR/KO232331/1) from the British Heart Foundation, Cancer Research UK, Economic and Social Research Council, Medical Research Council, the Welsh Government and the Wellcome Trust, under the auspices of the UK Clinical Research Collaboration, is gratefully acknowledged.

We would also like to thank Hayley Reed, the Involving Young People Research Officer at DECIPHer, for her advice on this project, and all the young people who worked with us on this project during our consultations with the ALPHA young people's public input advisory group based in DECIPHer.

Contributions of authors

Chris Bonell planned the project and led on synthesis and report drafting.

Kelly Dickson managed the review day-to-day, led on consultation with policy stakeholders and contributed to searching, synthesising process evaluations and drafting the report.

Kate Hinds led on synthesising theory and contributed towards drafting the report.

GJ Melendez-Torres conducted the review and synthesis of outcome evaluations and contributed towards drafting the report.

Claire Stansfield provided guidance on and executed the search.

Adam Fletcher contributed to project planning, led the consultation with young people and commented on report drafts.

James Thomas contributed to project planning, advised on EPPI-Reviewer 4 and commented on report drafts.

Katrina Lester contributed to screening studies at title and abstract.

Elizabeth Oliver contributed to screening studies at title and abstract.

Simon Murphy contributed to project planning and commented on report drafts.

Rona Campbell contributed to project planning and commented on report drafts.

Policy/practitioner stakeholders

Eustace de Sousa (National Lead – Children, Young People and Families, Directorate for Health and Wellbeing Public Health England)

Geoff Dessent (Deputy Director for Health and Wellbeing, Directorate for Health and Wellbeing Public Health England)

Jessica Urwin (Information officer, National Youth Agency)

Ann Hagell (Research Lead, Association for Young People's Health)

Georgina Parry-Crooke (Professor of Social Research and Evaluation, London Metropolitan University)

Publications

Melendez-Torres GJ, Dickson K, Fletcher A, Thomas J, Hinds K, Campbell R, *et al.* A systematic review and meta-analysis of effects of positive youth development interventions on violence outcomes. *J Epidemiol Community Health* 2016; in press.

Bonell C, Hinds K, Dickson K, Thomas J, Fletcher A, Murphy S, *et al.* What is positive youth development and how might it reduce substance use and violence? A systematic review and synthesis of theoretical literature *BMC Public Health* 2016;**16**:135.

Melendez-Torres GJ, Dickson K, Fletcher A, Thomas J, Hinds K, Campbell R, *et al.* Positive youth development programmes to reduce substance use in young people: systematic review. *Int J Drug Policy* 2016; in press.

Bonell C, Hinds K, Dickson K, Thomas J, Fletcher A, Murphy S, *et al.* Systematic review and synthesis of theories of how Positive Youth Development interventions can reduce substance use and violence. *BMC Public Health* 2016; in press.

Data sharing statement

Because all data are already in the public realm, no separate data sharing plans have been made.

References

1. The United Nations Children's Emergency Fund. *An Overview of Child Wellbeing in Rich Countries: A Comprehensive Assessment of the Lives and Wellbeing of Children and Adolescents in the Economically Advanced Nations*. Florence: Innocenti Research Centre; 2007.
2. Hagel A, Coleman J, Brooks F. *Key Data on Adolescence 2013*. London: Association for Young People's Health; 2013.
3. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. *Lancet* 2012;**379**:1641–52. [http://dx.doi.org/10.1016/S0140-6736\(12\)60149-4](http://dx.doi.org/10.1016/S0140-6736(12)60149-4)
4. Sawyer SM, Afifi RA, Bearinger LH, Blakemore SJ, Dick B, Ezech AC. Adolescence: a foundation for future health. *Lancet* 2012;**379**:1630–40. [http://dx.doi.org/10.1016/S0140-6736\(12\)60072-5](http://dx.doi.org/10.1016/S0140-6736(12)60072-5)
5. Department for Work and Pensions. *Households Below Average Income: An Analysis of the Income Distribution 1994/5–2010/11*. London: Department for Work and Pensions; 2012.
6. Hibbel B, Anderson B, Bjarnsson T, Ahlstrom S, Balakirev O, Kokkevi A, et al. *The ESPAD report 2003. Alcohol and Other Drug Use Among Students in 35 Countries*. Stockholm: Swedish Council for Information on Alcohol and Other Drugs; 2004.
7. Fuller E. *Smoking, Drinking and Drug Use Among Young People in England in 2012*. London: Natcen; 2013.
8. Fuller E. *Smoking, Drinking and Drug Use Among Young People in England in 2011*. London: Natcen; 2012.
9. National Treatment Agency for Substance Misuse. *Drug Treatment Activity in England 2006/07*. London: Department of Health; 2007.
10. Volkow ND, Baler RD, Compton WM, Weiss SR. Adverse health effects of marijuana use. *N Engl J Med* 2014;**370**:2219–27. <http://dx.doi.org/10.1056/NEJMr1402309>
11. Galea S, Nandi A, Vlahov D. The social epidemiology of drug use. *Epidemiol Rev* 2004;**26**:36–52. <http://dx.doi.org/10.1093/epirev/mxh007>
12. Donaldson L. *Tackling the Health of the Teenage Nation: Chief Medical Officer's Annual Report 2007*. London: Department of Health; 2008.
13. Patton GC, Coffey C, Carlin JB, Degenhardt L, Lynskey M, Hall W. Cannabis use and mental health in young people: cohort study. *BMJ* 2002;**325**:1195–8. <http://dx.doi.org/10.1136/bmj.325.7374.1195>
14. World Health Organization. *Evidence-Based Strategies and Interventions to Reduce Alcohol-Related Harm: Global Assessment of Public-Health Problems Caused by Harmful Use of Alcohol*. Geneva: World Health Organization; 2007.
15. Department for Children, Schools and Families. *Safe to Learn: Embedding Antibullying Work in Schools*. London: Her Majesty's Stationery Office; 2008.
16. Department for Children, Schools and Families. *Your Child, Your Schools, Our Future: Building a 21st Century Schools System*. London: Her Majesty's Stationery Office; 2009.
17. Department of Health. *Healthy Lives, Brighter Futures*. London: Her Majesty's Stationery Office; 2009.

18. Beinart S, Anderson B, Lee S, Utting D. *Youth at Risk? A National Survey of Risk Factors, Protective Factors and Problem Behaviour among Young People in England, Scotland and Wales (JRF Findings 432)*. York: Joseph Rowntree Foundation; 2002.
19. Jansen DE, Veenstra R, Ormel J, Verhulst FC, Reijneveld SA. Early risk factors for being a bully, victim, or bully/victim in late elementary and early secondary education: the longitudinal TRAILS study. *BMC Public Health* 2011;**11**:440. <http://dx.doi.org/10.1186/1471-2458-11-440>
20. Arseneault L, Walsh E, Trzesniewski K, Newcombe R, Caspi A, Moffitt T. Bullying victimization uniquely contributes to adjustment problems in young children: a nationally representative cohort study. *Pediatrics* 2006;**118**:130–8. <http://dx.doi.org/10.1542/peds.2005-2388>
21. Forero R, McLellan L, Rissel C, Bauman A. Bullying behaviour and psychosocial health among school students in New South Wales, Australia: cross sectional survey. *BMJ* 1999;**319**:344–8. <http://dx.doi.org/10.1136/bmj.319.7206.344>
22. Kaltiala-Heino R, Rimpelä M, Rantanen P, Rimpelä A. Bullying at school—an indicator of adolescents at risk for mental disorders. *J Adolesc* 2000;**23**:661–74. <http://dx.doi.org/10.1006/jado.2000.0351>
23. Juvonen J, Graham S, Schuster MA. Bullying among young adolescents: the strong, the weak, and the troubled. *Pediatrics* 2003;**112**:1231–7. <http://dx.doi.org/10.1542/peds.112.6.1231>
24. Bond L, Carlin JB, Thomas L, Rubin K, Patton G. Does bullying cause emotional problems? A retrospective study of young teenagers. *BMJ* 2001;**323**:480–4. <http://dx.doi.org/10.1136/bmj.323.7311.480>
25. Hawker DS, Boulton MJ. Twenty years' research on peer victimization and psychosocial maladjustment: a meta-analytic review of cross-sectional studies. *J Child Psychol Psychiatry* 2000;**41**:441–55. <http://dx.doi.org/10.1111/1469-7610.00629>
26. Van der Wal MF, De Wit CA. Psychosocial health among young victims and offenders of direct and indirect bullying. *Pediatrics* 2003;**111**:1312–7. <http://dx.doi.org/10.1542/peds.111.6.1312>
27. Dmitrieva J, Gibson L, Steinberg L, Piquero A, Fagan J. Predictors and consequences of gang membership: comparing gang members, gang leaders, and non-gang-affiliated adjudicated youth. *J Res Adolesc* 2014;**24**:1–15. <http://dx.doi.org/10.1111/jora.12111>
28. Sainsbury Centre for Mental Health. *Diversion: a Better Way for Criminal Justice and Mental Health*. London: Sainsbury Centre for Mental Health; 2009.
29. Scott S, Knapp M, Henderson J, Maughan B. Financial cost of social exclusion: follow up study of antisocial children into adulthood. *BMJ* 2001;**323**:191. <http://dx.doi.org/10.1136/bmj.323.7306.191>
30. Catalano RF, Berglund LM, Ryan JAM, Lonczak HS, Hawkins JD. Positive youth development in the United States: research findings on evaluations of positive youth development programs. *Prev Treatment* 2002;**5**:1–166. <http://dx.doi.org/10.1037/1522-3736.5.1.515a>
31. Jackson C, Sweeting H, Haw S. Clustering of substance use and sexual risk behaviour in adolescence: analysis of two cohort studies. *BMJ Open* 2012;**2**:e000661. <http://dx.doi.org/10.1136/bmjopen-2011-000661>
32. Buck D. *Clustering of Unhealthy Behaviours Over Time*. London: The King's Fund; 2012.
33. Kipping RR, Campbell RM, MacArthur GJ, Gunnell DJ, Hickman M. Multiple risk behaviour in adolescence. *J Pub Health* 2012;**34**:i1–i2. <http://dx.doi.org/10.1093/pubmed/fdr122>
34. National Youth Agency. *The NYA Guide to Youth Work in England*. Leicester: National Youth Agency; 2007.

35. Schulman S, Davies T. *Evidence of the Impact of the 'Youth Development Model' on Outcomes for Young People – A Literature Review*. Leicester: National Youth Agency; 2007.
36. Roth JL, Brooks-Gunn J. Promoting healthy adolescents: synthesis of youth development program evaluations. *J Res Adolesc* 1998;**8**:423–59. http://dx.doi.org/10.1207/s15327795jra0804_2
37. Philliber S, Kaye JW, Herrling S, West E. Preventing pregnancy and improving health care access among teenagers: an evaluation of the Children's Aid Society-Carrera program. *Perspect Sex Reprod Health* 2002;**34**:244–51. <http://dx.doi.org/10.2307/3097823>
38. Department for Education. *Positive for Youth – A New Approach to Cross-Government Policy for Young People Aged 13 to 19*. London: Her Majesty's Stationery Office; 2011.
39. Department of Health. *Healthy Lives, Healthy People. Our Strategy for Public Health in England*. London: Her Majesty's Government; 2010.
40. Mayor's Fund For London. *Mayor's Fund for London: Annual Report 2011*. London: Mayor's Fund For London; 2011.
41. Scottish Government. *Valuing Young People: Principles and Connections to Support Young People Achieve their Potential*. Edinburgh: Scottish Government; 2009.
42. Welsh Assembly Government. *The Youth Work Curriculum Statement for Wales*. Cardiff: Welsh Assembly Government; 2007.
43. Durlak JA, Weissberg RP. *The Impact of After-School Programs that Promote Personal and Social Skills*. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning; 2007.
44. Gavin LE, Catalano RF, David-Ferdon C, Gloppen KM, Markham CM. A review of positive youth development programs that promote adolescent sexual and reproductive health. *J Adolesc Health* 2010;**46**:S75–91. <http://dx.doi.org/10.1016/j.jadohealth.2009.11.215>
45. Harden A, Brunton G, Fletcher A, Oakley A, Burchett H, Backhans M. *Young people, Pregnancy and Social Exclusion: A Systematic Synthesis of Research Evidence to Identify Effective, Appropriate and Promising Approaches for Prevention and Support*. London: Institute of Education; 2006.
46. Wiggins M, Bonell C, Sawtell M, Austerberry H, Burchett H, Allen E, et al. Health outcomes of youth development programme in England: prospective matched comparison study. *BMJ* 2009;**339**. <http://dx.doi.org/10.1136/bmj.b2534>
47. Bonell C, Maisey R, Speight S, Purdon S, Keogh P, Wollny I, et al. Randomized controlled trial of 'Teens and Toddlers': a teenage pregnancy prevention intervention combining youth development and voluntary service in a nursery. *J Adolesc* 2013;**36**:859–70. <http://dx.doi.org/10.1016/j.adolescence.2013.07.005>
48. Dishion TJ, McCord J, Poulin F. When interventions harm. *Am Psychol* 1999;**54**:755–64. <http://dx.doi.org/10.1037/0003-066X.54.9.755>
49. Weiss B, Caron A, Ball S, Tapp J, Johnson M, Weisz JR. Iatrogenic effects of group treatment for anti-social youth. *J Consult Clin Psychol* 2005;**73**:1036–44. <http://dx.doi.org/10.1037/0022-006X.73.6.1036>
50. Roth JL, Brooks-Gunn J. Youth development programs: risk, prevention and policy. *J Adolesc Health* 2003;**32**:170–82. [http://dx.doi.org/10.1016/S1054-139X\(02\)00421-4](http://dx.doi.org/10.1016/S1054-139X(02)00421-4)
51. Bonell C, Thomas J, Campbell R, Murphy S, Fletcher A. *The Effects of Positive Youth Development Interventions on Substance Use, Violence and Inequalities: Systematic Review of Theories of Change, Processes and Outcomes*. URL: www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013005439 (accessed 22 August 2013).

52. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta analyses: the PRISMA statement. *PLOS Med* 2009;**6**:e1000097. <http://dx.doi.org/10.1371/journal.pmed.1000097>
53. World Health Organization. *Health for the World's Adolescents*. Geneva: World Health Organization; 2014.
54. Hummel S, Naylor P, Chilcott J, Guillaume L, Wilkinson A, Blank L. *Cost Effectiveness of Universal Interventions Which Aim to Promote Emotional and Social Wellbeing in Secondary Schools: Report for NICE*. Sheffield: University of Sheffield; 2009.
55. Durlak JA, Weissberg RP, Dymnicki AB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Dev* 2011;**82**:405–32. <http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x>
56. Bonell C, Jamal F, Harden A, Wells H, Parry W, Fletcher A, et al. Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. *Public Health Res* 2013;**1**.
57. Faggiano F, Vigna-Taglianti FD, Versino E, Zambon A, Borraccino A, Lemma P. School-based prevention for illicit drugs use: a systematic review. *Prev Med* 2008;**46**:385–96. <http://dx.doi.org/10.1016/j.ypmed.2007.11.012>
58. Foxcroft DR, Ireland D, Lowe G, Breen R. Primary prevention for alcohol misuse in young people. *Cochrane Database Syst Rev* 2002;**2**:CD003020. <http://dx.doi.org/10.1002/14651858.cd003024>
59. Hahn R, Fuqua-Whitley D, Wethington H, Lowy J, Crosby A, Fullilove M. Effectiveness of universal school-based programs to prevent violent and aggressive behavior: a systematic review. *Am J Prev Med* 2007;**33**:S114–S29. <http://dx.doi.org/10.1016/j.amepre.2007.04.012>
60. Thomas R, Perera R. School-based programmes for preventing smoking. *Cochrane Database Syst Rev* 2006;**3**:CD001293. <http://dx.doi.org/10.1002/14651858.cd001293.pub2>
61. Thomas J, Brunton J, Graziosi S. *EPPI-Reviewer 4.0: Software for Research Synthesis*. EPPI-Centre Software. London: Social Science Research Unit, Institute of Education; 2010.
62. Brunton G, Stansfield C, Thomas J. Finding Relevant Studies. In Gough D, Oliver S, Thomas J, editors. *An Introduction to Systematic Reviews*. London: Sage; 2012. pp. 107–34.
63. Peersman G, Oliver S, Oakley A. *EPPI-Center Review Guidelines: Data Collection for the EPIC Database*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London; 1997.
64. Michie S, West R, Campbell R, Brown J, Gainforth H. *ABC of Behaviour Change Theories: An Essential Resource for Researchers, Policy Makers and Practitioners*. Surrey: Silverback Publishing Ltd; 2014.
65. Davis R, Campbell R, Hildon Z, Hobbs L, Michie S. Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychol Rev* 2014;**8**:1–22.
66. Shepherd J, Harden A, Rees R, Brunton G, Garcia J, Oliver S, et al. *Young People and Healthy Eating: A Systematic Review of Barriers and Facilitators*. London: EPPI-Centre, Social Science Research Unit; 2001.
67. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. Oxford: The Cochrane Collaboration; 2011. URL: www.cochrane-handbook.org (accessed 22 August 2013).
68. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008;**336**:1106–10. <http://dx.doi.org/10.1136/bmj.39500.677199.AE>

69. Au W. High-stakes testing and curricular control: a qualitative metasynthesis. *Educ Res* 2007;**36**:258–67. <http://dx.doi.org/10.3102/0013189X07306523>
70. Crabtree BF, Miller WL. Using Codes and Code Manuals: A Template Organizing Style of Interpretation. In Crabtree BF, Miller WL, editors. *Doing Qualitative Research*. 2nd edn. Thousand Oaks, CA: Sage; 1999. pp. 163–78.
71. King N. Template Analysis. In Symon G, Cassell C, editors. *Qualitative Methods and Analysis in Organizational Research: A Practical Guide*. London: Sage; 1998. pp. 118–34.
72. King N. *What is Template Analysis?* Huddersfield: University of Huddersfield School of Human and Health Sciences; 2006. URL: www.hud.ac.uk/hhs/research/template_analysis/whatis.htm (accessed 27 April 2007).
73. Arai L, Roen K, Roberts H, Popay J. It might work in Oklahoma but will it work in Oakhampton? Context and implementation in the effectiveness literature on domestic smoke detectors. *Inj Prev* 2005;**11**:148–51. <http://dx.doi.org/10.1136/ip.2004.007336>
74. Noyes J, Popay J, Garner P. *What Can Qualitative Research Contribute to a Cochrane Systematic Review of DOT for Promoting Adherence to Tuberculosis Treatment?* Paper presented at the Qualitative Research and Systematic Reviews workshop, Continuing Professional Development Centre, Oxford, June 2005.
75. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol* 2008;**8**:45. <http://dx.doi.org/10.1186/1471-2288-8-45>
76. Schutz A. *Collected Papers: Vol. 1*. Leiden: The Hague, Martinus Nijhoff; 1962.
77. Schwartz SEO, Rhodes JE, Spencer R, Grossman JB. Youth initiated mentoring: investigating a new approach to working with vulnerable adolescents. *Am J Community Psychol* 2013;**52**:155–69. <http://dx.doi.org/10.1007/s10464-013-9585-3>
78. Rodriguez-Planas N. *Longer-term Impacts of Mentoring, Educational Services, and Incentives to Learn: Evidence from a Randomized Trial in the United States*. Working Paper 449. Barcelona Economics Working Paper Series. Bonn: Institute for the Study of Labor; 2010.
79. Tebes JK, Feinn R, Vanderploeg JJ, Chinman MJ, Shepard J, Brabham T, et al. Impact of a positive youth development program in urban after-school settings on the prevention of adolescent substance use. *J Adolesc Health* 2007;**41**:239–47. <http://dx.doi.org/10.1016/j.jadohealth.2007.02.016>
80. Morris SB, DeShon RP. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods* 2002;**7**:105–25. <http://dx.doi.org/10.1037/1082-989X.7.1.105>
81. Cheung MW. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. *Psychol Methods* 2014;**19**:211–29. <http://dx.doi.org/10.1037/a0032968>
82. Social Research Association. *Ethical Guidelines*. London: Social Research Association; 2003.
83. Centre for Social Justice and Community Action and National Co-ordinating Centre for Public Engagement. *Community-Based Participatory Research: A Guide to Ethical Principles and Practice*. Bristol: Centre for Social Justice and Community Action and National Coordinating Centre for Public Engagement; 2012.
84. Armstrong T, Armstrong G. The organizational, community and programmatic characteristics that predict the effective implementation of after-school programs. *J Sch Violence* 2004;**3**:93–109. http://dx.doi.org/10.1300/J202v03n04_07

85. Baker K, Pollack M, Kohn I. Violence prevention through informal socialization: an evaluation of the South Baltimore youth center. *Stud Crime Crime Prev* 1995;**4**:61–85.
86. Berg M, Coman E, Schensul J. Youth action research for prevention: a multi-level intervention designed to increase efficacy and empowerment among urban youth. *Am J Community Psychol* 2009;**43**:345–59. <http://dx.doi.org/10.1007/s10464-009-9231-2>
87. Bloomberg L, Ganey A, Alba V, Quintero G, Alcantara LA. Chicano-Latino youth leadership institute: an asset-based program for youth. *Am J Health Behav* 2003;**27**:S45–S54. <http://dx.doi.org/10.5993/AJHB.27.1.s1.5>
88. Bulanda JJ, McCrea KT. The promise of an accumulation of care: disadvantaged African-American youths' perspectives about what makes an after school program meaningful. *Child Adolesc Social Work J* 2013;**30**:95–118. <http://dx.doi.org/10.1007/s10560-012-0281-1>
89. Cross AB, Gottfredson DC, Wilson DM, Rorie M, Connell N. Implementation quality and positive experiences in after-school programs. *Am J Community Psychol* 2010;**45**:370–80. <http://dx.doi.org/10.1007/s10464-010-9295-z>
90. Cross AB, Gottfredson DC, Wilson DM, Rorie M, Connell N. The impact of after-school programs on the routine activities of middle-school students: results from a randomized, controlled trial. *Criminol Public Policy* 2009;**8**:391–412. <http://dx.doi.org/10.1111/j.1745-9133.2009.00555.x>
91. Gottfredson DC, Cross A, Wilson D, Rorie M, Connell N. *A Randomized Trial of the Effects of an Enhanced After-School Program for Middle-School Students*. University of Maryland: Baltimore, MD; 2010.
92. Gottfredson DC, Cross AB, Wilson D, Rorie M, Connell N. An experimental evaluation of the All Stars prevention curriculum in a community after school setting. *Prev Sci* 2010;**11**:142–54. <http://dx.doi.org/10.1007/s11121-009-0156-7>
93. Gottfredson DC, Gerstenblith SA, Soulé DA, SC Womer, Lu S. Do after school programs reduce delinquency? *Prev Sci* 2004;**5**:253–66. <http://dx.doi.org/10.1023/B:PREV.0000045359.41696.02>
94. Lee KS, Conigrave KM, Clough AR, Wallace C, Silins E, Rawles J. Evaluation of a community-driven preventive youth initiative in Arnhem Land, Northern Territory, Australia. *Drug Alcohol Rev* 2008;**27**:75–82. <http://dx.doi.org/10.1080/09595230701711124>
95. Kuperminc GP, Thomason J, DiMeo M, Broomfield-Massey K. Cool Girls, Inc.: promoting the positive development of urban preadolescent and early adolescent girls. *J Prim Prev* 2011;**32**:171–83. <http://dx.doi.org/10.1007/s10935-011-0243-y>
96. Rhodes JE, Reddy R, Grossman JB. The protective influence of mentoring on adolescents' substance use: direct and indirect pathways. *Appl Dev Sci* 2005;**9**:31–47. http://dx.doi.org/10.1207/s1532480xads0901_4
97. Grossman JB, Tierney JP. Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation Rev* 1998;**22**:403–26. <http://dx.doi.org/10.1177/0193841X9802200304>
98. Tierney JP. *Making a Difference. An Impact Study of Big Brothers/Big Sisters*. Philadelphia, PA: Public/Private Ventures; 1995.
99. Maxfield M, Castner L, Maralani V, Vencill M. *The Quantum Opportunities Program Demonstration: Implementation Findings*. Washington, DC: Mathematica Policy Research Inc.; 2003.
100. Maxfield M, Schirm A, Rodriguez-Planas N. *The Quantum Opportunity Program Demonstration: Implementation and Short-Term Impacts*. Washington, DC: Mathematica Policy Research; 2003.

101. Rodriguez-Planas N. *Mentoring, Educational Services, and Economic Incentives Longer-term Evidence on Risky Behaviors from a Randomized Trial*. Universitat Autònoma de Barcelona, FEDEA and IZA Discussion Paper No. 4968; May 2010.
102. Rodriguez-Planas N. Longer-term impacts of mentoring, educational services, and learning incentives: evidence from a randomized trial in the United States. *Am Econ J* 2012;**4**:121–39. <http://dx.doi.org/10.1257/app.4.4.121>
103. Schirm A, Rodriguez-Planas N, Maxfield M, Tuttle C. *The Quantum Opportunity Program Demonstration: Short-term impacts*. Washington, DC: Mathematica Policy Research; 2003.
104. Schirm A, Rodriguez-Planas N. *The Quantum Opportunity Program demonstration: Initial post-intervention impacts*. Washington, DC: Mathematica Policy Research; 2004.
105. Schirm A, Stuart E, McKie A. *The Quantum Opportunity Program Demonstration: Final Impacts*. Washington, DC: Mathematica Policy Research; 2006.
106. Bloom D, Gardenhire-Crooks A, Mandsager C. *Reengaging High School Dropouts: Early Results of the National Guard Youth ChalleNGe Program Evaluation. Full Report*. New York, NY: MDRC; 2009.
107. Millenky M, Bloom D, Dillon C. *Making the Transition: Interim Results of the National Guard Youth ChalleNGe Evaluation. [Executive Summary]*. New York, NY: MDRC; 2010.
108. Millenky M, Bloom D, Muller-Ravett S, Broadus J. *Staying on Course: Three-Year Results of the National Guard Youth ChalleNGe Evaluation. Executive Summary*. New York, NY: MDRC; 2011.
109. Millenky M, Schwartz SEO, Rhodes JE. Supporting the transition to adulthood among high school dropouts: an impact study of the national guard youth challenge program. *Prev Sci* 2014;**15**:448–59. <http://dx.doi.org/10.1007/s11121-013-0388-4>
110. Perez-Arce F, Constant L, Loughran DS, Karoly LA. *A Cost-Benefit Analysis of the National Guard Youth ChalleNGe Program. Technical Report*. Santa Monica, CA: RAND Corporation; 2012.
111. St Pierre TL, Kaltreider DL. Drug prevention in a community setting: a longitudinal study of the relative effectiveness of a three-year primary prevention program in boys and girls clubs across the nation. *Am J Community Psychol* 1992;**20**:673–706. <http://dx.doi.org/10.1007/BF01312603>
112. Wiggins M, Bonell CP, Burchett H, Sawtell M, Austerberry H, Allen E, et al. *Young People's Development Programme Evaluation: Final Report*. London: University of London, Institute of Education, Social Science Research Unit; 2008.
113. Benson PL. Developmental Assets: An Overview of Theory, Research and Practice. In Silbereisen RK, Lerner RM, editors. *Approaches to Positive Youth Development*. Thousand Oaks, CA: Sage; 2007. pp. 33–58. <http://dx.doi.org/10.4135/9781446213803.n2>
114. Benson PL, Mannes M, Pittman K, Ferber T. Youth Development, Developmental Assets, and Public Policy. In Lerner RM, Steinberg L, editors. *Handbook of Adolescent Psychology*. 2nd edn. Hoboken, NJ: Wiley; 2004. pp. 781–814. <http://dx.doi.org/10.1002/9780471726746.ch25>
115. Benson PL, Scales PC. Positive youth development and the prevention of youth aggression and violence. *Eur J Dev Sci* 2009;**3**:218–34.
116. Benson PL, Scales PC, Syvertsen AK. The contribution of the developmental assets framework to positive youth development theory and practice. *Adv Child Dev Behav* 2011;**41**:197–230. <http://dx.doi.org/10.1016/B978-0-12-386492-5.00008-7>
117. Busseri MA, Rose-Krasnor L. Breadth and intensity: salient, separable, and developmentally significant dimensions of structured youth activity involvement. *Br J Dev Psychol* 2009;**27**:907–33. <http://dx.doi.org/10.1348/026151008X397017>

118. Catalano RF, Hawkins JD, Berglund ML, Pollard JA, Arthur MW. Prevention science and positive youth development: competitive or cooperative frameworks? *J Adolesc Health* 2002;**31**:230–9. [http://dx.doi.org/10.1016/S1054-139X\(02\)00496-2](http://dx.doi.org/10.1016/S1054-139X(02)00496-2)
119. Ginwright S, Cammarota J. New terrain in youth development: the promise of a social justice approach. *Soc Justice* 2002;**29**:82–95.
120. Kia-Keating M, Dowdy E, Morgan ML, Noam G. Protecting and promoting: an integrative conceptual model for healthy development of adolescents. *J Adolesc Health* 2011;**48**:220–8. <http://dx.doi.org/10.1016/j.jadohealth.2010.08.006>
121. Kim S, Crutchfield C, Williams C, Hepler N. Toward a new paradigm in substance abuse and other problem behavior prevention for youth: youth development and empowerment approach. *J Drug Educ* 1998;**28**:1–17. <http://dx.doi.org/10.2190/5ET9-X1C2-Q17B-2G6D>
122. Lee TY. Construction of an integrated positive youth development conceptual framework for the prevention of the use of psychotropic drugs among adolescents. *Sci World J* 2011;**11**:2403–17. <http://dx.doi.org/10.1100/2011/315870>
123. Lerner RM, Lerner JV. Toward a New Vision and Vocabulary About Adolescence: Theoretical, Empirical, and Applied Bases of a 'Positive Youth Development' Perspective. In Balter L, Tamis-LeMonda CS, editors. *Child Psychology: A Handbook of Contemporary Issues*. 2nd edn. New York, NY: Psychology Press; 2006. pp. 445–69.
124. Lerner RM, Lerner JV, von Eye A, Bowers EP, Lewin-Bizan S. Individual and contextual bases of thriving in adolescence: a view of the issues. *J Adolesc* 2011;**34**:1107–14. <http://dx.doi.org/10.1016/j.adolescence.2011.08.001>
125. Perkins DF, Borden LM, Villarruel FA. Community youth development: a partnership for action. *Sch Comm J* 2001;**11**:39–56.
126. Schwartz SJ, Pantin H, Coatsworth JD, Szapocznik J. Addressing the challenges and opportunities for today's youth: toward an integrative model and its implications for research and intervention. *J Prim Prev* 2007;**28**:117–44. <http://dx.doi.org/10.1007/s10935-007-0084-x>
127. Bronfenbrenner U, Morris PA. The Bioecological Model of Human Development. In Lerner RM, Damon W, editors. *Handbook of Child Psychology. Vol. 1: Theoretical Models of Human Development*. 6th edn. Hoboken, NJ: John Wiley & Sons; 2006.
128. Bandura A. *Social Learning Theory*. Morristown, NJ: General Learning Press; 1977.
129. Erikson EH. *Identity: Youth and Crisis*. New York, NY: Norton; 1968.
130. Marcia JE. Ego identity status: relationship to change in self-esteem, general maladjustment, and authoritarianism. *J Pers* 1967;**35**:118–33. <http://dx.doi.org/10.1111/j.1467-6494.1967.tb01419.x>
131. Bonell C, Wells H, Harden A, Jamal F, Fletcher A, Thomas J, et al. Theories of how the school environment impacts on student health: systematic review. *Health Place* 2013;**24**:242–9. <http://dx.doi.org/10.1016/j.healthplace.2013.09.014>
132. Sayer A. Moral economy and political economy. *Stud Polit Econ* 2000;**61**:79–103.
133. Chen H. *Theory-Driven Evaluations*. London: Sage; 1990.
134. Roth JL, Brooks-Gunn J. What is a Youth Development Program? Identification and defining principles. In Lerner RM, Jacobs F, Wertlieb D, editors. *Handbook of Applied Developmental Science, Vol. 2: Promoting Positive Child, Adolescent, and Family Development Through Research, Policies, and Programs*. Thousand Oaks, CA: Sage; 2000. pp. 197–223.
135. Hirschi T. *Causes of Delinquency*. Berkeley, CA: University of California Press; 1969.
136. Freire P. *Education for Critical Consciousness*. New York, NY: Continuum; 1998.

137. Gramsci A. *Selections from the Prison Notebooks*. New York, NY: International Publishers; 1971.
138. Bourdieu P, Passeron JC. *Reproduction in Education, Society and Culture*. London: Sage; 1977.
139. Foucault M. Technologies of the Self. In Martin LH, Gutman H, Hutton PH, editors. *Technologies of the Self: A Seminar with Michel Foucault*. London: Tavistock; 1982.
140. Freire P. *Pedagogy of the Oppressed*. Middlesex: Penguin; 1972.
141. Freire P. Creating Alternative Research Methods: Learning to Do It By Doing It. In Hall BL, Gillette A, Tandon R, editors. *Creating Knowledge: A Monopoly?* New Delhi: Society for Participatory Research in Asia; 1982.
142. Freire P. *The Politics of Education: Culture, Power, and Liberation*. South Hadley, MA: Bergin and Garvey; 1985. <http://dx.doi.org/10.1007/978-1-349-17771-4>
143. Weiss C. Theory-based Evaluation: Past, Present, and Future. In Rog D, Fournier D, editors. *Progress and Future Directions in Evaluation: Perspectives on Theory, Practice, and Methods. New Directions for Evaluation, no. 76*. San Francisco, CA: Jossey-Bass; 1997. <http://dx.doi.org/10.1002/ev.1086>
144. Bronfenbrenner U, Morris P. The Bioecological Model of Human Development. In Damon W, Lerner RM, editors. *Handbook of Child Psychology, Vol. 1: Theoretical Models of Human Development*. 6th edn. Hoboken, NJ: John Wiley & Sons, Inc.; 2006. pp. 793–828.
145. Akers RL, Krohn MD, Lanza-Kaduce L, Radosevich M. Social learning and deviant behaviour: a specific test of a general theory. *Am Sociol Rev* 1979;**44**:636–55. <http://dx.doi.org/10.2307/2094592>
146. Campbell M, Egan M, Lorenc T, Bond L, Popham F, Fenton C, *et al*. Considering methodological options for reviews of theory: illustrated by a review of theories linking income and health. *Syst Rev* 2014;**3**:114. <http://dx.doi.org/10.1186/2046-4053-3-114>
147. Harris R, Mosedale S, Garner J, Perkins E. What factors influence the use of contracts in the context of NHS dental practice? *Soc Sci Med* 2014;**108**:54–9. <http://dx.doi.org/10.1016/j.socscimed.2014.01.032>
148. Krieger N. Got Theory? On the 21st CE rise of explicit use of epidemiologic theories of disease distribution: a review and ecosocial analysis. *Curr Epidemiol Rep* 2014;**1**:45–56. <http://dx.doi.org/10.1007/s40471-013-0001-1>
149. Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ* 2006;**333**:597–600. <http://dx.doi.org/10.1136/bmj.333.7568.597>
150. Kneale D. *Pathways to Parenthood: Exploring Context as a Predictor of Time to First Parenthood*. London: Institute of Education, University of London; 2008.
151. Simons-Morton B, Farhat T. Recent findings on peer group influences on adolescent substance use. *J Prim Prev* 2010;**31**:191–208. <http://dx.doi.org/10.1007/s10935-010-0220-x>
152. Flood-Page C, Campbell C, Harrington V, Miller J. *Youth Crime: Findings from the 1998/99 Youth Lifestyles Survey*. London: Home Office; 2000.
153. Langford R, Bonell CP, Jones HE, Poulidou T, Murphy SM, Waters E, *et al*. The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. *Cochrane Database Syst Rev* 2014;**4**:CD008958. <http://dx.doi.org/10.1002/14651858.cd008958.pub2>
154. Bronfenbrenner U. *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press; 1979.

155. Ford DH, Lerner RM. *Developmental Systems Theory: An Integrative Approach*. Newbury Park, CA: Sage; 1992.
156. Gottlieb G. *Synthesizing Nature–Nurture: Prenatal Roots of Instinctive Behavior*. Mahwah, NJ: Lawrence Erlbaum; 1997.
157. Brandstädter J. Action Perspectives on Human Development. In Damon W, Lerner RM, editors. *Handbook of Child Psychology, Vol. 1: Theoretical Models of Human Development*. 6th edn. Hoboken, NJ: John Wiley & Sons, Inc.; 2006. pp. 516–68.
158. Dryfoos JG. *Adolescents at Risk: Prevalence and Prevention*. New York, NY: Oxford University Press; 1990.
159. Phinney JS. Ethnic identity in adolescents and adults: review of research. *Psychol Bull* 1990;**108**:499–514. <http://dx.doi.org/10.1037/0033-2909.108.3.499>
160. Berger P, Luckmann T. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor; 1967.
161. Giroux H. Dialectics and the Development of Curriculum Theory. In Pinar WF, editor. *Contemporary Curriculum Discourses*. New York, NY: Peter Lang; 1999. pp. 7–23.
162. Gitlin T. *Inside Prime Time*. London: Routledge; 1994.
163. Martin R. Truth, Power, Self: An Interview with Michael Foucault. In Hall B, Gillette A, Tandon R, editors. *Technologies of the Self: A Seminar with Michel Foucault*. London: Tavistock; 1982. pp. 9–15.
164. Baltes PB. On the incomplete architecture of human ontogeny. *Am Psychol* 1997;**32**:366–80. <http://dx.doi.org/10.1037/0003-066X.52.4.366>
165. Baltes PB, Staudinger UM, Lindenberger U. Lifespan psychology: theory and application to intellectual functioning. *Annu Rev Psychol* 1999;**50**:471–507. <http://dx.doi.org/10.1146/annurev.psych.50.1.471>
166. Gibson JJ. *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton-Mifflin; 1966.
167. Gibson JJ. *The Ecological Analysis Approach to Visual Perception*. Boston, MA: Houghton-Mifflin; 1979.
168. Hawkins JD, Weis JG. The social development model: an integrated approach to delinquency prevention. *J Prim Prev* 1985;**6**:73–97. <http://dx.doi.org/10.1007/BF01325432>
169. Jessor R, Jessor SL. *Problem Behavior and Psychological Development: A Longitudinal Study of Youth*. New York, NY: Academic Press; 1977.
170. Foschi M, Warriner GK, Hart SD. Standards, expectations, and interpersonal influence. *Soc Psychol Q* 1985;**48**:108–17. <http://dx.doi.org/10.2307/3033606>
171. Blumer H. *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, NJ: University of California Press; 1969.
172. Skinner BF. *Science and Human Behavior*. New York, NY: Free Press; 1953.
173. Gottlieb G. Normally occurring environmental and behavioral influences on gene activity: from central dogma to probabilistic epigenesis. *Psychol Rev* 1998;**105**:792–802. <http://dx.doi.org/10.1037/0033-295X.105.4.792-802>
174. Heckhausen J. *Developmental Regulation in Adulthood: Age-Normative and Sociostructural Constraints as Adaptive Challenges*. New York, NY: Cambridge University Press; 1999.
175. Heckhausen J. *Motivational Psychology of Human Development: Developing Motivation and Motivating Development*. *Advances in Psychology*, 131. New York, NY: Elsevier Science; 2000.

176. Heckhausen J, Wrosch C, Schulz R. A motivational theory of lifespan development. *Psychol Rev* 2010;**117**:32–60. <http://dx.doi.org/10.1037/a0017668>
177. Elder GH Jr. The Life Course and Human Development. In Damon W, Lerner RM, editors. *Handbook of Child Psychology, Vol. 1: Theoretical Models of Human Development*. 6th edn. Hoboken, NJ: John Wiley & Sons, Inc.; 2006. pp. 665–715.
178. Thelen E, Smith LB. Dynamic Systems Theories. In Damon W, Lerner RM, editors. *Handbook of Child Psychology, Vol. 1: Theoretical Models of Human Development*. 6th edn. Hoboken, NJ: John Wiley & Sons, Inc.; 2006. pp. 563–634.
179. Magnusson D. Holistic Interactionism: A Perspective for Research on Personality Development. In Pervin L, John O, editors. *Handbook of Personality: Theory and Research*. 2nd edn. New York, NY: Guilford Press; 1999. pp. 219–47.
180. Baltes PB, Baltes MM. Psychological Perspectives on Successful Aging: The Model of Selective Optimization with Compensation. In Baltes PB, Baltes MM, editors. *Successful Aging: Perspectives from the Behavioral Sciences*. New York, NY: Cambridge University Press; 1990. pp. 1–34. <http://dx.doi.org/10.1017/CBO9780511665684.003>
181. Ajzen I, Fishbein M. *Understanding Attitudes and Predicting Social Behavior*. Engelwood Cliffs, NJ: Prentice Hall; 1980.

Appendix 1 Protocol

The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes

Background

This review will synthesise evidence on positive youth development (PYD) interventions aimed at reducing substance use and violence, and inequalities in these outcomes.

Description of the problem

Young people in the UK have among the worst health in Europe with marked inequalities across the social scale (1, 2). Health risk behaviours increase during adolescence (3) and lead to high rates of later chronic disease and other problems, and substantial economic costs (4). Child poverty is currently increasing, raising the possibility of upward trends in young people's risk behaviours, with worrying implications for future chronic disease rates and NHS costs (5). Substance use and violence are highly prevalent and damaging to young people's long-term health. Rates of youth smoking, drinking and illicit drug use, collectively termed substance use, in the UK are among the highest in Europe (6, 7). One in four 15-year-olds are regular or occasional smokers, with previous declines having ceased since 2006 and just under a third of 15-year-olds drink alcohol every week (8). One in five 15-year-olds report drug use in the past month and over 11,500 under-18s access treatment services each year (9). Youth substance use reinforces existing socioeconomic inequalities in health across the life-course: substance use is most prevalent among socially-disadvantaged young people and frequent use at a young age is strongly associated with more harmful use and chronic illness in adulthood (10-12). Aggression and violence are similarly challenging. One survey reports that by age 15-16, a quarter of young people have carried a weapon and 19% reported attacking someone with the intention to hurt them seriously (13). Violence is subject to marked health inequalities (14) and is associated with an increased risk of: physical health problems (15); engaging in health risk behaviours such as substance use (16-18); long-term emotional, behavioural and mental health problems (15, 19, 20); and self-harm and suicide (21). The economic costs associated with youth substance use and aggression are extremely high (10, 22, 23).

Description of the intervention

Interventions to reduce health risks in adolescence are potentially highly cost effective (24). There are increasing calls for adolescent health interventions to address multiple rather than single risk behaviours because such behaviours cluster together (25, 26) and because such interventions are potentially more feasible and efficient (27). PYD is one such intervention to address inter-clustered risk behaviours among young people. PYD is the dominant paradigm in youth work in the UK. The National Youth Agency (NYA), the major youth work organisation in the UK, defines such interventions as voluntary and informal educational activities aiming to bring about generalised youth development rather than merely remedying 'problem behaviours'. Such development is defined in terms of the promotion of positive skills, attitudes, relationships and identities (28). A literature review published by NYA developed a complex definition of PYD in terms of philosophy, constructs, domains and processes but similarly emphasised young people's positive attributes and competencies through structured voluntary activities (29).

Similarly, in the USA, PYD is defined in terms of its goal of developing: bonding; resilience; social, emotional, cognitive, behavioural or moral competence; self-determination; spirituality; self-efficacy; clear and positive identity; belief in the future; recognition for positive behaviour; opportunities for pro-social involvement; and/or pro-social norms (24, 40). academic, cognitive or vocational skills; confidence; connections to peers and adults; character in terms of self-control, respect and morality; and caring for others (30) Drawing on these definitions, we define PYD as involving voluntary education provided by youth workers outside of normal

school time targeting young people age 11-18 not merely the prevention of problem behaviour but addressing generalised, positive development in terms of: bonding; resilience; social, emotional, cognitive, behavioural or moral competence; self-determination; spirituality; self-efficacy; clear and positive identity; belief in the future; recognition for positive behaviour; opportunities for pro-social involvement; and/or pro-social norms. Our definition excludes PYD delivered in school time because this has been the subject of recent reviews (31, 32). It also excludes interventions delivered in custodial or probationary settings, clinical settings, employment training for school leavers or that primarily target families. PYD has the potential to reduce substance use and violence through various complex pathways. First, PYD can address some of the underlying, social determinants of these outcomes, such as disengagement from education, lack of social support and low aspirations for the future (24). Second, PYD can divert young people away from substance use and violence through engaging them in more positive forms of recreation (30). Third, PYD can promote social and emotional competences, which are an important protective factor against adolescent health-risk behaviours (24). Fourth, PYD providers can provide credible health messages and signpost health services (33).

Even in a context of public-sector cuts, there is major investment in such interventions. The UK government's Positive for Youth (34) report announced a multi-million pound investment in youth work, youth centres, the National Citizen Service and other youth volunteering projects. The most recent public health white paper (35) cited such work as a key element in promoting young people's health. The London mayor and local government across the UK are also investing millions of pounds in various PYD interventions (36). The devolved governments in Scotland and Wales also emphasise these principles and promote investment in PYD (37, 38).

However, despite this widespread investment and potential, the evidence base for the public health benefits of such interventions is unclear. While a systematic review examining non-health outcomes (39) reported benefits for self-confidence and self-esteem, school bonding, positive social behaviours, school grades and achievement test scores, the review did not systematically examine health effects. Systematic reviews of health outcomes have so far only focused on sexual health (40, 41), reporting sustained effects but with considerable unexplained variability between programmes. For example, the Children's AID Society Carrera programme reduced teenage pregnancy in some US sites but not others (33), while two evaluations of PYD interventions in the UK respectively suggested adverse and no effects on sexual health (42, 43). US researchers have argued that some youth programmes which target 'delinquent' young people and which are relatively unstructured may actually reinforce violence and anti-social behaviours via 'peer deviancy training' (44).

Others have disputed this, referring to meta-analyses of interventions addressing youth delinquency (45) which suggest that targeting and structure of sessions do not moderate effects. However, no systematic review focused on PYD interventions has examined these questions. Non-systematic review of PYD effects on violence and drug use (24, 46) have reported benefits as well as variability, but their findings must be treated with caution given that they were unsystematic and are now quite old.

Rationale for current study

This review will fill two timely and important knowledge gaps and provide important evidence to local government commissioners of youth services and public health. First, it will aim to synthesize evidence on the effectiveness and cost-effectiveness of PYD interventions in reducing substance use and violence. Second, it will aim to examine how effects vary according to the characteristics of participants (in order to assess what works for whom and estimate effects on health inequalities) and context (in order to assess what works in what

contexts and determine likely generalisability to different settings). Addressing the first gap, concerning intervention effects, is timely and important because, as described above, UK young people have among the worst health in Europe with marked inequalities across the social scale and PYD interventions are receiving significant policy attention and investment despite a lack of evidence of health benefits from systematic reviews. Addressing the second gap, concerning moderators of effects is also important given the possibility discussed above that PYD effects will vary, and given our interest in assessing the potential of PYD to reduce health inequalities.

Our synthesis of theories of change and process evaluations will inform the development of a taxonomy of interventions and of hypotheses to be tested regarding how intervention effects are moderated by characteristics of participants and context. We will then use a combination of meta-regression and qualitative comparative analysis to test these hypotheses. Thus, our synthesis will facilitate a more informed view of the likely impact of PYD on public health and health inequalities internationally and in the UK, and which approaches have the most potential for public health improvement in different settings and populations across the UK. If appropriate, our review will inform our development of a research proposal to evaluate the effectiveness of a PYD intervention in the UK. We anticipate that such an intervention would be coordinated by the National Youth Agency, which is a collaborating institution on this proposal. Our research will also make a major contribution to the rapidly developing field of "implementation science" (47) through its application of meta-regression and qualitative comparative analysis to public health interventions for young people.

Research aims

To search systematically for, appraise the quality of and synthesise evidence to address the following review questions: RQ1. What theories of change inform PYD interventions delivered to young people aged 11-18 addressing substance use and violence?

RQ2. What characteristics of participants and contexts are identified as barriers and facilitators of implementation and receipt in process evaluations of PYD?

RQ3. What is the effectiveness and cost-effectiveness of PYD when compared to usual or no treatment in reducing substance use (smoking, alcohol, drugs), and violence (perpetration and victimization)?

RQ4. What characteristics of participants and contexts appear to moderate, or are necessary and sufficient for PYD effectiveness?

Research objectives

(1) To conduct electronic and other searches for studies of PYD interventions by December 2013.

(2) To screen references and reports for inclusion in the review by February 2014.

(3) To extract data from and assess the quality of included studies by May 2014.

(4) To synthesise thematically theories of change of PYD interventions to produce a taxonomy and theory of change of PYD interventions by July 2014.

(5) To synthesise process evaluations of PYD interventions by September 2014.

(6) To consult with policy/practice and young people to validate the resultant taxonomy and theory of change by October 2014.

(7) To synthesise outcome and economic evaluation data and undertake meta-regression and qualitative comparative analyses by December 2014.

(8) To draw on these syntheses to draft a report addressing our review questions by February 2015.

(9) To consult with policy/practice and young people on the draft report to inform amendments and dissemination by March 2015.

(10) To submit the final report to NIHR by May 2015.

Research design

Our proposal is for a multi-method systematic review of theories of change, and process, outcome and economic evaluations of PYD interventions delivered to young people age 11-18 addressing substance use (smoking, alcohol, drugs) and violence (perpetration and victimization). The review will follow existing general criteria for the good conduct and reporting of systematic reviews (e.g. Centre for Reviews and Dissemination; Preferred Reporting Items for Systematic Reviews and Meta-Analyses). The review protocol will be registered with PROSPERO International Prospective Register of Systematic Review (<http://www.crd.york.ac.uk/Prospero/>).

Our four components fit together as follows. Our review and thematic synthesis of theories of change (RQ1) will enable us to create a taxonomy of PYD interventions and theory of change of how these are implemented, and aim to achieve their effects in different contexts and subgroups. Our review and thematic synthesis of process evaluations (RQ2) will enable us to refine this theory of change to incorporate hypotheses about potential barriers and facilitators of implementation and receipt relating to characteristics of participants and contexts. Our review of outcome and economic evaluations (RQ3) will enable us to estimate the effectiveness and cost-effectiveness of PYD. Depending on the results of the taxonomy and the heterogeneity of evaluation studies found in relation to this taxonomy, this will either examine the effectiveness of PYD overall or of sub-types of PYD.

Our use of meta-regression and qualitative comparative analysis will allow us to develop hypotheses about what factors relating to participants and contexts moderate / are necessary and sufficient for intervention effects (RQ4).

Size of available literature

On 22 March 2013, we conducted a search in PubMed using the search string given in appendix 2. This was a limited search in that: (i) it searched only medical journals when it is likely that relevant studies are published in sources other than journals as well as in journals in the fields of criminology, sociology, psychology, education and public/social policy; (ii) it is not possible to use adjacency terms within PubMed; and (iii) we narrowed the search by including terms for study design which would not occur in a full search. Our search identified 2,209 references which were screened on title and abstract only. Of these, 57 studies appeared very likely to meet our inclusion criteria while 138 might possibly meet these criteria but would require screening of the full report to decide. Of the 57 studies which were deemed likely to be included, 25 were outcome evaluations, 23 process evaluations and 9 were theoretical frameworks. The preliminary search and screening confirmed the utility and applicability of our inclusion criteria. While this was by no means an exhaustive search, it does suggest that we will be able to identify a sufficiently large number of relevant studies which will enable us to answer our review questions using the methods specified.

Criteria for considering studies for this review

Types of participant

We will include studies conducted where a majority of participants are age 11-18 years. We will exclude studies of populations targeted on the basis of pre-defined physical and mental health conditions but not by pre-existing risk behaviour or other forms of targeting (e.g. area-level deprivation).

Types of intervention

We will include PYD interventions involving voluntary education which aim not merely to prevent problem behaviour but aim to address generalised (beyond health) and positive (beyond avoiding risk) development in terms of promoting: bonding (developing the child's relationship with a healthy adult, positive peers, school, community, or culture); resilience (strategies for adaptive coping responses to change and stress, and promoted psychological flexibility and capacity); social competence (developmentally appropriate interpersonal skills, and rehearsal strategies for practicing these skills including communication, assertiveness, refusal and resistance, conflict-resolution, and interpersonal negotiation strategies for use with peers and adults); emotional competence (identifying feelings in self or others, skills for managing emotional reactions or impulses, or skills for building the youth's self-management strategies, empathy, self-soothing, or frustration tolerance); cognitive competence (cognitive abilities, processes, or outcomes, including academic performance, logical and analytic thinking, problem-solving, decision-making, planning, goal-setting, and self-talk skills); behavioural competence (skills and reinforcement for effective verbal, non-verbal and other actions); moral competence (empathy, respect for cultural or societal rules and standards, a sense of right and wrong, or a sense of moral or social justice); self-determination (capacity for empowerment, autonomy, independent thinking, or self-advocacy, or their ability to live and grow by self-determined internal standards and values which may or may not include group values); spirituality (beliefs in a higher power, internal reflection or meditation, or supported youth in exploring a spiritual belief system, or sense of spiritual identity, meaning, or practice); self-efficacy (personal goal-setting, coping and mastery skills, or techniques to change negative self-efficacy expectancies or self-defeating cognitions); clear and positive identity (healthy identity formation and achievement in youth, including positive identification with a social or cultural sub-group that supports their healthy development of sense of self); belief in the future (belief in his or her future potential, goals, options, choices, or long range hopes and plans were classified as promoting belief in the future, including guaranteed tuition to post-secondary institutions, school-to-work linkages, future employment opportunities, or future financial incentives to encourage continued progress on a pro-social trajectory; or optimism about a healthy and productive adult life); recognition for positive behaviour (response systems for rewarding, recognizing, or reinforcing children's pro-social behaviors were classified as using recognition for positive behaviour); opportunities for pro-social involvement (activities and events in which youths could actively participate, make a positive contribution, and experience positive social exchanges); and/or pro-social norms (clear and explicit standards for behavior that minimized health risks and supported pro-social involvement). Informed by (24, 40). PYD interventions address one of these but applied to different domains (family, community, school) or more than one of these in a single domain. Our definition excludes PYD delivered in school time, or in custodial or probationary setting, clinical settings or employment training for school leavers. It also excludes interventions that target parents/carers alongside young people in order to focus on family functioning.

Types of outcome

We will include studies addressing:

Substance use (smoking, alcohol and/or drug use); or Violence (perpetration and/or victimization).

Informed by existing systematic reviews focused on substance use and violence among young people (48-51), outcome measures may draw on dichotomous or continuous variables, and self-report or observational data. They may use measures of frequency (monthly, weekly or daily), the number of episodes of use or an index constructed from multiple measures. Alcohol measures may examine alcohol consumption or problem drinking. Drug outcomes may examine drugs in general or specific illicit drugs. Measures of violent and aggressive behaviour may examine the perpetration or victimization of physical violence including violent crime. Our Data Analysis section describes how we will combine measures.

Types of studies

In order to address RQ1, we will include studies describing PYD intervention theory of change. We will define theory as we did in our previous NIHR/PHR funded review of the effects of schools and school-environment interventions on health (52). Included studies may deal exclusively with theory of change or might address it alongside the reporting of empirical data. In order to address RQ2, we will include studies reporting on process evaluation of PYD intervention. This would include studies reporting on the planning, delivery, receipt or causal pathways of PYD using quantitative and/or qualitative data. These studies may report exclusively on process evaluations or report process alongside outcome or economic data. In order to address RQ3, we will include studies reporting on outcome and economic evaluations of PYD interventions. We will include experimental (randomized controlled trials) and quasi-experimental studies (employing non-randomized prospective comparison groups). Control groups will receive usual care or no treatment. Economic studies addressing RQ3 will be defined in terms of their comparison of the costs and consequences of two or more interventions or, where there is good reason to believe outcomes are similar, involve cost-minimisation analyses. In order to address RQ4, we will draw on syntheses of all of the above study types.

Language

We will only include studies published in English because these interventions have been overwhelmingly developed in English speaking countries (24, 40).

Dates

We will only search as far back as studies published in 1985 since this is when PYD interventions first began to be developed (24, 40).

Search methods for the identification of studies

In appendix 1, we provide the search string that we have used in a preliminary search in PubMed. As explained above, this was a limited search but it will inform the development of a more sophisticated search strategy maximising sensitivity as recommended by the Cochrane Handbook for Systematic Reviews of Interventions (53).

We will also learn from previous systematic reviews focused on the effects of PYD on sexual health (40, 41) in developing our search strategy. The studies sought by this review are not likely to be reliably indexed in databases with controlled vocabularies. So we anticipate our searches involving a large number of free text terms. We will take the following essential concepts of the inclusion criteria to develop the search string: young people; and positive youth development. If these searches elicit over 30,000 hits in pilot searches on ASSIA and Medline we will consider introduce a set of terms for our outcome measures. Our searches will involve different free text and controlled vocabulary terms for each of these concepts using the Boolean operator "AND". The concepts will be linked by the Boolean operator "OR". The combination of these three concepts is considered specific enough to include all available studies regardless of study design. We will not restrict the searches by date, language or publication type.

Electronic searches

We will search the following databases from inception to present: ASSIA (Applied Social Science Index and Abstracts); Australian Educational Index; BiblioMap (Database of health promotion research); British Educational Index; CENTRAL (Cochrane Central Register of Controlled Trials); The Campbell Library; CINAHL; CISDOC (The Health and Safety Information Centre of the International Labour Office); Cochrane Controlled Trials Database; DARE (Database of Abstracts of Reviews of Effects); Database of Educational Research; Econlit; ERIC; Health Management Information Consortium; IBSS (International Bibliography of the Social Sciences); International Clinical Trials Registry Platform; Medline; NHS Economic Evaluation Database; OpenGrey (System for Information on Grey Literature in Europe); Proquest Dissertations and Theses; PsycInfo; Social Policy and Practice including Child Data and Social Care Online; Social Science Citation Index/Web of Knowledge; Sociological Abstracts; Dissertation Abstracts/Index to Theses; and TRoPHI (Trials Register of Promoting Health Interventions).

Searching other resources

1. We will carefully search reference lists from all studies that meet the inclusion criteria.
2. We will only hand search those journals which; i) contain studies that we include, ii) which are found only via reference checking and iii) which are not indexed on databases we have searched. We will hand search these initially for the last 5 years and if these elicit >1 new included study hand search for a further 5 years back.
3. We will search for references to relevant studies in international government reports and non-governmental organization publications via a Google search.
4. We will contact subject experts to identify unpublished or on-going research.
5. We will search all available clinical trials registers (e.g. clinicaltrials.gov) for relevant ongoing and unpublished trials.

Data collection and assessment

Selection of studies

Search results will be downloaded to EPPI-Reviewer 4. A worksheet with the inclusion criteria operationalised into exclusion criteria and guidance notes will be prepared and piloted by two reviewers screening 100 references. Pilot screening results will be discussed by pairs of reviewers involved in screening to ensure consistency in applying the criteria. A 90% agreement rate will be required before proceeding to independent screening of the full data set. If a single reviewer cannot reach a decision regarding inclusion of a specific article, judgement for selection will be referred to a second reviewer. If both reviewers disagree and cannot reach a consensus a third reviewer will be consulted. Full reports will be obtained for those references judged as meeting our inclusion criteria or where there is insufficient information from the title and abstract to judge inclusion. A second round of screening will then occur using the same approach but based on full study reports in order to determine which studies are included in the review. We will maintain a record of the selection process for all screened material.

Data extraction and management

Two reviewers will independently extract data from studies meeting the inclusion criteria, using a piloted data extraction form with guidance developed for this review. Where the two authors disagree, they will meet to discuss this and if possible reach a consensus. If the reviewers cannot reach consensus regarding the particulars of data extraction for a specific study, judgement will be referred to a third reviewer. Included studies will be described using the EPPI-Centre classification system for health promotion and public health research (54),

supplemented by additional codes developed for this review. For all studies where relevant, we will extract information pertaining to: basic study details (individual and organizational participant characteristics, study location, timing and duration, research questions or hypotheses); study design and methods (design, allocation, blinding, sample size, control of confounding, accounting for data clustering, data collection, attrition, analysis); intervention characteristics (including timing and duration, programme development, theoretical framework/logic model, content and activities, providers and details of any intervention offered to the control group); process evaluation of the intervention (feasibility, fidelity/quality, intensity, coverage/accessibility, acceptability, mechanism and context using an adapted version of an existing tool (55)); outcome measures at follow-up(s) (reliability of measures, effect size both overall and where available by age, sex, socio-economic status and ethnic sub-group). For economic analyses, we will extract data in order to complete a number of data tables. These will include data on the key study design elements and results for each identified study; for example, the intervention costs, price year, time horizon, base case assumptions and perspective taken regarding cost and effectiveness estimates, and estimates of cost-effectiveness such as incremental cost-effectiveness ratios and net benefit statistics. The two reviewers will independently enter data from the data extraction forms into EPPI-Reviewer 4.

Published reports may be incomplete in a wide range of ways. For example: they may not report sufficient detail about their participants for our equity analysis; they may not present information on all the outcomes that were measured (possibly resulting in outcome reporting bias); they may not provide sufficient information about the intervention for accurate characterisation; and they may not report the necessary statistical information for the calculation of effect sizes. In all cases where there is a danger of missing data affecting our analysis, we will contact authors of papers wherever possible to request additional information. If authors are not traceable or sought information is unavailable from the authors within two months of contacting them, we will record that the study information is missing on the data extraction form, and this will be captured in our risk of bias assessment of the study.

Assessment of quality and risk of bias

We will assess the quality of theoretical literature using a framework previously developed in our NIHR-funded systematic review of the health effects of schools and school –environment interventions, which assesses factors such as clarity and parsimony (52). We will assess the quality of qualitative studies using standard Critical Appraisal Skills Program and EPPI-Centre tools for qualitative studies (56). Quality tools for qualitative studies address 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. A final step in the quality assessment of qualitative studies will be to assign studies two types of ‘weight of evidence’.

First, reviewers will be asked to assign a weight (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). Second, reviewers will also be asked to assign an additional weight (low, medium, high) to rate the usefulness of the findings for shedding light on factors relating to the review questions. Guidance will be given to reviewers to help them reach an assessment on each criterion and the final weight of evidence. For outcome evaluations, we will assess risk of bias within each included study using the tool outlined in the Cochrane Handbook for Systematic Reviews of Interventions (53). For each study, two reviewers will independently judge the likelihood of bias in seven domains: sequence generation; allocation concealment; blinding (of participants, personnel, or outcome assessors); incomplete outcome data; selective outcome reporting; and other sources of bias (e.g. recruitment bias in cluster-randomised studies); and intensity/type of comparator. Each

study will subsequently be allocated a score of 'high risk', 'low risk' or 'unclear risk' within each domain. In cases of disagreement, the reviewers will meet to establish consensus but where the two authors cannot reach consensus regarding categorisation for risk of bias for a specific study, we will refer judgement to a third reviewer. For economic evaluations, we will assess quality using the Drummond checklist (57), supplemented with the Philips checklist (58) if an evaluation contains a decision model component. The Drummond checklist asks general questions about issues such as the comprehensiveness of the descriptions regarding the interventions at hand and the quality of the effectiveness evidence underpinning the evaluation. The Philips checklist asks detailed questions regarding factors such as the rationale for the model structure and the overall quality of the modelling methodology. We will also consider adding extra questions to the Drummond checklist for issues that are particularly pertinent to PYD interventions such as costs accruing to participants' families, where we believe use of the Drummond checklist alone might miss in terms of critical appraisal (59). We will assess reporting bias according to Sterne's guidance (60). We will reduce the effect of reporting bias by including studies and not publications in order to avoid the introduction of duplicated data (i.e. two articles could represent duplicate publications of the same study). Following the Cho statement on redundant publications (61), we will attempt to detect duplicate studies and, if multiple articles report on the same study, we will extract data only once. We will prevent location bias by searching across multiple databases. We will prevent language bias by not excluding any article based on language.

Data analysis

RQ1 and 2: Thematic synthesis of theory and qualitative data

Syntheses of theories of change (RQ1) and process evaluations (RQ2) will employ thematic synthesis methods (62-64) to develop and refine: a taxonomy of PYD interventions; theory of change underlying these interventions; and characteristics of participants and context acting as potential barriers and facilitators of implementation and receipt. These syntheses will not be restricted to studies judged to be of high quality but instead conclusions drawing on poorer quality evidence will be given less weight. Detailed evidence tables will be prepared to describe: the methodological quality of each study; details of the intervention examined; study site/population; and full findings. Two reviewers will read and re-read data contained within the evidence tables, apply line-by-line codes and memos to capture the content of the data, and then group and organise codes into higher-order themes. These themes will be used to generate an explanatory framework to address

RQ2. RQ3: Synthesis of overall quantitative outcome and economic data

Outcome evaluations

In order to address RQ3, we will produce pooled estimates using EPPI-Reviewer 4 and a narrative account of the effectiveness interventions. We will regard follow-up times of less than three months, three months to one year and more than one year as different outcomes. Once we know the number of studies and the extent of heterogeneity amongst the studies (as determined by a Q test), we will make a decision whether to calculate an effect size across all outcomes (i.e., combining substance use and violent behaviour outcomes), or run analyses separately for the different outcomes, or conduct a multivariate meta-analysis (65). Once that decision has been made, or each study grouping that has a sufficient number of studies to undertake meta-analysis, we will produce forest plots. A forest plot includes the point estimates and standard errors for each study, such as risk ratios for dichotomous outcomes or standardised mean differences for continuous outcomes, with the estimates weighted by a function of the sample sizes. When the results cannot be plotted, we will describe them in the 'characteristics of included studies' table, or enter the data into additional tables. We will use the chi-squared test and the I^2 statistic to measure heterogeneity. The results of these statistical tests will be evaluated in accordance with the Cochrane handbook (53). If we

consider that we have unexplained statistical heterogeneity in any of our study groupings, we will investigate it further using subgroup and sensitivity analyses. When studies are found to be statistically heterogeneous, we will use a random effects model; otherwise we will use a fixed-effect model. When using the random-effects model, we will conduct a sensitivity check by using the fixed-effect model to reveal differences in results. If an indication of substantial heterogeneity is determined (e.g. I^2 value of greater than 50%), that cannot be explained through analyses of variance or meta-regressions (see analysis plan for RQ4, below), then we will not produce a pooled estimate and will present a narrative summary of our findings. The narrative report will classify and present studies according to: intervention content; research design; time to outcome measurement; outcome measure; and intervention effect.

Prior to synthesis, we will check for correct analysis by cluster and report values of: intra-cluster correlation coefficients (ICC), cluster size, data for all participants or effect estimates and standard errors. Where proper account has not been taken of data clustering, we will correct for this by inflating the standard error by the square root of the design effect [63] and, for dichotomous outcomes, adjusting the numerators and denominators (51). Where ICCs are not reported we will contact authors to request this information or impute one, based on values reported in other studies. Where imputation is necessary, we will undertake sensitivity analyses to assess the impact of a range of possible values. In other instances of missing data (such as missing population information) it may not be possible to include a study in a particular analysis if, for example, it is impossible to classify the population using our equity tool.

We will use the GRADE approach as described in the Cochrane Handbook for Systematic Reviews of Interventions to present the quality of evidence and 'Summary of findings' tables. The downgrading of the quality of a body of evidence for a specific outcome will be 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 will draw funnel plots to assess the presence of possible publication bias (trial effect versus standard error). While funnel plot asymmetry may indicate publication bias, this can be misleading with a small number of studies. We will discuss possible explanations for any asymmetry in the review in light of our number of included studies.

We will undertake a sensitivity analysis to explore whether the findings of the review are robust in light of the decisions made during the review process. We will also assess 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. Where meta-analysis establishes that PYD has effects on primary outcomes and where data allows, we will undertake additional exploratory meta-analyses to determine PYD effects on social outcomes to examine the plausibility that such outcomes might lie on causal pathways. Such analyses will be informed by the synthesis of process evaluation findings to avoid data dredging, and compare those findings to the results of analyses with studies of all quality levels included.

Economic evaluations

Our team includes a health economist who will appraise the evidence to determine the most appropriate approach to synthesis. The overall approach to critiquing and synthesising the information on cost-effectiveness will be narrative using standard checklists (57). However, in the event that there is substantial variation across study results, particular emphasis will be placed on identifying the reasons for these differences. Cost estimates will be inflated to current prices and converted into UK currency using purchasing power parity statistics to aid the comparison of results. Where included studies report on cost effectiveness using non-monetary outcome indicators such as QALYs, we will examine this.

However, we anticipate that this will be rare and most studies will report cost-effectiveness in terms of the costs of averting our primary outcomes of substance use and violence. The main limitation of this approach is that trade-offs between different outcomes are not made explicit and results are difficult to interpret when a particular intervention is associated with 'better' outcomes on some scales, but 'poorer' scores on others. We will, if feasible, undertake supplementary decision modelling alongside reviewing to link observed trial effects to longer-term health outcomes and their associated costs. However, this may not be possible. In 2009, Hummel et al. built an economic model for the NICE public health programme to assess the cost-effectiveness of universal interventions that aimed to promote emotional and social wellbeing in secondary schools (32). While they identified a number of published longitudinal studies, they concluded that existing data-sets did not contain appropriate information to estimate these longer-term effects and emphasised the many caveats with respect to the number of assumptions made by the modelling required to estimate long-term effects. The NICE public health methods guidance clearly states that complex modelling should be avoided if it is likely to create cost-effectiveness estimates that are highly uncertain (66).

RQ4: Meta-regression and qualitative comparative analysis

If sufficient studies are found, we will employ meta-regression using Stata 12 to investigate what factors moderate intervention effects (67, 68) in order to examine RQ4. It may not be feasible to apply this method if we judge there are too many confounders, insufficient data or meta-regression is unable to account for inter-dependencies in complex interventions. Hence, we will complement meta-regression with qualitative comparative analysis, adapted for use in research synthesis (69) to assess necessary and sufficient conditions for intervention effectiveness. We will employ EPPI-Reviewer (70) and "fsQCA" software as appropriate (71) to do this. The use of initial hypotheses derived from work addressing RQ2 will protect us from 'dredging' the data for spurious statistically significant results. The required steps of 'qualitatively anchoring' outcomes in qualitative comparative analysis will ensure that changes in outcomes are meaningful and not simply statistical artefacts with little relevance for decision-making (71). These analyses will include examination of how intervention effects are moderated by characteristics of participants (for example in terms of individual socioeconomic status, sex and ethnicity) and contexts (for example in terms of area deprivation), in order to examine potential impacts on health inequalities. This will draw on existing methods involving an 'equity lens' (72) employing meta-regression to examine effects by participant sub-group. We will examine whether participant socio-economic status, sex and ethnicity, and area deprivation moderate PYD effectiveness. We should stress that meta-regression and qualitative comparative analysis will be exploratory analyses oriented towards hypothesis development rather than testing since these will draw on observational rather than experimental comparisons.

Reporting

Our aim is to provide research outputs which provide rigorous evidence on the potential effectiveness and cost-effectiveness of PYD in preventing violence and substance use, and reducing health inequalities, and to inform the future optimization of PYD. Our synthesis will enable a more informed view of the likely impact of PYD on public health and health inequalities internationally and in the UK, and which approaches have the most potential for public health improvement in different settings across the UK. If appropriate, our review will inform our development of a research proposal to evaluate the effectiveness of a PYD intervention in the UK. We anticipate that such an intervention would be coordinated by the National Youth Agency, which is a collaborating institution on this proposal.

We will produce three reports: a full technical report for NIHR; a briefing report for policy and practice audiences (particularly targeting local authority commissioners of public health and youth services); and a concise young people's report in consultation with the 'ALPHA' (Advice

Leading to Public Health Advancement) young people's group. All three reports will be published online. Stakeholders will be able to use these latter two reports as a yard-stick in assessing local services and lobbying for more effective services. The research will be launched at an event organised through the Association for Young People's Health group, which is a collaborating partner in the research. In addition, we will disseminate the research via open-access scientific journals, and via academic and policy conferences. We will undertake seminars involving representatives of the UK and devolved national government departments of health and education to present the research to discuss policy implications and next steps. We will also use stakeholder and academic networks to support dissemination as well as existing web and social media platforms based at the EPPI-Centre (Institute of Education) and DECIPHer (Cardiff University).

Socioeconomic position and inequalities

Socioeconomic status and health inequalities are central to the research. PYD interventions are generally targeted towards disadvantaged settings and populations with the aim of reducing inequalities. Our review will examine evidence for them doing so effectively. As described above, as well as examining overall effectiveness, where data allow our review will use an 'equity lens' (72), as described above. We will examine whether participant socioeconomic status, sex and ethnicity, and area deprivation moderate PYD effectiveness.

Research governance and ethics

The principal investigator is Chris Bonell, Professor of Sociology and Social Policy, Department of Childhood, Families and Health, Institute of Education, University of London who is responsible for the conduct and delivery of the work. The sponsor of the research is Steve Denton, Pro-Director of

Strategy and Organisation at the Institute of Education. The co-applicants will form an investigator committee which will meet monthly throughout the project, overseeing its conduct. These meetings will be minuted to keep a record of tasks, deadlines and responsibilities. Since the research involves no human participants and draws solely on evidence already in the public realm, there is no requirement for review by research ethics committee. While recognising that systematic reviews do not have the same potential for direct harm that primary research can have, it is essential that the team has considered relevant issues. The team will therefore follow relevant guidelines and best practice. EPPI-Centre staff follow the Social Research Association's (SRA) ethical guidelines (73) and refer also to guidance recommended by the National Coordinating Centre for Public Engagement (74). The SRA guidelines emphasise four obligations to: society; funders and employers; colleagues; and research participants. The second set of guidelines emphasise seven principles of particular relevance for work requiring public engagement: mutual respect; equality and inclusion; democratic participation; active learning; making a difference; collective action; and personal integrity.

Project timetable and milestones

- October-December 2013: Electronic and other searches for studies of PYD interventions.
- December 2013-February 2014: Screening of found references and reports for inclusion in the review.
- February-May 2014: Data extraction and quality assessment of included studies.
- May-July 2014: Thematic synthesis of theories of change of PYD interventions to produce taxonomy and theory of change.
- July-September 2014: Synthesis of process evaluations of PYD interventions.
- October 2014: Consultation with policy/practice and youth stakeholders to validate the taxonomy and theory of change.
- October-December 2014: Synthesis of outcome and economic evaluation data and

- undertake meta-regression and qualitative comparative analyses.
- December 2014-February 2015: Draft report addressing our review questions.
- March 2015: Consultation with policy/practice and youth stakeholders on the draft report to inform amendments and dissemination.
- May 2015: Submission of the final report to NIHR.

Expertise

The team for this project involves investigators with recognised expertise and experience in: synthesis of theories of change, and of quantitative and qualitative evidence; evaluation of PYD and other complex interventions aiming to promote young people's health; thematic and statistical meta-analysis, meta-regression and qualitative comparative analysis, and economic evaluation. Prof. Chris Bonell (Professor of Sociology and Social Policy, Institute of Education) will direct the project, lead on searching, appraisal and qualitative synthesis, and act as a second reviewer. Prof. Bonell will provide supervision to the research officer employed to work on the project. This will involve weekly meetings with all actions minuted. Dr James Thomas (Reader in Social Policy, Institute of Education) will lead statistical meta-analyses and meta-regressions, and qualitative comparative analysis, and act as a second reviewer on some studies. Dr Adam Fletcher (Senior Lecturer in Social Science and Health, Cardiff University) will advise on synthesis of qualitative evidence, lead consultation with young people, and act as second reviewer on some studies. Prof. Rona Campbell (Professor of Public Health Research, University of Bristol) will advise on the systematic review of complex interventions. Dr Simon Murphy (Senior Research Fellow, DECIPHer, Cardiff University) will be an unfunded co-applicant, advising on stakeholder consultation, dissemination and knowledge transfer. Dr Alec Miners and Ms Claire Stansfield are not co-applicants but are nonetheless key, costed members of our team. Dr Miners is an experienced health economist who will lead our economic analyses. Ms Stansfield is an information scientist who will lead our search strategy.

Partner collaboration

We have consulted with potential users of the research to gauge their interest in the work, identify additional priorities and establish whether they might sit on our stakeholder consultative body. These stakeholders include the Department of Health (Richard Sangster), Department for Education (Richard White), National Youth Agency (Alex Stutz), Action on Smoking and Health (ASH) Wales and Association for Young People's Health (John Coleman), all of whom support the work and will collaborate on it. We also consulted with the ALPHA young people's public input advisory group based in DECIPHer (Development and Evaluation of Complex Public Health Interventions for Public Health Improvement) across the universities of Cardiff, Bristol and Swansea. This consultation informed our view that synthesising evidence on PYD was a priority, and that this should include assessments of processes alongside outcomes in order to consider the acceptability of interventions as well as their potential transferability across different settings. This consultation also informed our decision to prioritise evidence of effects on substance use (smoking, alcohol and drugs) and violence (perpetration and victimization).

As with our previous evidence synthesis of the effects on health of schools and school environment interventions, funded by the NIHR Public Health Research Programme, we will consult with policy/practice and youth stakeholders in the course of the project. We will convene an advisory group of the above policy/practice stakeholders, and we will consult separately with the ALPHA young people's public input advisory group based in DECIPHer. Consultations with each of these two bodies will occur at two points: first, when we have synthesised evidence addressing theories of change and process evaluations, in order to validate and refine our theory of change; and second, during the write up of the research, to

inform how research outputs are structured and disseminated. If appropriate, our review will inform our development of a research proposal to evaluate the effectiveness of a PYD intervention in the UK. We anticipate that such an intervention would be coordinated by the National Youth Agency, a collaborating institution on this proposal.

References

1. UNICEF. Child poverty in perspective: an overview of child well-being in rich countries. Innocenti Report Card 7. Florence: UNICEF; 2007.
2. Coleman J, Brooks F, Threadgold P. Key Data on Adolescence 2011: The Latest Information and Statistics about Young People Today London: Association for Young People's Health; 2011.
3. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. *Lancet*. 2012;379(9826):1641-52.
4. Sawyer SM, Afifi RA, Bearinger LH, Blakemore SJ, Dick B, Ezech AC, et al. Adolescence: a foundation for future health. *Lancet*. 2012;379(9826):1630-40.
5. DWP. Households Below Average Income: An Analysis Of The Income Distribution 1994/5-2010/11. London: Department for Work and Pensions; 2012.
6. Hibell B, Guttormsson U, Ahlström S, Balakireva O, Bjarnason T, Kokkevi A, et al. The 2011 ESPAD Report: Substance Use Among Students in 36 European Countries. Stockholm: The Swedish Council for Information on Alcohol and other Drugs; 2011.
7. Fuller E. Smoking, drinking and drug use among young people in England in 2011. London: National Centre for Social Research; 2012.
8. NatCen/NFER. Smoking, Drinking and Drug Use among Young People in England in 2009. London: NHS Health and Social Care Information Centre; 2010.
9. National Treatment Agency for Substance Misuse. Drug Treatment Activity in England 2006/07. London: Department of Health; 2007.
10. Donaldson L. Tackling the health of the teenage nation: Chief Medical Officer's Annual Report 2007. London: Department of Health, England; 2008.
11. Patton GC, Coffey C, Carlin JB, Degenhardt L, Lynskey M, Hall W. Cannabis use and mental health in young people: cohort study. *British Medical Journal*. 2002;325(7374):1195-8.
12. World Health Organisation. Evidence-based strategies and interventions to reduce alcohol - related harm: Global assessment of public -health problems caused by harmful use of alcohol. Geneva: World Health Organisation; 2007.
13. Beinart S, Anderson B, Lee S, Utting D. Youth at Risk?: A National Survey of Risk Factors, Protective Factors and Problem Behaviour among Young People in England, Scotland and Wales (JRF Findings 432). York: Joseph Rowntree Foundation; 2002.
14. Jansen DE, Veenstra R, Ormel J, Verhulst FC, Reijneveld SA. Early risk factors for being a bully, victim, or bully/victim in late elementary and early secondary education: the longitudinal TRAILS study. *BMC Public Health*. 2011;11((1):440.

15. Arseneault L, Walsh E, Trzesniewski K, Newcombe R, Caspi A, Moffitt T. Bullying victimization uniquely contributes to adjustment problems in young children: a nationally representative cohort study. *Pediatrics*. 2006;118(1):130-8.
16. Forero R, McLellan L, Rissel C, Bauman A. Bullying behaviour and psychosocial health among school students in New South Wales, Australia: cross sectional survey. *British Medical Journal*. 1999;319(7206):344-8.
17. Kaltiala-Heino R, Rimpelä M, Rantanen P, Rimpelä A. Bullying at school-an indicator of adolescents at risk for mental disorders. *Journal of Adolescence*. 2000;23(6):661-74.
18. Juvonen J, Graham S, Schuster MA. Bullying among young adolescents: the strong, the weak, and the troubled. *Pediatrics*. 2003;112(6):1231-7.
19. Bond L, Carlin JB, Thomas L, Rubin K, Patton G. Does bullying cause emotional problems? A retrospective study of young teenagers. *British Medical Journal*. 2001;323:480-84.
20. Hawker DS, Boulton MJ. Twenty years' research on peer victimization and psychosocial maladjustment: a meta-analytic review of cross-sectional studies. *Journal of Child Psychology and Psychiatry*. 2000;41(4):441-55.
21. Van der Wal MF, De Wit CA, Hirasing RA. Psychosocial health among young victims and offenders of direct and indirect bullying. *Pediatrics*. 2003;111:1312-7.
22. Sainsbury Centre for Mental Health. *Diversion: a better way for criminal justice and mental health*. London: Sainsbury Centre for Mental Health.; 2009.
23. Scott S, Knapp M, Henderson J, Maughan B. Financial cost of social exclusion: follow up study of antisocial children into adulthood. *BMJ*. 2001;323(7306):191.
24. Catalano RF, Berglund LM, Ryan JAM, Lonczak HS, Hawkins JD. Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Programs. *Prevention and Treatment*. 2002;5(1):1-166.
25. Jackson C, Sweeting H, Haw S. Clustering of substance use and sexual risk behaviour in adolescence: analysis of two cohort studies. *British Medical Journal Open*. 2012;2(e000661).
26. Buck D. *Clustering of unhealthy behaviours over time*. London: The King's Fund; 2012.
27. Kipping RR, Campbell RM, MacArthur GJ, Gunnell DJ, M. H. Multiple risk behaviour in adolescence. *Journal of Public Health*. 2012;34(s1): i1-i2.
28. National Youth Agency. *The NYA Guide to Youth Work in England*. Leicester: National Youth Agency; 2007.
29. Schulman S, Davies T. *Evidence of the impact of the 'youth development model' on outcomes for young people - a literature review*. Leicester: National Youth Agency; 2007.
30. Roth J, Brooks-Gunn J. Promoting healthy adolescents: synthesis of youth development program evaluations. *Journal of Research on Adolescence*. 1998;8(4):423-59.
31. Durlak JA, Dymnicki AB, Taylor RD, Weissberg RP, Schellinger KB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development*. 2011;82(1):405-32.

32. Hummel S, Naylor P, Chilcott J, Guillaume L, Wilkinson A, Blank L. Cost effectiveness of universal interventions which aim to promote emotional and social wellbeing in secondary schools: report for NICE Sheffield: University of Sheffield; 2009.
33. Philiber S, Kaye JW, Herrling S. The National Evaluation of the Children's Aid Society Carrera Model Program to Prevent Teen Pregnancy. New York: Philiber Research Associations; 2001.
34. Department for Education. Positive for Youth - A new approach to cross-government policy for young people aged 13 to 19. London: Stationery Office; 2011.
35. Department of Health. Healthy Lives, Healthy People. Our Strategy for Public Health in England. London: HM Government; 2010.
36. Mayor's Fund For London. Mayor's Fund for London: Annual Report 2011. London: London Mayor; 2011.
37. Scottish Government. Valuing Young People: Principles and connections to support young people achieve their potential. Edinburgh: Scottish Government.; 2009.
38. Welsh Assembly Government. The Youth Work Curriculum Statement for Wales. Cardiff: Welsh Assembly Government; 2007.
39. Durlak JA, Weissberg RP. The impact of after-school programs that promote personal and social skills. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning; 2007.
40. Gavin LE, Catalano RF, David-Ferdon C, Gloppen KM, Markham CM. A review of positive youth development programs that promote adolescent sexual and reproductive health. *J Adolesc Health*. 2010;46(3 Suppl):S75-91.
41. Harden A, Brunton G, Fletcher A, Oakley A, Burchett H, Backhans M. Young people, pregnancy and social exclusion: A systematic synthesis of research evidence to identify effective, appropriate and promising approaches for prevention and support. London: Institute of Education; 2006.
42. Wiggins M, Bonell C, Sawtell M, Austerberry H, Burchett H, Allen E, et al. Health outcomes of youth development programme in England: prospective matched comparison study. *BMJ*. 2009;339
43. Bonell C, Maisey R, Speight S, Purdon S, Keogh P, Wollny I, et al. Randomized controlled trial of 'Teens and Toddlers': a teenage pregnancy prevention intervention combining youth development and voluntary service in a nursery. *Journal of Adolescence* (in press).
44. Dishion TJ, McCord J, Poulin F. When interventions harm. *American Psychologist*. 1999;54(9):755-64.
45. Weiss B, Caron A, Ball S, Tapp J, Johnson M, Weisz JR. Iatrogenic effects of group treatment for anti-social youth. *Journal of Consulting and Clinical Psychology*. 2005;73:1036-44.
46. Roth JL, Brooks-Gunn J. Youth development programs: risk, prevention and policy. *Journal of Adolescent Health*. 2003;32:170-82.

47. Gaglio B, Glasgow RE. Evaluation approaches for dissemination and implementation research. In: Brownson RC, Colditz G, Proctor E, editors. *Dissemination and Implementation Research in Health: Translating Science to Practice*: Oxford University Press; 2012.
48. Faggiano F, Vigna-Taglianti FD, Versino E, Zambon A, Borraccino A, Lemma P. School-based prevention for illicit drugs use. *Cochrane Database of Systematic Reviews Issue 2, Art No CD003020 DOI@: 101002/14651858 CD 003020 pub2*. 2005.
49. Foxcroft DR, Ireland D, Lowe G, Breen R. Primary prevention for alcohol misuse in young people. *Cochrane Database of Systematic Reviews Issue 2, Art No CD003020 DOI@: 101002/14651858 CD 003020 pub2*. 2002;3:Art. No. CD003024.
50. Hahn R, Fuqua-Whitley D, Wethington H, Lowy J, Crosby A, Fullilove M. Effectiveness of universal school-based programs to prevent violent and aggressive behavior: a systematic review. *Am J Prev Med*. 2007;33(2 Suppl):S114-S29.
51. Thomas R, Perera R. School-based programmes for preventing smoking (Cochrane Review). *Cochrane Database of Systematic Reviews 3: CD001293*. DOI:10.1002/14651858.CD001293. 2006.
52. Bonell C, Jamal F, Harden A, Wells H, Parry W, Fletcher A, et al. Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. *Public Health Research*. 2013;1(1).
53. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. Oxford: The Cochrane Collaboration; 2011.
54. Peersman G, Oliver S, Oakley A. *EPPI-Center Review Guidelines: Data Collection for the EPIC Database*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London; 1997.
55. Egan M, Bambra C, Petticrew M, Whitehead M. Reviewing evidence on complex social interventions: development and testing of a new tool for appraising implementation. *Journal of Epidemiology and Community Health*. 2009;63:4-11.
56. Shepherd J, Harden A, Rees R, Brunton G, Garcia J, Oliver S, et al. *Young People and Healthy Eating: A systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit; 2001.
57. Drummond MF, Sculpher MJ, Torrance GW, O'Brien B, Stoddart GL. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University Press; 2005.
58. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technology Assessment* 2004;8(36).
59. Miners AH, Yao G, Raftery J, Taylor RS. Economic evaluations of calcineurin inhibitors in renal transplantation: A literature review. *Pharmacoeconomics*. 2007;25(11):935-47.
60. Sterne J, Egger M, Moher D. Addressing reporting biases. In: Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: Wiley; 2008. p. 297-334.

61. Cho BK, Rosenfeldt F, Turina MI, Karp RB, Ferguson TB, Bodnar E, et al. Joint statement on redundant (duplicate) publication by the Editors of the undersigned cardiothoracic journals. *Ann Thorac Surg*. 2000;69:663.
62. Arai L, Roen K, Roberts H, Popay J. It might work in Oklahoma but will it work in Oakhampton? Context and implementation in the effectiveness literature on domestic smoke detectors. *Inj Prev*. 2005;11:148-51.
63. Noyes J, Popay J, Garner P. What can qualitative research contribute to a Cochrane systematic review of DOT for promoting adherence to tuberculosis treatment? *Qualitative Research and Systematic Reviews workshop*; 2005 28-29 June; Continuing Professional Development Centre, University of Oxford. 2005.
64. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*. 2008;8:45.
65. Kalaian HA, Raudenbush SW. A multivariate mixed linear model for meta-analysis. *Psychological Methods*. 1996;1:227-35.
66. NICE. *NICE Guide to the Methods of Technology Appraisal*. London: National Institute of Health and Clinical Excellence; 2008.
67. Thompson S, Sharp S. Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine*. 1999;18(2):693-708.
68. Borenstein M, Hedges L, Higgins J, Rothstein H. *Introduction to Meta-Analysis*. Chichester: John Wiley and Sons Ltd; 2009.
69. Ragin CC. *Redesigning Social Inquiry: Fuzzy Sets and Beyond*. Chicago, IL: University of Chicago Press; 2008.
70. Thomas J, Brunton J, Graziosi S. *EPPI-Reviewer 4.0: software for research synthesis*. EPPI-Centre software. London: Social Science Research Unit, Institute of Education; 2010 Contract No.: Document Number].
71. Ragin CC, Drass KA, Davey S. *Fuzzy-Set / Qualitative Comparative Analysis 2.0*. Tucson AZ: Department of Sociology, University of Arizona; 2006.
72. Kavanagh J, Oliver S, Lorenc T, Caird J, Tucker H, Greaves A, et al. School-based cognitive behavioural interventions: a systematic review of effects and inequalities. *Health Sociology Review* 2009;18:61-78.
73. Social Research Association. *Ethical Guidelines*. London: Social Research Association; 2003.
74. Centre for Social Justice and Community Action and National Co-ordinating Centre for Public Engagement. *Community-based participatory research: a guide to ethical principles and practice, national coordinating centre for public engagement*. Bristol: National Coordinating Centre for Public Engagement; 2012.

Appendix 2 Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist

TABLE 11 Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist

Section/topic	Checklist item	Reported
Title		
<i>Title</i>	1 Identify the report as a systematic review, meta-analysis, or both	<i>Title page</i>
Abstract		
<i>Structured summary</i>	2 Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number	<i>Abstract</i>
Introduction		
<i>Rationale</i>	3 Describe the rationale for the review in the context of what is already known	<i>Chapter 1, Rationale for this review</i>
<i>Objectives</i>	4 Provide an explicit statement of questions being addressed with reference to PICOS	<i>Chapter 1, Review aims and objectives</i>
Methods		
<i>Protocol and registration</i>	5 Indicate if a review protocol exists, if and where it can be accessed (e.g. web address), and, if available, provide registration information including registration number	<i>Chapter 2, About this chapter and Appendix 1</i>
<i>Eligibility criteria</i>	6 Specify study characteristics (e.g. PICOS, length of follow-up) and report characteristics (e.g. years considered, language, publication status) used as criteria for eligibility, giving rationale	<i>Chapter 2, Criteria for considering studies for this review and Appendix 3 exclusion criteria and guidance</i>
<i>Information sources</i>	7 Describe all information sources (e.g. databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	<i>Chapter 2, Search strategy</i>
<i>Search</i>	8 Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	<i>Appendix 4</i>
<i>Study selection</i>	9 State the process for selecting studies (i.e. screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	<i>Chapter 2, Study selection</i>
<i>Data collection process</i>	10 Describe method of data extraction from reports (e.g. piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	<i>Chapter 2, Data extraction</i>
<i>Data items</i>	11 List and define all variables for which data were sought (e.g. PICOS, funding sources) and any assumptions and simplifications made	<i>Chapter 2, Synthesis of results</i>
<i>Risk of bias in individual studies</i>	12 Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level and how this information is to be used in any data synthesis)	<i>Chapter 2, Synthesis of results</i>

continued

TABLE 11 Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist (continued)

Section/topic	Checklist item	Reported
<i>Summary measures</i>	13 State the principal summary measures (e.g. risk ratio, difference in means)	<i>Chapter 2, Synthesis of results</i>
<i>Synthesis of results</i>	14 Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g. I^2) for each meta-analysis	<i>Chapter 2, Synthesis of results</i>
<i>Risk of bias across studies</i>	15 Specify any assessment of risk of bias that may affect the cumulative evidence (e.g. publication bias, selective reporting within studies)	<i>Chapter 2, Quality assessment</i>
<i>Additional analyses</i>	16 Describe methods of additional analyses (e.g. sensitivity or subgroup analyses, metaregression), if done, indicating which were pre-specified	<i>Chapter 2, Synthesis of results</i>
Results		
<i>Study selection</i>	17 Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram	<i>Chapter 3, Results of the search</i>
<i>Study characteristics</i>	18 For each study, present characteristics for which data were extracted (e.g. study size, PICOS, follow-up period) and provide the citations	<i>Chapter 3, Study characteristics; Appendix 13; Appendix 14; and Appendix 16</i>
<i>Risk of bias within studies</i>	19 Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12)	<i>Chapter 6, Risk of bias and quality of evidence and Appendix 16</i>
<i>Results of individual studies</i>	20 For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and CIs, ideally with a forest plot	<i>Chapter 6, Effects of interventions on substance use and Chapter 6, effects of interventions on violence</i>
<i>Synthesis of results</i>	21 Present results of each meta-analysis done, including CIs and measures of consistency	<i>Chapter 6, Effects of interventions on substance use and Chapter 6, effects of interventions on violence</i>
<i>Risk of bias across studies</i>	22 Present results of any assessment of risk of bias across studies (see item 15)	<i>Chapter 6, Risk of bias and quality of evidence and Appendix 16</i>
<i>Additional analysis</i>	23 Give results of additional analyses, if done (e.g. sensitivity or subgroup analyses, metaregression; see item 16)	<i>Chapter 6, Effects of interventions on substance use and Chapter 6, Effects of interventions on violence</i>
Discussion		
<i>Summary of evidence</i>	24 Summarise the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g. health-care providers and users and policy-makers)	<i>Chapter 7, Key results</i>
<i>Limitations</i>	25 Discuss limitations at study and outcome level (e.g. risk of bias) and at review level (e.g. incomplete retrieval of identified research, reporting bias)	<i>Chapter 7, Review limitations</i>
<i>Conclusions</i>	26 Provide a general interpretation of the results in the context of other evidence and implications for future research	<i>Chapter 7, Implications</i>
Funding		
<i>Funding</i>	27 Describe sources of funding for the systematic review and other support (e.g. supply of data; role of funders for the systematic review)	<i>Acknowledgements</i>
PICOS, participants, interventions, comparisons, outcomes and study design.		

Appendix 3 Exclusion criteria and guidance

TABLE 12 Exclusion criteria and guidance

Exclusion criteria	Guidance
1. Date Exclude studies published before 1985	However, if date is not an available filter on databases/websites, please apply this criterion first
2. Language Exclude studies not published in English	Studies are not excluded on geographical location (e.g. all high-, middle- and low-income countries are relevant). However, studies not available in English are excluded
3. Population: Exclude studies where:	We are not interested in studies in which CYP
<ul style="list-style-type: none"> the majority of the population are not enrolled in full-time education 	<ul style="list-style-type: none"> are enrolled in a school or educational facility (e.g. 16-year-olds who are in employment or juvenile detention centres, or unemployed teenage parents)
AND/OR	AND/OR
<ul style="list-style-type: none"> the mean sample of participants does not fall within the age range of 11–18 years 	<ul style="list-style-type: none"> the majority of the population are within the bounds of the age range, that is, 11 years of age or younger (e.g. 6- to 12-year-olds) or 18 years of age or over (e.g. 18- to 21-year-olds)
AND/OR	AND/OR
<ul style="list-style-type: none"> CYP are diagnosed with pre-existing medical conditions 	<ul style="list-style-type: none"> have been diagnosed with a pre-existing medical condition [e.g. physical illnesses such as diabetes, cancer, juvenile arthritis OR mental health conditions (e.g. eating disorders, mood disorders, OCD or behavioural diagnosis such as conduct disorder, ADHD, ODD)]
4. Intervention Exclude studies that:	Studies may refer to programmes as ‘positive youth development’ and others may be called ‘youth development programs’ – which do or do not solely aim to prevent problems and do or do not address positive adolescent development according to the definition provided
<ul style="list-style-type: none"> focus only on the prevention of problem behaviours 	Setting:
OR	programmes may be identified by their setting (e.g. ‘youth clubs, ASPs, ‘drop-in centres’ which include a range of activities with the aim of promoting positive development as defined)
<ul style="list-style-type: none"> promote only ONE development goal in ONE domain 	Activity-based:
AND	programmes may not be identifiable by the term ‘PYD’ or setting but focus on a single activity, for example:
<ul style="list-style-type: none"> are delivered in school time OR criminal justice OR clinical settings 	<ul style="list-style-type: none"> sports-based programmes promoting ‘bonding’ and ‘identity’ after school chess clubs promoting ‘cognitive skills’ and ‘bonding’ eco-conscious programme providing ‘recognition for pro-social involvement’ and ‘beliefs in the future’
AND/OR	Other PYD labels:
<ul style="list-style-type: none"> targets parents/carers (alongside CYP) in order to focus on family functioning 	programmes may be called service learning, or focus on civic engagement and take a PYD approach without specifically using those terms (e.g. improving community engagement via school-volunteer action programmes)

continued

TABLE 12 Exclusion criteria and guidance (continued)

Exclusion criteria	Guidance
	<p>Studies excluded because they are delivered</p> <ul style="list-style-type: none"> in schools: refer to programmes conducted during normal school hours, not after school criminal justice settings (e.g. prisons, probation centres or police stations) clinical/health-care settings (e.g. hospitals, primary care, clinical psychology departments or out-patient clients)
<p>5. Type of studies</p> <p>We will exclude studies that do not:</p> <ul style="list-style-type: none"> RQ1: describe PYD intervention theory of change RQ2: report a process evaluation of PYD intervention RQ3: report outcome and economic evaluations of PYD interventions using the following research designs: <ul style="list-style-type: none"> experimental (RCTs) <p>OR</p> <ul style="list-style-type: none"> quasi-experimental studies (employing non-randomised prospective comparison groups) with control groups which receive usual care or no treatment 	<p>We are not interested in:</p> <ul style="list-style-type: none"> correlational studies or studies 'exploring the association between factors and outcomes (e.g. studies investigating the relationships between participating in certain PYD activities and whether this led to an increase in particular social and developmental outcomes using linear regression methods. They are not evaluating process or outcome. They may also have a predictive analysis to identify what factors might contribute to the success or failure of PYD programmes for particular individuals) methodological studies (e.g. which types of methods of evaluating PYD are most appropriate, or investigation into how well particular tools captures PYD or other outcomes) 'views' studies not part of a formal process evaluation (e.g. collecting data on young people's experience separate from evaluating specific implementation processes) surveys of PYD coverage (e.g. the number of programmes being delivered in particular geographical areas) policy documents: on PYD interventions not an empirical study opinion/think pieces (e.g. the value of providing PYD programmes for CYP)
<p>6. Reporting data: outcomes/Process</p> <p>Exclude studies which do not:</p> <ul style="list-style-type: none"> address substance use <u>or</u> violence as part of the PYD programme <p>AND do not:</p> <ul style="list-style-type: none"> report the impact of a PYD intervention on substance use or violence as a behavioural outcome <p>OR do not:</p> <ul style="list-style-type: none"> collect and report data on the process of delivering interventions 	<p>All study types</p> <p>The PYD programme being investigated MUST address substance use and/or violence</p> <p>Outcome evaluations must report data on the substance use or violence at a behavioural level. Studies that measure ONLY variables considered to mediate behaviour change (e.g. knowledge, attitudes, beliefs, intentions) are excluded if they do not also report/measure a behavioural outcome</p> <p>Process evaluations: must collect and report data on the process of delivering PYD interventions (e.g. implementation procedures, accessibility and acceptability of programmes, the content, materials provided, etc.)</p>
<p>ADHD, attention deficit and hyperactivity disorder; CYP, children and young people; OCD, obsessive-compulsive disorder; ODD, oppositional defiant disorder.</p>	

Appendix 4 Search strategies for bibliographic databases

PsycINFO (EBSCOhost)

Search date: 7 November 2013.

Searcher: CS.

Search modes: Boolean/Phrase.

Number of records: 8227.

#	Query	Results
S1	TI ("Young people#" OR "young person#" OR "young offender#" OR adolescent# OR adolescence OR youth# OR minors OR teen OR teens OR teenage OR teenaged OR teenager# OR juvenile# OR pupil# OR boy# OR girl# OR underage# OR (school AND dropout#) OR (school AND "drop out#") OR "school aged") OR AB ("Young people#" OR "young person#" OR "young offender#" OR adolescent# OR adolescence OR youth# OR minors OR teen OR teens OR teenage OR teenaged OR teenager# OR juvenile# OR pupil# OR boy# OR girl# OR underage# OR (school AND dropout#) OR (school AND "drop out#") OR "school aged")	311,506
S2	TI "youth opportunit*" OR AB "youth opportunit*"	36
S3	TI "youth work*" OR AB "youth work*"	413
S4	TI "youth club#" OR AB "youth club#"	74
S5	TI "youth centre#" OR AB "youth centre#"	38
S6	TI "youth center#" OR AB "youth center#"	145
S7	TI (youth# N2 empower*) OR AB (youth# N2 empower*)	185
S8	TI (adolescent# N2 empower*) OR AB (adolescent# N2 empower*)	112
S9	TI ("young people#" N3 empower*) OR AB ("young people#" N3 empower*)	54
S10	TI ("youth led") OR AB ("youth led")	51
S11	TI youth N1 voice# OR AB youth N1 voice#	85
S12	TI youth N1 advoca* OR AB youth N1 advoca*	78
S13	TI (youth# N1 engagement) OR TI ("engaging youth#") OR AB (youth# N1 engagement) OR AB ("engaging youth#")	249
S14	TI (adolescent# N1 engagement) OR TI ("engaging adolescent#") OR AB (adolescent# N1 engagement) OR AB ("engaging adolescent#")	286
S15	TI ("young people#" N1 (engagement OR engaging)) OR AB ("young people#" N1 (engagement OR engaging))	92
S16	TI (youth# N1 involvement) OR TI ("involving youth#") OR AB (youth# N1 involvement) OR AB ("involving youth#")	402
S17	TI ("young people#" N1 (involving OR involvement)) OR AB ("young people#" N1 (involving OR involvement))	111
S18	TI "youth participation" OR AB "youth participation"	191
S19	TI "adolescent participation" OR AB "adolescent participation"	85
S21	TI "youth program*" OR AB "youth program*"	306

#	Query	Results
S22	TI youth N5 asset# OR AB youth N5 asset#	178
S23	TI youth N12 asset# OR AB youth N12 asset#	245
S24	TI "young people#" N12 asset# OR AB "young people#" N12 asset#	32
S25	TI adolescent# N12 asset# OR AB adolescent# N12 asset#	153
S26	AB ("positive development") N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#)	174
S27	TI "extra curricular" OR AB "extra curricular"	419
S28	TI supplement* N1 education OR AB supplement* N1 education	98
S29	TI (nonformal* N1 education) OR AB (nonformal* N1 education) OR TI ("non formal*" N1 education) OR AB ("non formal*" N1 education)	92
S30	TI informal* N1 education OR AB informal* N1 education	281
S31	TI "additional education" OR AB "additional education"	144
S32	TI (community N1 empower*) OR AB (community N1 empower*)	512
S33	TI "civic engagement" OR AB "civic engagement"	744
S34	TI "Asset# building" OR AB "Asset# building"	87
S35	TI "Developmental asset#" OR AB "Developmental asset#"	172
S36	TI "Psychosocial asset#" OR AB "Psychosocial asset#"	24
S37	TI "peer led" OR AB "peer led"	414
S38	TI "peer engagement" OR AB "peer engagement"	15
S39	TI service N1 learning OR AB service N1 learning	1601
S40	TI afterschool OR AB afterschool	184
S41	TI ("after school" N5 (program# OR club# OR initiative# OR scheme# OR center# OR centre# OR event# OR intervention# OR service# OR project# OR service# OR session#)) OR AB ("after school" N5 (program# OR club# OR initiative# OR scheme# OR center# OR centre# OR event# OR intervention# OR service# OR project# OR service# OR session#))	957
S42	S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34 OR S35 OR S36 OR S37 OR S38 OR S39 OR S40 OR S41	5553
S43	TI (mentor* N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#)) OR AB (mentor* N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#))	826
S44	TI (coach* N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#)) OR AB (coach* N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#))	523
S45	TI ("development program*" N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#)) OR AB ("development program*" N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#))	344
S46	TI ((community N1 service) N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#)) OR AB ((community N1 service) N12 ("young people#" OR adolescen* OR youth# OR minors OR teen* OR juvenile* OR pupil* OR boy# OR girl# OR "school aged" OR "drop out#" OR dropout# OR underage#))	327
S47	DE "Social Programs" OR DE "Mentor" OR DE "Coaching"	6995

#	Query	Results
S48	TI "positive youth" OR AB "positive youth"	810
S49	TI "youth development" OR AB "youth development"	1750
S50	S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S43 OR S44 OR S45 OR S46 OR S48 OR S49	6245
S51	S42 OR S47	12,467
S52	S51 AND S1	2564
S53	S50 OR S52	7286
	Limiters - Publication Year: 1985-2013 Search modes - Boolean/Phrase	
S54	S51	2516
	Limiters - Publication Year: 1985-2013; Age Groups: School Age (6-12 yrs), Adolescence (13-17 yrs)	
S55	S54 OR S53	8227

Notes

#, wildcard of 1 or 0 characters; *, truncation; N1, words within one place of each other in any order.

Appendix 5 Data extraction tool for theory studies

1. Aim of paper

Guidance: one sentence summing up the aim of paper.

2. Does the paper offer a theory of change for PYD interventions?

Guidance: by 'theory of change', we mean a description of the intended core components of an intervention, how these interact to produce changes and anticipated short-, medium- and long-term outcomes.³ 'PYD' is as defined as per our inclusion criteria.

- i. YES
- ii. NO

Guidance: if no then cease data extraction and discuss if this paper should in fact be excluded from the review

3. Does the paper cover mostly the same ground as papers we have already reviewed?

- i. YES

Guidance: if yes then only extract novel information and elsewhere simply cite already reviewed papers covering the same ground

- i. NO

4. How does the paper describe PYD?

Guidance: as general approach/specific intervention; specify

5. Which theories are used to inform PYD theory of change?

- i. None mentioned
- ii. Name and author of theory

Guidance: Please just give this info – we do not need a description of the theory.

6. Does the study describe how PYD is intended to act on the individual? Guidance Action on individual – in other words, what are the proximal pathways via which intervention components are intended to affect the individual? Please provide one brief paragraph.

- i. YES (one brief paragraph of summary)

Guidance: specify mechanisms

- i. NO

7. Does the study describe how PYD is intended to act on the individual's environment?

Guidance: action on environment – in other words what are the proximal pathways via which intervention components are intended to affect aspects of the individual's environment such as families, peer groups, schools and other institutions, local communities and other levels?

i. YES (one brief paragraph of summary)

Guidance: specify levels/mechanisms

i. NO

8. Does the study describe how PYD is intended to reduce risk of substance misuse/violence?

Guidance: risk reduction – in other words what are the more distal pathways via which the intervention's effects on individuals/environments are intended to reduce risk of smoking, alcohol and drug use and violence perpetration or victimisation.

i. YES (one brief paragraph of summary)

Guidance: specify mechanisms

i. NO

9. Any other comments

Guidance: write here any points not covered above (e.g. your view of the novelty and importance of this paper)

10. References checked

i. Checked: none

ii. Checked: follow-up

Appendix 6 Data extraction tool for extracting process information

1. Study location (country, area of country)
 - i. Not stated
 - ii. Details
2. The nature of the intervention(s)/components and how it was delivered to intervention group(s)
 - i. Not stated
 - ii. Details
3. How intervention was developed

Guidance (e.g. background to the programme; founder, etc.)

- i. Not stated
 - ii. Details
4. Timing of intervention
 - i. Not stated
 - ii. Details
5. Provider organisation description
 - i. Not stated
 - ii. Details
6. Provider staff characteristics/qualifications/experience
 - i. Not stated
 - ii. Details

7. Target population

Guidance: target of the programme

- i. Not stated
 - ii. Details

8. Timing of process evaluation

Guidance: prospective or retrospective

- i. Not stated
 - ii. Details

9. Sampling methods
 - i. Not stated
 - ii. Details
10. Sample size, response rates at baseline
 - i. Not stated
 - ii. Details
11. Sample size, response rates at follow-up
 - Not stated
 - Details
 - i. Not appropriate
12. Sociodemographic characteristics of participants at baseline/follow-up
 - i. Not stated
 - ii. Details
13. Methods of data collection
 - i. Not stated
 - ii. Details
14. Methods of data analysis
 - i. Not stated
 - ii. Details
15. Data/findings on contextual facilitators/barriers to intervention processes (feasibility, fidelity, quality, coverage/accessibility, acceptability/satisfaction, intensity/dose, mechanism)

Guidance: how; setting; characteristic of the population

- i. Not stated
- ii. Details

Appendix 7 Data extraction tool for outcome evaluations

1. Study location (country, area of country)
 - i. Not stated
 - ii. Details
2. The nature of the intervention(s)/components and how it was delivered to intervention group(s)
 - i. Not stated
 - ii. Details
3. The nature of the intervention(s)/components delivered to control group(s)
 - i. Waitlist/delayed treatment
 - ii. Attention placebo/alternative intervention (please specify)

Use if the comparison group receives a different intervention to the treatment group that is not the same as usual care and which has different aims to the main intervention
 - i. Usual treatment/care, with assignment
 - ii. Matched group from target population or other inactive, without assignment
4. How intervention was developed
 - i. Not stated
 - ii. Details
5. Timing of interventions
 - i. Not stated
 - ii. Details
6. Provider organisation description
 - i. Not stated
 - ii. Details
7. Target population
 - i. Not stated
 - ii. Details
8. Timing of outcome evaluation
 - i. Not stated
 - ii. Details

9. Outcome evaluation study design
 - i. RCT
 - ii. nRCT
10. Unit of allocation
 - i. Individual
 - ii. Other (please specify)
11. Generation of allocation sequence: any stratification, minimisation, etc.?
 - i. Not applicable
 - ii. Not stated
 - iii. No
 - iv. Yes

Guidance: for example, the investigators describe a random component in the sequence generation process such as: referring to a random number table; using a computer random number generator; coin tossing; shuffling cards or envelopes; throwing dice; drawing of lots; minimisation.

Note: minimisation may be implemented without a random element and this is considered equivalent to being random.

12. Concealment of allocation (provide details)
 - i. Not applicable
 - ii. Not stated
 - iii. No
 - iv. Yes

Guidance: for example, participants and investigators enrolling participants could not foresee assignment because one of the following, or an equivalent method, was used to conceal allocation: central allocation (including telephone, web-based and pharmacy-controlled randomisation); sequentially numbered drug containers of identical appearance; sequentially numbered, opaque, sealed envelopes.

13. Blinding of intervention provider, outcome assessor
 - i. Not stated
 - ii. No
 - iii. Yes

14. Sample size, overall response rates at baseline
 - i. Not stated
 - ii. Details

15. Sample size, overall response rates at follow-up
 - i. Not stated
 - ii. Details

16. Sociodemographic characteristics at baseline/follow-up
 - i. Not stated
 - ii. Details
17. Were baseline equivalence/differences between arms reported?
 - i. Yes
 - ii. No
18. How were differences between intervention and comparison groups controlled?
 - i. Not applicable (e.g. RCT)
 - ii. Matching (please specify)
 - iii. Adjustment (please specify)
 - iv. Not controlled
 - v. Not stated/Not clear (please specify)
19. Outcome measures (1): for each one answer the following
 - i. Description
 - ii. Pre-hypothesisation
 - Primary outcome
 - Secondary outcome
 - Other
 - No
20. Evidence of reliability/validity (provide details)
 - i. Yes
 - ii. No
21. Data collection methods
 - i. Not stated
 - ii. Details
22. Baseline response rate
 - i. Not stated
 - ii. Details
23. Follow-up response rate
 - i. Not stated
 - ii. Details
24. Rates of outcome by arm at follow-up (this is for each outcome, by intervention group and control group – differing from overall follow-up rates)
 - i. Not stated
 - ii. Details

25. Effect sizes

- i. Overall

Guidance: if multiple choose ITT, adjusted analysis accounting for any clustering

- by gender

Guidance: if multiple choose ITT, adjusted analysis accounting for any clustering

- i. by age

Guidance: if multiple choose ITT, adjusted analysis accounting for any clustering

- i. by SES

Guidance: if multiple choose ITT, adjusted analysis accounting for any clustering

- i. by ethnic subgroup

Guidance: if multiple choose ITT, adjusted analysis accounting for any clustering

- i. No effect size?

26. Study analysis was intention-to-treat?

- i. Not stated
- ii. No
- iii. Yes

27. Study analysis appropriately accounted for clustering

- i. Not stated
- ii. No
- iii. Yes

28. Study analysis adjusted for confounders

- i. Not stated
- ii. No
- iii. Yes

Appendix 8 Author contact for missing data

Dear X,

I am a researcher from the EPPI-Centre (<http://eppi.ioe.ac.uk/cms/>). We are currently undertaking a systematic review on 'The effects of positive youth development interventions on substance use, violence and inequalities'.

Please find the protocol here: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013005439

I am writing to you today as the **lead author** of the following study included in our review:

Add individual study details

to **request further information** that we could not answer during our data extraction process.

If possible, could you provide data for the following unanswered questions collated in the attached excel sheet. Ideally we would appreciate if you could return this form by (date)

However, if this is not possible, please could you indicate if and when you would be able to respond by?

If you have any questions, please do not hesitate to get in touch. I have attached the full data extraction form for you to get a sense of the type of information we are extracting for all studies in the review.

Thank you in advance for your assistance on this matter.

Kind regards

Kelly Dickson

Appendix 9 Theory studies: coding templates

BOX 1 NVivo Node Structure Report: Kate Hinds

Hierarchical name

Definition of PYD interventions

Definition of PYD interventions\Aspects of interventions

Definition of PYD interventions\Aspects of interventions\Age-related progression.

Definition of PYD interventions\Aspects of interventions\Developmental outcomes and related activities.

Definition of PYD interventions\Aspects of interventions\Interaction between activities and context.

Definition of PYD interventions\Aspects of interventions\Programme atmosphere.

Definition of PYD interventions\Aspects of interventions\Programme characteristics.

Definition of PYD interventions\Aspects of interventions\Programme characteristics\Multilevel, feedback loops, negotiation.

Definition of PYD interventions\Aspects of interventions\Structure of activities.

Definition of PYD interventions\Aspects of interventions\Structure of activities\Breadth of activities.

Definition of PYD interventions\Aspects of interventions\Structure of activities\Intensity.

Definition of PYD interventions\Aspects of interventions\Targeting vs. universal.

Definition of PYD interventions\Comparisons and integrations

Definition of PYD interventions\Comparisons and integrations\Comparison with deficit model.

Definition of PYD interventions\Comparisons and integrations\Comparison with traditional youth programmes.

Definition of PYD interventions\Comparisons and integrations\Integration between PYD and resilience.

Definition of PYD interventions\Comparisons and integrations\Prevention programmes.

Definition of PYD interventions\Comparisons and integrations\PYD vs. prevention interventions.

Definition of PYD interventions\Comparisons and integrations\Relationship between protective factors and developmental assets.

Definition of PYD interventions\Comparisons and integrations\Relationship between protective factors and developmental assets\Action of risk and protective factors.

BOX 1 NVivo Node Structure Report: Kate Hinds (*continued*)

Definition of PYD interventions\Comparisons and integrations\Relationship between risk-protection approach and applied developmental science.

Definition of PYD interventions\Comparisons and integrations\Treatment programmes.

Definition of PYD interventions\Components of PYD

Definition of PYD interventions\Components of PYD\Asset-building community.

Definition of PYD interventions\Components of PYD\Asset-building society.

Definition of PYD interventions\Components of PYD\Developmental assets.

Definition of PYD interventions\Components of PYD\Developmental assets\sixth C: contribution.

Definition of PYD interventions\Components of PYD\Developmental assets\sixth C: contribution\Contribution as link between individual and context.

Definition of PYD interventions\Components of PYD\Developmental assets\Connectedness.

Definition of PYD interventions\Components of PYD\Developmental assets\Thriving.

Definition of PYD interventions\Components of PYD\Developmental assets\Thriving\Self-regulation.

Definition of PYD interventions\Components of PYD\Domains.

Definition of PYD interventions\Components of PYD\Domains\Behavioural domain.

Definition of PYD interventions\Components of PYD\Domains\Cognitive domain.

Definition of PYD interventions\Components of PYD\Domains\Educational domain.

Definition of PYD interventions\Components of PYD\Domains\Emotional domain.

Definition of PYD interventions\Components of PYD\Domains\Moral domain.

Definition of PYD interventions\Components of PYD\Domains\Physiological domain.

Definition of PYD interventions\Components of PYD\Domains\Social domain.

Definition of PYD interventions\Components of PYD\Domains\Structural domain.

Definition of PYD interventions\Components of PYD\notion of hope.

Definition of PYD interventions\Components of PYD\notion of optimism.

BOX 1 NVivo Node Structure Report: Kate Hinds (*continued*)*Definition of PYD interventions\Developmental science*

Definition of PYD interventions\Developmental Science\Associations in behavioural science.

Definition of PYD interventions\Developmental Science\Definitions.

Definition of PYD interventions\Developmental Science\Definitions\Thriving.

Definition of PYD interventions\Developmental Science\Developmental affordances.

Definition of PYD interventions\Developmental Science\Identity development.

Definition of PYD interventions\Developmental Science\Plasticity.

Definition of PYD interventions\Developmental Science\Resiliency.

Definition of PYD interventions\Developmental Science\Resiliency\Protective factors.

Definition of PYD interventions\Developmental Science\Role of developmental assets in developmental science.

Definition of PYD interventions\Developmental Science\Selection, optimisation and compensation.

Definition of PYD interventions\Developmental Science\System theory.

Definition of PYD interventions\Developmental Science\System theory\Cultural-ecological framework.

Definition of PYD interventions\Developmental Science\System theory\Developmental regulation.

Definition of PYD interventions\Developmental Science\System theory\intentionality.

Definition of PYD interventions\Developmental Science\System theory\Interaction between individual and their milieu.

Definition of PYD interventions\Developmental Science\Variation in developmental trajectories.

*Mechanism of action**Mechanism of action\Action on risk*

Mechanism of action\Action on risk\Action of protective factors.

BOX 1 NVivo Node Structure Report: Kate Hinds (*continued*)*Mechanism of action\Action on Thriving*

Mechanism of action\Action on Thriving\Agency.

Mechanism of action\Action on Thriving\Developmental affordances.

Mechanism of action\Action on Thriving\Ecological modelling.

Mechanism of action\Action on Thriving\Group formation and group work.

Mechanism of action\Action on Thriving\Identity development.

Mechanism of action\Action on Thriving\Interaction between individual and context.

Mechanism of action\Action on Thriving\Selection, optimisation and compensation.

*Mechanism of action\Depth of participation**Mechanism of action\Importance of structured activities**Mechanism of action\Multilevel changes**Mechanism of action\Partnerships**Mechanism of action\Plasticity**Moderation by context*

Moderation by context\Generalisability.

Moderation by context\Moderation by cultural background.

Moderation by context\Moderation by maturity.

Moderation by context\Moderation by risk.

*Taxonomy**Taxonomy\Emphasis on adult relationships**Taxonomy\Programme variations responding to cultural difference**Taxonomy\Schools programmes**Taxonomy\Social action programmes*

Taxonomy\Social action programmes\Programmes to address inequity.

BOX 2 NVivo Node Structure Report: Chris Bonell

Hierarchical name*Definition of PYD interventions**Definition of PYD interventions\Influences*

Definition of PYD interventions\Influences\Developmental systems theory.

Definition of PYD interventions\Influences\Policy context of deficit thinking and erosion of traditional supports.

Definition of PYD interventions\Influences\Policy context of discovery of childhood and adolescence.

Definition of PYD interventions\Principles

Definition of PYD interventions\Principles\All young people have potential.

Definition of PYD interventions\Principles\PYD compensates for environmental deficits.

Definition of PYD interventions\Principles\PYD compensates for environmental deficits\But limits on this.

Definition of PYD interventions\Principles\PYD more than prevention

Definition of PYD interventions\Principles\PYD more than prevention\Problem free is not fully prepared.

Definition of PYD interventions\Principles\PYD more than prevention\PYD must enable strengths or thriving.

Definition of PYD interventions\Principles\PYD more than prevention\PYD must enable strengths or thriving\
Expanded definition of health to include PYD assets.

Definition of PYD interventions\Principles\PYD more than prevention\PYD prevention may be compatible.

Definition of PYD interventions\Principles\PYD more than prevention\PYD prevention may be compatible\
But not empirically established.

Definition of PYD interventions\Principles\PYD more than prevention\PYD prevention may be compatible\
Counterargument that incompatible since prevention reinforces deficits.

Definition of PYD interventions\Principles\PYD more than prevention\PYD prevention may be compatible\
Prevention model is not itself static.

Definition of PYD interventions\Principles\PYD more than prevention\PYD prevention may be compatible\
Real programmes often combine prevention and promotion.

BOX 2 NVivo Node Structure Report: Chris Bonell (*continued*)*Definition of PYD interventions\What PYD must provide*

Definition of PYD interventions\What PYD must provide\Overarching frameworks.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Activities.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Activities\Opportunities for real challenges, skills and new horizons.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Atmosphere of hope.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Atmosphere of hope\Atmosphere of hope (2).

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Atmosphere of hope\Elements of atmosphere of hope.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Atmosphere of hope\Elements of atmosphere of hope\Empowerment.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Atmosphere of hope\Elements of atmosphere of hope\Empowerment\Burt unclear how defined.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Atmosphere of hope\Elements of atmosphere of hope\Empowerment\Burt unclear how defined\PYD as positive reinforcement, positive opportunities, positive norms.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Atmosphere of hope\Elements of atmosphere of hope\Positive expectations.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Goals = 5Cs (see below).

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Goals, atmosphere and activities\Goals = 5Cs (see below)\PYD must address Cs other than just competence.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Skills plus participation plus caring adults.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\Skills plus participation plus caring adults\Adults needs knowledge, observation, discipline, to support, have own support.

Definition of PYD interventions\What PYD must provide\Overarching frameworks\20 external assets PYD should offer.

BOX 2 NVivo Node Structure Report: Chris Bonell (*continued*)**Grouping (i.e. taxonomy)****Grouping (i.e. taxonomy)\Explicit taxonomies**

Grouping (i.e. taxonomy)\Explicit taxonomies\Categorisation by breadth vs. depth.

Grouping (i.e. taxonomy)\Explicit taxonomies\Categorisation by how many of 5Cs they addressed.

Grouping (i.e. taxonomy)\Explicit taxonomies\Categorisation by type, duration, quality.

Grouping (i.e. taxonomy)\Explicit taxonomies\Categorisation by whether aim to prevent as well as promote.

Grouping (i.e. taxonomy)\Explicit taxonomies\Categorisation by whether modifies only individual or environment too.

Grouping (i.e. taxonomy)\Explicit taxonomies\Categorisation by whether modifies only individual or environment too\Debate as to whether should modify wider environments.

Grouping (i.e. taxonomy)\Implicit taxonomies

Grouping (i.e. taxonomy)\Implicit taxonomies\Challenge or accept societal arrangements.

Grouping (i.e. taxonomy)\Implicit taxonomies\Focus on group or individual attributes.

Mechanism of action**Mechanism of action\Assumptions i.e. bases**

Mechanism of action\Assumptions i.e. bases\Individual-environment interactions.

Mechanism of action\Assumptions i.e. bases\Plasticity of development.

Mechanism of action\Assumptions i.e. bases\Plasticity of development\But one paper states strong genetic (presumably non-plastic) influence.

Mechanism of action\Key mechanisms

Mechanism of action\Key mechanisms\PYD enables developmental regulation.

Mechanism of action\Key mechanisms\PYD enables developmental regulation\Modifies intention self-regulation.

Mechanism of action\Key mechanisms\PYD enables developmental regulation\Modifies intention self-regulation
\Selection, optimisation and compensation.

Mechanism of action\Key mechanisms\PYD enables developmental regulation\Modifies intention self-regulation
\Selection, optimisation and compensation\Breadth and depth required former in early and latter in late adolescence.

Mechanism of action\Key mechanisms\PYD enables developmental regulation\Modifies intention self-regulation
\Selection, optimisation and compensation\Breadth and depth required former in early and latter in late adolescence\Balance between these not clear.

BOX 2 NVivo Node Structure Report: Chris Bonell (*continued*)

Mechanism of action\Key mechanisms\PYD enables developmental regulation\Positive interactions between individual and environment.

Mechanism of action\Key mechanisms\PYD enables developmental regulation\Positive interactions between individual and environment\Positive feedback can ensue.

Mechanism of action\Key mechanisms\PYD enables developmental regulation\Positive interactions between individual and environment\Requires environmental not merely individual intervention.

Mechanism of action\Key mechanisms\PYD offers affordances in environment or developmental attentiveness.

Mechanism of action\Key mechanisms\PYD promotes identity development via exploration and commitment.

Mechanism of action\Key mechanisms\PYD promotes identity development via exploration and commitment\Erikson's theory of identity development in childhood and adolescence.

Mechanism of action\Key mechanisms\PYD promotes identity development via exploration and commitment\Linear and cyclic aspects.

Mechanism of action\Key mechanisms\PYD promotes identity development via exploration and commitment\Requires breadth and depth.

Mechanism of action\Key mechanisms\PYD reinforces bonding to conventional institutions.

Mechanism of action\Key PYD assets

Mechanism of action\Key PYD assets\PYD as promoting assets.

Mechanism of action\Key PYD assets\PYD as promoting assets\5Cs widely cited as key assets.

Mechanism of action\Key PYD assets\PYD as promoting assets\5Cs widely cited as key assets\Character.

Mechanism of action\Key PYD assets\PYD as promoting assets\5Cs widely cited as key assets\Competence.

Mechanism of action\Key PYD assets\PYD as promoting assets\5Cs widely cited as key assets\Confidence.

Mechanism of action\Key PYD assets\PYD as promoting assets\5Cs widely cited as key assets\Connections.

Mechanism of action\Key PYD assets\PYD as promoting assets\5Cs widely cited as key assets\Originally 4 Cs but caring (i.e. compassion) added.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\20 youth assets.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Commitment but may be covered by confidence in identity.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Critical thinking.

BOX 2 NVivo Node Structure Report: Chris Bonell (*continued*)

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Critical thinking\Critical consciousness.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Critical thinking\Critical consciousness\3 stages: self-awareness, social awareness, global awareness.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Deficits in these function as risk factor for substance use.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Extra individual assets.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Extra individual assets\Community as well as individual assets.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Extra individual assets\Group as well as individual assets.

Mechanism of action\Key PYD assets\PYD as promoting assets\Some studies cite other assets\Learning to be productive, connect, navigate.

Mechanism of action\PYD distal products

Mechanism of action\PYD distal products\Contribution.

Mechanism of action\PYD distal products\Contribution\Behavioural and ideological components.

Mechanism of action\PYD distal products\Contribution\Rationale for this contribution.

Mechanism of action\PYD distal products\Domains of PYD impact.

Mechanism of action\PYD distal products\Horizontal and vertical pile up of assets (i.e. accumulation hypothesis).

Mechanism of action\PYD distal products\Horizontal and vertical pile up of assets (i.e. accumulation hypothesis)\Assets reduce risks in short term which determine longer term reductions.

Mechanism of action\PYD distal products\Molecular impacts (i.e. differentiation hypothesis).

BOX 2 NVivo Node Structure Report: Chris Bonell (*continued*)*Mechanism of action*\Risk behaviours

Mechanism of action\Risk behaviours\Evidence cited that PYD does reduce risk.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Accumulation hypothesis (i.e. horizontal and vertical pile-up).

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts).

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Direct and indirect mechanisms.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Direct and indirect mechanisms\Bonding key asset which indirectly reduces problem behaviours via reducing association with delinquent peers.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Direct and indirect mechanisms\Positive expectations modify self-concept and behaviour directly.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Raises question of whether assets are protective factors.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Raises question of whether assets are protective factors\Assets are protective factors that also promote thriving.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Raises question of whether assets are protective factors\Assets are protective factors that also promote thriving\ (e.g. relationship with adult reduces substance use).

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Raises question of whether assets are protective factors\Assets are protective factors that also promote thriving\Some domains of impact are protective, deficits raise risk.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Raises question of whether assets are protective factors\Protective factors interact with risks but assets do not necessarily.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Raises question of whether assets are protective factors\PYD and risk could coexist.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Differential hypothesis (i.e. molecular impacts)\Raises question of whether assets are protective factors\Two kinds of protection: buffering (interaction) and oppositional (not interaction).

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Identity hypothesis.

Mechanism of action\Risk behaviours\Few papers describe mechanisms to risk\Other theories fill gap.

BOX 2 NVivo Node Structure Report: Chris Bonell (*continued*)***Moderation by context***

Moderation by context\By age

Moderation by context\By duration of relationship

Moderation by context\By programme features

Moderation by context\By programme features\More assets addressed, more effective the programme.

Moderation by context\High risk may overwhelm assets

Moderation by context\Mechanisms and assets may vary culturally

Moderation by context\Mechanisms and assets may vary culturally\But that framework itself is generalisable.

Moderation by context\Mechanisms and assets may vary culturally\But variation may not be addressed sufficiently in theories.

Moderation by context\Mechanisms and assets may vary culturally\e.g. in more group-based cultures.

Appendix 10 Process studies: coding templates

CB: process synthesis coding template

1. Staffing continuity/quality/communication key to model
 - i. Lack of opportunities for intersite communication surmountable challenge
 - ii. Staff retention challenging
 - Limited hours a surmountable challenge
 - Limited retention could lead to more poorly trained staff
 - Retention could lead to failure of sustained relationships
 - Retention aided by high salaries
 - iii. Good leadership particularly key
 - iv. Some programs required 24-hour worker availability but challenging
 - v. Difficulty having staff expertise across a range of areas
 - Lack of training in some areas a surmountable challenge
 - But not always surmountable
2. Community integration critical for implementation.
 - i. Importance of outreach to key community members and parents
 - ii. Ethnicity and language critical for implementation
 - Recruit community members as staff
 - But volunteer staff not always reliable
 - iii. Importance of community generating ideas
 - iv. Parents feeling threatened by programme was barrier to uptake
 - v. Community needs
 - Need for childcare/support key determinant of uptake
3. Collaboration with other community agencies
 - i. Collaboration can include community funders
 - ii. Collaboration can include other providers or services
 - iii. Collaboration with schools critical
 - For recruitment: importance of single point of liaison with schools
 - For activities

4. Young people empowered to determine activities
 - i. Recipients reject some components
 - ii. Reject components focused on problems of potential
 - iii. Reject uninteresting components
 - iv. Some components actually alienating to some young people
 - v. Some based on pragmatism rather than empowerment
 - vi. Young people's choices could cause problems
 - But this moderated by site mix at outset
 - But financial incentives to ensure participation in key activities
 - vii. Young people choosing mentors may increase success
 - viii. Not all programmes enabled empowerment
5. Community can become topic of action
6. Recipient behaviour/relationships did not derail PYD
 - [In well-run sites] Poor recipient behaviour did not distract staff from PYD model
 - Group differences not barrier to collaboration (as long as well-run site)
 - Different young people required different styles
7. Determinants of fidelity
 - i. Fidelity compromised where sites merely relabelled existing services
 - ii. Fidelity strengthened by affinity with prior work
8. Sustained relationships also compromised by yp mobility
9. But staff in some sites worked hard to retain recipients
10. Pattern of provision reflected yp needs
11. Intensity requirement causing perverse effects.

KD: process synthesis coding template

12. Community-based relationships to enhance implementation and young people's engagement
 - Cultural relevance and integration to support implementation
 - Local community member to enhance implementation and young people's engagement
 - Utilising local resources
 - Local collaborations
 - Consultation with the community
 - Crossing language barriers
 - i. Parental co-operation

13. Communication and co-located After School sites
 - i. Informal meetings, communicate and share experience
 - ii. School liaison
 - iii. Communication of programme goals
14. Staffing issues
 - i. Difficulty hiring PT staff in after-school settings
 - ii. Lack of training and utilising staff skill set
 - iii. Hiring and staff turnover
 - iv. Training and staff turnover
 - v. Salary and staff retention
 - vi. Staff stability
 - vii. Engaged and unengaged staff
 - viii. Management turnover and leadership issues
 - ix. 24/7 'on call' burn-out
15. Youth–Staff relationships
 - i. Staff as role models
 - ii. Youth-centred approach
 - iii. Qualities and features 'valued' or 'desired' in the relationship
 - iv. Relationship compromised by youth geographical movement
 - v. Building on staff interests to re-engage youth
16. Peer-to-peer relationships
 - i. Overcoming differences
 - ii. Bonding/friendships
17. Funding
 - i. Sufficient funding
 - ii. Grant
 - iii. Long-term state investment
 - iv. Stipend to increase engagement
18. Case management
 - i. Problem-focused
19. Youth led components
 - i. Activities
 - ii. Research topics reflecting personal social experience
 - iii. Choosing mentors
20. Programme dose/intensity/implementation levels
 - i. Student attendance and outcomes

21. Fidelity

- i. Variation in fidelity depending on existing programmes

22. Flexible services

- i. Based on individual needs
- ii. Challenges/difficulties: 'hours target'
- iii. Based on age group

Appendix 11 Data transformation

TABLE 13 Data transformation

Programme	References used	Data transformation
All Stars	Gottfredson <i>et al.</i> ⁹²	Effect sizes were calculated from ORs and <i>p</i> -values as presented
MAPs	Gottfredson <i>et al.</i> ⁹³	Because of baseline imbalances in this non-randomised trial, gain scores were computed. Using gain scores, SDs from pre-test and post-test and the average of participants across both measurement occasions by arm, we calculated a between-groups standardised mean difference in change scores using $r = 0.5$ and sensitivity analysed for upper and higher correlations ($r = 0.1$, $r = 0.9$) between pre-test and post-test
BBBS	Tierney ⁹⁸	<p>A smaller subset of the analysis sample was used for initiation of substance-use outcomes. We assumed that this was balanced across arms (specific group <i>ns</i> were not provided) and estimated ORs using group percentages and sample sizes</p> <p>'Number of times hit someone' provided a difference between groups with an indicator of significance but no specific <i>p</i>-value. We pegged the <i>p</i>-value at 0.049 and also used group sizes to estimate a standardised mean difference</p> <p>'Number of times involved in a fight' provided little information. We set the standardised mean difference at 0 and took the largest variance from within included outcomes reported in the study. We sensitivity analysed this with the smallest within-study variance</p>
Cool Girls, Inc.	Kuperminc <i>et al.</i> ⁹⁵	ORs, but no significance or CIs, were reported. We used the OR with <i>p</i> -value = 0.5 and sensitivity analysed at <i>p</i> -value = 0.25 and <i>p</i> -value = 0.75
NGYCP	Millenky <i>et al.</i> ¹⁰⁷ Millenky <i>et al.</i> ¹⁰⁸	We used estimated probabilities from the linear probability models presented in the reports together with sample sizes to estimate ORs
PYDC	Tebes <i>et al.</i> ⁷⁹	Effect sizes were not altered, but see <i>Results</i> regarding the need to sensitivity analyse change scores
QOP	Schirm <i>et al.</i> ¹⁰³ Schirm and Rodriguez-Planas ¹⁰⁴ Schirm <i>et al.</i> ¹⁰⁵	We used estimated probabilities from the linear probability models presented in the reports together with sample sizes to estimate ORs
YARP	Berg <i>et al.</i> ⁸⁶	We interpreted the standardised path between intervention allocation and frequency of cannabis use as a point-biserial correlation and converted this into a standardised mean difference
YPDP	Wiggins <i>et al.</i> ¹¹²	We used weighted adjusted ORs as presented in the report

SD, standard deviation.

Appendix 12 User involvement

The policy and practitioner advisory group were provided with background information about the review, a summary of the draft theory and process synthesis, a list of potential hypotheses and key questions to focus their feedback. The material sent to the group below and a write-up of the summary of feedback from young people is included below.

Policy and practitioner stakeholders

The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes

Background

There are increasing calls for adolescent health interventions to address multiple- rather than single-risk behaviours because such behaviours cluster together^{31,32} and because such interventions are potentially more feasible and efficient.³³ PYD is one such intervention to address interclustered risk behaviours among young people. PYD is the dominant paradigm in youth work in the UK. The NYA, the major youth work organisation in the UK, defines such interventions as voluntary and informal educational activities that aim to bring about generalised youth development rather than merely to remedy 'problem behaviours'. Similarly, in the USA, PYD is defined in terms of its goal of developing bonding; resilience; social, emotional, cognitive, behavioural or moral competence; self-determination; spirituality; self-efficacy; clear and positive identity; belief in the future; recognition for positive behaviour; opportunities for pro-social involvement; and/or pro-social norms;^{30,44} academic, cognitive or vocational skills; confidence; connections to peers and adults; character in terms of self-control, respect and morality; and caring for others.³⁶

The UK government's Positive for Youth³⁸ report announced a multimillion pound investment in youth work, youth centres, the National Citizen Service and other youth volunteering projects. The most recent public health White Paper³⁹ cited such work as a key element in promoting young people's health. The Mayor of London and local government across the UK are also investing millions of pounds in various PYD interventions.⁴⁰ The devolved governments in Scotland and Wales also emphasise these principles and promote investment in PYD.^{41,42}

However, despite this widespread investment and potential, the evidence base for the public health benefits of such interventions is unclear. Although a systematic review examining non-health outcomes⁴³ reported benefits for self-confidence and self-esteem, school bonding, positive social behaviours, school grades and achievement test scores, the review did not systematically examine health effects. Systematic reviews of health outcomes have so far focused only on sexual health,^{44,45} reporting sustained effects but demonstrating considerable unexplained variability between programmes.

Aims

To search systematically for, appraise the quality of and synthesise evidence to address the following RQs:

- RQ1: what theories of change inform PYD interventions delivered to young people aged 11–18 years addressing substance use and violence?
- RQ2: what characteristics of participants and contexts are identified as barriers and facilitators of implementation and receipt in process evaluations of PYD?
- RQ3: what is the effectiveness and cost-effectiveness of PYD compared with usual or no treatment in reducing substance use (smoking, alcohol, drugs) and violence (perpetration and victimisation)?
- RQ4: what characteristics of participants and contexts appear to moderate, or are necessary and sufficient for, PYD effectiveness?

Overview of methods

Inclusion criteria

Types of participant Majority of participants are aged 11–18 years excluding studies of populations targeted on the basis of pre-defined physical and mental health conditions.

Types of intervention Voluntary education that aims not merely to prevent problem behaviour but also to address generalised (beyond health) and positive (beyond avoiding risk) development in terms of promoting bonding; resilience; social competence; emotional competence; cognitive competence; behavioural competence; moral competence; self-determination; spirituality; self-efficacy; clear and positive identity; belief in the future; recognition for positive behaviour; opportunities for pro-social involvement; and/or pro-social norms. Included studies of PYD interventions may address one of these interventions but applied to different domains (family, community, school) or more than one of these interventions in a single domain. Our definition excludes PYD delivered in school time or in custodial, probationary, clinical or post-school employment training settings. It also excludes interventions that target parents/carers alongside young people and that are focused on family functioning.

Types of outcome Substance (smoking, alcohol and/or drug) use or violence (perpetration and/or victimisation).

Types of studies Studies describing PYD intervention theory of change in relation to our outcomes (RQ1), process evaluation of PYD intervention (RQ2) and quasi-experimental or experimental evaluations of PYD intervention effectiveness or cost-effectiveness (RQ3).

Searches and screening

Electronic databases and websites were searched and consultation with experts was undertaken. Single-reviewer screening of titles and abstracts and then screening of full reports that could not be excluded based on title and abstract were undertaken.

Quality assessment, data extraction and synthesis

Quality assessment was assessed using existing criteria.^{56,64,66,67} Theory and process evaluations were subject to qualitative synthesis involving line-by-line coding and elaboration of coding template. We then used these syntheses to inform hypotheses to test in metaregressions and/or qualitative comparative analyses of outcome evaluations.

Results

Searching and screening

We provided the review group in *Figure 1*.

Data extraction

We included two tables (see *Appendix 13*), one describing the included theory papers and another outlining the quality assessment of theory papers and the characteristics of the process evaluations.

Summary of theory synthesis

Theory of PYD impacts on substance use and violence:

1. How interventions enable PYD

i. PYD interventions provide a positive environment

- positive expectations: all young people have the potential to develop healthily;
- affective relationships with adults: long-term, personal and caring compared with short-term, impersonal and instrumental relationships, with a family-like programme atmosphere;
- diverse activities and settings: offer diverse activities, settings and relationships; and
- active participation: youth empowered to choose activities, take on responsibilities.

ii. Taxonomy of PYD

Interventions vary according to:

- the breadth, depth and duration of activities offered;
- whether they aim to transform individual young people only or also the neighbourhood environments in which they live;
- whether they address only positive development or also prevention of risk behaviours; and
- whether they aim to enable young people to contribute to or challenge the existing social order.

iii. Interventions offer 'affordances' and opportunities for social learning

The positive environment interventions provide offer:

- 'affordances': resources that individuals make use of in the course of their development (e.g. relationships, challenges, education) – different individuals at different points in their maturation will make use of different affordances (e.g. younger adolescents require breadth whereas older adolescents require depth);
- opportunities for social learning: young people learn skills and pro-social conventional behaviours through (1) observing and imitating positive adult role models; and (2) the reinforcement of positive behaviour through material rewards or non-material celebrations.

iv. Young people improve their 'intentional self-regulation'

These opportunities enable young people to improve their 'intentional self-regulation' a cognitive and behavioural process which involves:

- intentionality (reflecting on and learning from existing behaviour to inform intentions for future behaviour);
- selection (from a range of opportunities) of those activities that provide the young person with psychological and social assets to meet their ultimate personal goals;
- optimisation (acquiring, refining, coordinating and applying goal-relevant resources towards achieving the selected targets); and
- compensation (redirecting behaviour if initial efforts are thwarted).

v. Young people accrue 'positive assets'

Improve 'intentional self-regulation' enables young people to accrue positive assets, for example the '5 Cs':

- competence (includes self-efficacy, resilience, or social, cognitive, behavioural and moral competence);
- confidence (includes self-determination and a clear and positive identity);
- connection (positive, social including bonding);
- character; and
- caring (or compassion, character includes spirituality and a belief in the future).

vi. A virtuous cycle of 'developmental regulation'

As young people's positive assets accrue, they can make better use of the opportunities available in their environment, which leads to:

- 'developmental regulation', that is, a positive feedback cycle in which individuals gain more benefit from opportunities in their environment;
- this enables them in turn to make a positive contribution to their communities and societies or, as a few authors would argue, enables them to contribute by critiquing and challenging inequities present in the existing social order; and
- these commitments and contributions signal the development of a stable and positive adolescent identity.

2. How PYD reduces risk of substance use and violence

i. Relationship between positive development and risk reduction

Different authors variously present this as complementary albeit with comparatively little theorisation of how risks are reduced. Many argue that positive assets and protective factors are often the same thing, such as communication skills and self-efficacy.

ii. Various functions

- Buffering function: risk factors in a young person's environment have less impact on those who have positive assets; and
- Compensation: even if young person engages in a risk behaviour, their possession of positive assets ameliorates its impact either on their health/well-being or on their engagement in other risk behaviours (this use of 'compensation' is quite distinct from that within the description of intentional self-regulation in *Theory of PYD impacts on substance use and violence*, iv. *Young people improve their 'intentional self-regulation'*) – but this function is regarded as less important than buffering.

iii. Various mechanisms

- Molecular mechanism: a specific asset exerts specific protection against a specific risk; and
- Pile-up mechanism: the accumulation of multiple assets is protective regardless of the specific assets involved. Pile-up can involve (1) vertical pile-up – the greater number of different assets at a point in time or over time, the less risk; and (2) horizontal pile-up – the more settings that provide the same assets, the less risk.

3. Assessment of theory of PYD effects on substance use and violence

i. Limitations of PYD theory

Obvious limitations are that:

- the literature focuses more on asserting the normative value of PYD assets than on specifying causal mechanisms;
- although notions of intentional self-regulation and developmental regulation offer a general mechanism for positive development, they offer little insight into how specific assets develop; and
- there is even less attention given to how these interventions reduce risk behaviours; where causal mechanisms are specified this largely draws on theories from conventional prevention science such as social learning theory.

ii. Usefulness of PYD theory

Despite this, the PYD literature offers three useful insights:

- interventions may produce significant reductions in risk in the longer term even if effects in the short term are small because the interventions place the young person on a trajectory whereby more and more positive assets accrue leading to more significant pile up and molecular impacts on risk reduction;
- the focus on pile up of multiple positive assets might ensure that PYD interventions are more likely than traditional prevention interventions to address multiple determinant of problem behaviours; and
- interventions which emphasise hope in young people's positive potential may be less likely to inadvertently reinforce risk behaviours than prevention interventions which emphasise risk.

iii. Hypotheses about what factors might affect the impact of PYD

The above might lead us to hypothesise that:

- interventions that offer and breadth of activities may be more effective for younger adolescents, whereas those which emphasise depth may be more effective for older adolescents;
- interventions that combine prevention and positive development may be less effective than those that focus only on positive development;
- interventions of more than 1 year's duration may be more effective than those of shorter duration; and interventions may be more effective for participants with low or moderate levels of baseline risk given that there is more scope for stimulating 'intentional self-regulation'.

Summary of process synthesis

Synthesis of process evaluations of Positive Youth Development interventions

Theme 1: collaboration with the community

A major theme across a number of studies was the importance of working with local communities.

Cultural sensitivity The importance of cultural sensitivity when implementing programmes in ethnic minority communities was a subtheme. Formal and informal community engagement was a key factor in ensuring that programmes were culturally sensitive, accessible and appealing to young people and their parents, as well as to the wider community. This can generate ideas for culturally appropriate activities and allay fears among minority ethnic groups that the programme is an outside imposition. The latter could be important in ensuring that young people accessed an intervention. It is also important that interventions overcome potential language barriers by translating key proceedings and communicating with indigenous members in their language. In some cases, programmes recruited community members as staff. Employing a respected community member was perceived by many participants as pivotal to successful implementation. Such actions could also be seen as providing the additional benefit of providing local role models.

Challenges with community engagement However, challenges in community engagement and establishing trust also formed a subtheme. In attempting to involve parents and local members of the community, staff reported difficulties in establishing good working relationships. Volunteer mentors, for example, were not always reliable in maintaining contact, which could leave participants feeling disappointed.

Building trusting and openly communicative relationships with parents could also be challenging. Providers could find that parents remained anxious about admitting to family problems and could perceive staff as intrusive or felt threatened by mentors developing close relationships with their children, and this sometimes limited access.

Theme 2: collaborating with other community agencies

Another theme was the importance of collaboration with other community agencies to enable programme implementation. This could expand the range of activities offered, a crucial element in PYD. Examples include local libraries helping to deliver a reading programme; local youth services hosting some events and offering anger management training; local funding bodies providing resources; and local health services providing training in substance-abuse prevention and conflict resolution, etc. This could be particularly important when there were gaps in programme providers' expertise.

Collaboration with schools A subtheme was that collaboration with schools was crucial to implementation but could raise challenges. This was obviously important in recruiting to ASPs but lines of communication could sometimes be unclear. Working with schools could sometimes also lead to unplanned changes in what youth development consisted of, for example, seeing it as an alternative rather than a complement to normal schooling where schools sought this.

Theme 3: young people's relationship with programme providers and peers

Another theme was the importance of young people's relationships with programme providers and peers in maximising the acceptability and impact of interventions.

Calm and authoritative programme providers A key subtheme was the importance of programme providers attending to young people in a calm and nurturing yet authoritative way in the context of any challenging behaviour which arose from participants. One study reported that successful implementation was associated with staff responding to challenging behaviour from young participants by signalling their continued commitment to providing unconditional positive regard. Implementation was successful when staff attended to young people as individuals, were consistent and used a balance between nurture and discipline. Other studies reported a range of poor staff responses across different programme sites, including staff disengagement and apathy, excessive familiarity and excessive and inconsistent discipline.

Positive peer relations A further subtheme was the importance of positive peer relations, supported by each other and by staff. Some projects succeeded in developing cohesion and support between participants even where there were social differences such as differences in gang memberships. Sites achieved this by facilitating the young people to recognise common problems, concerns and needs. This could occur only in programmes that were well managed. One study reported that a lack of organisation and high staff turnover were key factors in young people not developing good intergroup relationships.

Theme 4: staff retention

Staff retention was another key theme across studies, which reported on the importance of staff continuity to ensure that programmes were implemented as planned. This was particularly important in those programmes that aimed to develop durable relationships between staff and participants, such as mentoring programmes. Evaluations did not report on whether or not context in terms of study sites or participants was a crucial determinant of staff retention. Nonetheless, we include this theme in our synthesis because it may well be that local context (e.g. in terms of demographic and economic factors) might be an important influence, but also because this theme could be of interest to those implementing programmes in future.

Difficulty offering full-time posts in the youth work field One subtheme was that staff retention was particularly challenging in youth programmes because most could not offer full-time positions, making it difficult to retain qualified staff in only part-time jobs. Some projects aimed to overcome this by recruiting those not looking for full-time work, such as college students interested in gaining experience of youth work, or other professionals who were interested in working with young people in addition to their primary employment.

Retention and training Another subtheme was that retention problems could mean that staff were insufficiently trained in the short periods in which they were employed. One study found that sites with high employee turnover were less likely to have staff who were highly trained. Creative attempts to compensate for this included drawing on existing skills that happened to be held by staff members, such as outdoor activities and skills in art, and incorporating these opportunistically into programme activities. However, such compensation was not always possible and this could mean that programmes might not be able to provide the range of activities normally expected of PYD. This was reported to be a particular problem in terms of academic educational components.

Theme 5: youth-led empowerment

Our final cross-cutting theme concerns the importance of and challenges to ensuring that young people are empowered to make decisions about their engagement in programme activities. Again, it is questionable whether or not this theme engaged with notions of context sufficiently to justify its inclusion here. We have, however, retained it because the extent to which empowerment raises challenges for programme implementation and acceptability will inevitably depend on the characteristics of the young people participating.

Young people determining their own engagement in activities One subtheme relates simply to the extent to which young people were empowered to determine the PYD activities in which they participated. To ensure that this occurred, staff needed to be trained in facilitation techniques to halt any tendency to be too directive. In many projects, young people's decision-making processes were considered more important than their final choice of activity. One study reported that when activities were imposed in a top-down manner they failed and were abandoned. However, other studies reported that young people in some particular interventions had very limited empowerment to shape and determine their involvement in positive youth activities. For example, some studies gave young people no or extremely limited options.

Tensions arising from choice A subtheme was that tensions could arise when empowering young people to choose the activities in which to engage. Less appealing programme components were often rejected by young people on the basis that they were uninteresting and unengaging to them. Often, these were found to be activities with a vocational or academic learning component. Some studies reported that such components did not sit well with the wider culture of some programmes that aimed to be distinct from school in terms of atmosphere. Some sites offered more recreational activities (e.g. outings to the cinema, ice-skating, swimming, etc.) because they attracted recruits and then encouraged participants to try more education-focused activities. However, participants continued to opt for recreational activities and resisted attempts at educational engagement. Some sites provided financial incentives to participants to engage in educational activities but these were effective for only short periods of time and for only students already inclined to engage educationally. These data suggest a tension between PYD's orientation towards empowering young people to choose their own activities and its commitment to ensuring that young people engage in a variety of activities and develop a range of competencies including those related to academic learning and careers.

Hypotheses arising from this synthesis

- Interventions that have specific methods to engage communities will be more effective.
- Projects that engage with schools will achieve better recruitment.
- Interventions that are delivered by well-trained staff will be more effective.
- Interventions that have better staff retention will be more effective.
- Interventions that offer some choices but require some engagement with educational components will be more effective.

Questions for stakeholders

The following questions were sent to stakeholders and provided a focus for the informal interviews.

- Do the themes that we have identified in the theory and process evaluation literature make sense to you?
- Is our synthesised theory of change clear? Is it useful in thinking about real PYD interventions?
- Do our syntheses resonate with your own experience of youth programmes?
- In terms of the relevance of our literature to England and the rest of the UK, are we covering all the key issues?
- Are there other issues that we need to bear in mind when thinking about relevance to the UK?
- Do our hypotheses make sense? Are there any others that we should bear in mind?

Consulting with young people

The members of the ALPHA group were presented with the following PowerPoint presentation:

'Positive Youth Development' projects by Adam Fletcher

- Voluntary and informal educational activities aiming to bring about generalised youth development rather than merely remedying 'problem behaviours.'
- Seen as important in promoting young people's health
- In 2011 the UK Government announced a multimillion pound investment in (interventions) youth work, youth centres, National Citizen Service, youth volunteering projects, etc.
- So, what do research projects that studied PYD tell us?

How can we enable Positive Youth Development? The research says . . .

Provide a positive environment through:

1. positive expectations;
2. real relationships with adults;
3. diverse activities and settings;
4. active participation.

Positive Youth Development methods are distinctive because . . .

1. Breadth, depth and duration.
2. Focus on positive development not just risk behaviours.
3. Young people contribute to or challenge society.

Young people change through . . .

1. Reflecting on and learning from behaviour.
2. Selecting activities that help them to meet goals.
3. Changing behaviour if not successful.

Young people gain positive assets such as . . .

1. Competence.
2. Confidence.
3. Connection.
4. Character.
5. Caring.

Do you agree?

How might Positive Youth Development decrease smoking, drinking, drug use and violence? The research says . . .

- Seeing positive behaviours and having links to positive adults, peers and institutions can reduce peer influence on substance use and violence.
- Learning how to resolve conflict from youth projects can reduce violence.
- Feeling a part of society, connected to people and believing in yourself protects against behaviour problems.
- Young people learn to make choices based on internal values rather than peer influence.

What affects how well Positive Youth Development works? The research says . . .

- Interventions that offer a breadth of activities may be better for younger adolescents. Those interventions with 'depth' may be better for older adolescents? What do you think?
- Interventions that prevent risk behaviours as well as promote positive development may work less well than those that focus only on positive development? What do you think? Different for different groups?
- Interventions that last more than 1 year may be best?

How important are these if you had to rank them?

- Youth work projects that have specific methods by which to engage local communities.
- Projects that engage with schools.
- Projects that are delivered by well-trained staff.
- Project that have better staff retention (less turnover and new people).
- Projects that offer some choices but require some engagement with educational components.

Summary of feedback provided by young people

Introduction

Adam Fletcher came to talk to the ALPHA group about the findings of the PYD systematic review (28 March 2015). A total of 13 young people (10 boys and 3 girls), aged 14–16 years, attended.

Findings from the outcome evaluation

When considering the findings, young people agreed that PYD programmes can have an effect on their drug use but not on their alcohol consumption unless programmes provide a social venue to offer an alternative to drinking. They both agreed and disagreed that PYD programmes can influence smoking and have a beneficial influence on violence.

Findings from the process evaluation

Young people discussed the five themes and wrote messages for the research team to think about.

Theme 1: collaboration with the community

Young people said a mix of external and community members is best, as external people bring other experiences. They agreed that cultural sensitivity is important.

Theme 2: collaborate with other community agencies

Young people said that focusing on one public health topic is better than having different sessions and that it is good to be on-site (school and youth club) but that it needs to be separate otherwise they will worry about confidentiality.

Theme 3: focus on young people's relationship with Positive Youth Development providers and peers

Young people agreed that providers need to relate in a calm and nurturing but authoritative way, that skilled providers are required to bridge social differences with peers and that relationship with workers need to offer a balance between friend and teacher and leader.

Theme 4: retain programme staff

Young people agreed that staff retention is needed to increase trust and to build relationships, because staff loss affects the attendance of young people and their focus when they do attend. There is a need to ensure that there are always senior staff and younger staff who are relatable.

Theme 5: youth-led empowerment

Young people thought that there is a need to give them lots of options to choose between, but that it is also important to impose certain limits within budget possibilities while taking into consideration what young people want to learn. It is also important to allow for some recreational time for talking to friends.

What are the implications of this research for youth work in the UK?

1. Youth centre closures: youth centres closures and staff turnover because youth workers need to find more permanent positions. New staff are brought in but it affects the youth work as it takes time to build a relationship and trust with the young people and can lead to a decrease in young people attending. Young people are not empowered because the people who run youth services close a centre no matter what the young people say.
2. The value of youth work: not enough people know how youth work is used, valued and appreciated by young people. The service is important because when youth clubs are shut there is nowhere to go. They need to stay open so that young people have somewhere to go and to learn about real world issues.
3. Promoting what youth services do including in the community: youth services need to better promote what they are doing so that young people attend and to ensure that services are less likely to be cut. However, if services were to open to young people in new areas, it is not clear if would be used because people already have routines in the evenings and at weekends. In some areas services do not work with community members but this should be promoted.
4. Somebody to talk to: need people other than parents and teachers to whom to talk. As being a youth worker involves good training and having the right personality, youth workers know that they are dealing with young people so need to be nurturing and caring but also authoritative if needed.

How, and to whom, should we communicate this research?

- People in the council, as they are making the cuts to youth services because they do not know the value of them.
- Communities in general need to know that the youth services are keeping young people off the streets. Then people in the community will value the youth club/services.
- Tell youth fora, youth leaders and youth centres. Send information to youth fora in Wales and they should disseminate them to individual youth workers.
- Send research directly to youth clubs/youth workers if you can find out who they are.

Appendix 13 Characteristics of theory reports

TABLE 14 Characteristics of theory reports (studies included in the review to answer RQ1)

Study, title	Stated aims of the report	Existing theories cited (theory name, author)	Normative or causal focus
Benson <i>et al.</i> ¹¹⁴ Youth development, developmental assets and public policy	The chapter describes the strength-based youth development approach in large part by comparing and contrasting it with the deficit-based orientation to successful development. It also discusses the theoretical and empirical basis of the developmental asset framework as a prime example of PYD, a comprehensive conceptualisation of developmental well-being and a generator of knowledge regarding the developmental pathways of young people. We identify relevant social and cultural dynamics affecting youth, consider their implication for youth development policy and highlight a number of public policies from around the country that reflect the tenets and unfolding wisdom of healthy youth development. The chapter concludes by assessing the sociopolitical prospects for developmental principles and knowledge to actually inform and shape public policy for young people	Ecological model of human development: Bronfenbrenner ¹⁵⁴	Normative
Benson ¹¹³ Developmental assets: an overview of theory, research and practice	To discuss the concepts of developmental assets, asset-building communities and asset-building society To discuss the 4-H survey in relation to assets and damaging behaviours	Developmental systems theory: Ford and Lerner; ¹⁵⁵ Gottlieb ¹⁵⁶ Action theory of development: Brandtstädter ¹⁵⁷	Normative
Benson and Scales ¹¹⁵ PYD and the prevention of youth aggression and violence	In this paper, we conduct analyses on several databases of 6th- to 12th-grade students in the USA to explore the linkage of positive relationships, opportunities, skills and values, called Developmental Assets, to prevention of youth aggressive and violent behaviours	Ecological model of human development: Bronfenbrenner ¹⁵⁴ Ecological theory: Bronfenbrenner and Morris ¹⁴⁴	Normative
Benson <i>et al.</i> ¹¹⁶ The contribution of the developmental assets framework to PYD theory and practice	A considerable body of literature on developmental assets has emerged in the past two decades, informing research and practice in education, social work, youth development, counselling, prevention and community psychology. In addition to synthesising this literature, this chapter discusses: the recent development of the Developmental Asset Profile, an instrument designed, in part, to assess change over time; the utilisation of asset measures in international research; the expansion of the assets framework to early childhood and young adults; and new research using latent class analysis to identify classes or subgroups of youth	–	Normative

continued

TABLE 14 Characteristics of theory reports (studies included in the review to answer RQ1) (continued)

Study, title	Stated aims of the report	Existing theories cited (theory name, author)	Normative or causal focus
Berg <i>et al.</i> ⁸⁶ YARP: a multilevel intervention designed to increase efficacy and empowerment among urban youth	YARP, a federally funded research and demonstration intervention, utilises youth empowerment as the cornerstone of a multilevel intervention designed to reduce and/or delay onset of drug and sex risk, while increasing individual and collective efficacy and educational expectations. The intervention, located in Hartford, CT, served 114 African Caribbean and Latino high school youth in a community education setting and a matched comparison group of 202 youth from 2001 to 2004. The strategy used in YARP begins with individuals, forges group identity and cohesion, trains youth as a group to use research to understand their community better (formative community ethnography), and then engages them in using the research for social action at multiple levels in community settings (policy, school-based, parental, etc.) Engagement in community activism has, in turn, an effect on individual and collective efficacy and individual behavioural change. This approach is unique insofar as it differs from multilevel interventions that create approaches to attack multiple levels simultaneously. We describe the YARP intervention and employ qualitative and quantitative data from the quasi-experimental evaluation study design to assess the way in which the YARP approach empowered individual youth and groups of youth (youth networks) to engage in social action in their schools, communities and at the policy level, which in turn affected their attitudes and behaviours	Ecological systems theory: Bronfenbrenner; ¹⁵⁴ Dryfoos ¹⁵⁸ Identity theory: Phinney ¹⁵⁹ Social learning theory: Bandura ¹²⁸ Social construction theory: Berger and Luckmann ¹⁶⁰ Critical transformational theories: Giroux; ¹⁶¹ Gitlin; ¹⁶² Foucault, cited in Martin, ¹⁶³ Bourdieu and Passeron; ¹³⁸ Freire; ¹⁴⁰ Gramsci ¹³⁷	Normative Causal
Busseri <i>et al.</i> ¹¹⁷ Breadth and intensity: salient, separable and developmentally significant dimensions of structured youth activity involvement	We present a theory-based framework for studying structured activity involvement as a context for PYD based on two key dimensions: breadth and intensity of involvement. Our main goal is to demonstrate the separability, salience and developmental significance of these two dimensions	Identity development theory: Erikson; ¹²⁹ Marcia ¹³⁰ Life-span development processes of selective optimisation with compensation (e.g. Baltes, ¹⁶⁴ Baltes <i>et al.</i> ¹⁶⁵) Concept of 'affordances' in Gibson's ¹⁶⁶ ecological theory of human perception ¹⁶⁷	Causal
Catalano <i>et al.</i> ¹¹⁸ Prevention science and positive youth development: competitive or cooperative frameworks?	To examine the convergence in the critiques and recommendations for the future of programmes to promote healthy development and prevent problem behaviours among children and adolescents	Attachment theory Identity development theory: Erikson ¹²⁹ Ecological model of human development: Bronfenbrenner ¹⁵⁴	Causal

TABLE 14 Characteristics of theory reports (studies included in the review to answer RQ1) (continued)

Study, title	Stated aims of the report	Existing theories cited (theory name, author)	Normative or causal focus
Ginwright and Cammarota ¹¹⁹ New Terrain in Youth Development: The Promise of a Social Justice Approach	Presents a youth development model that addresses structures of power and teaches youth to understand how their opportunities are circumscribed by larger political, economic and social forces. Critiques two dominant approaches to youth development which have oppressed urban youth of colour. The proposed model views youth as agents of social change, fostering 'the praxis of critical consciousness and social action' by taking youth through self-awareness, social awareness and global awareness	Critical consciousness: Freire ¹⁴⁰	Normative
Kia-Keating <i>et al.</i> ¹²⁰ Protecting and promoting: an integrative conceptual model for healthy development of adolescents	This article draws on extant research to delineate links between the risk and resilience and PYD literatures	–	Causal
Kim <i>et al.</i> ¹²¹ Towards a new paradigm in substance abuse and other problem behaviour prevention for youth: youth development and empowerment approach	The purpose of this article is to: (1) address a paradigm shift taking place in the field of substance-abuse prevention directed at youth; and (2) to introduce an innovative approach to substance-abuse and other problem behaviour prevention that reflects this shift in prevention paradigm	Social control theory: Hirshi ¹³⁵ Social learning theory: Bandura ¹²⁸ Social development model: Hawkins and Weiss ¹⁶⁸ Problem behaviour theory: Jessor and Jessor ¹⁶⁹ Expectations-states theory: Foschi ¹⁷⁰	Causal
Lee ¹²² Construction of an integrated positive youth development conceptual framework for the prevention of the use of psychotropic drugs among adolescents	This is a theoretical paper which aims to construct an integrated conceptual framework for the prevention of adolescents' use and abuse of psychotropic drugs. This paper provides empirical support for integrating a PYD perspective in the revised model. It further explains how the 15 empirically sound constructs identified by Catalano <i>et al.</i> ³⁰ and used in a PYD programme, the Project P.A.T.H.S., relate generally to the components of the revised model to formulate an integrated PYD conceptual framework for primary prevention of adolescent drug use	Social learning theory: Bandura ¹²⁸ Symbolic interaction: Blumer ¹⁷¹ Operant conditioning theory: Skinner ¹⁷²	Causal
Lerner and Lerner ¹²³ Towards a New Vision and Vocabulary About Adolescence: Theoretical, Empirical, and Applied Bases of a 'Positive Youth Development' Perspective	Towards a New Vision and Vocabulary About Adolescence: Theoretical, Empirical, and Applied Bases of a 'Positive Youth Development' Perspective	Developmental systems theory: Bronfenbrenner; ¹⁵⁴ Gottlieb ^{156,173}	Normative

continued

TABLE 14 Characteristics of theory reports (studies included in the review to answer RQ1) (continued)

Study, title	Stated aims of the report	Existing theories cited (theory name, author)	Normative or causal focus
Lerner <i>et al.</i> ¹²⁴ Individual and contextual bases of thriving in adolescence: a view of the issues	We introduce this special issue on the individual and contextual bases of adolescent thriving by describing the relational developmental systems theory-based, PYD perspective that frames much of contemporary research about health and positive development across the adolescent period and that, more specifically, frames the 4-H study of PYD, the data set from which the empirical work in this special issue is drawn	Bioecological theory: Bronfenbrenner and Morris ¹²⁷ Action theory models of intentional, goal-directed behaviours Baltes; ¹⁶⁴ Brandtstädter; ¹⁵⁷ Heckhausen; ^{174,175} Heckhausen <i>et al.</i> ¹⁷⁶ Life-course theory: Elder ¹⁷⁷ Dynamic systems theory: Thelen and Smith ¹⁷⁸ Holistic person–context interaction theory: Magnusson ¹⁷⁹ Developmental systems formulations: Ford and Lerner; ¹⁵⁵ Gottlieb ¹⁵⁶	Normative
Perkins <i>et al.</i> ¹²⁵ Community Youth Development: A Partnership for Action	The concept of community youth development is introduced and explained to raise the level of accountability, significance, and urgency for developing comprehensive responses to the epidemic of risk facing America's youth. The two theoretical models of adolescence (i.e. PYD and Risk and Resiliency) that are employed as the pillars of this approach are also presented	–	Normative
Roth and Brooks-Gunn ⁵⁰ Youth development programs: risk, prevention and policy	We focus on the promise and reality of youth development programmes. After a brief review of the available evidence about programme effectiveness, we define the elements of youth development programmes based on theoretical writings and ethnographic studies. We then investigate the reality in two ways. First, we map the defining principles of youth development to practice by looking at which elements are present in successful programmes. Second, we investigate the relation between these elements and programme outcomes. We conclude with directions for the future	–	Causal
Schwartz <i>et al.</i> ¹²⁶ Addressing the challenges and opportunities for today's youth: towards an integrative model and its implications for research and intervention	This article calls for, and proposes some tenets of, model building in adolescent psychosocial development. Specifically, it is suggested that there is a need for a model that draws from the risk-protection approach, from which many prevention science approaches are drawn, and the applied developmental science perspective, from which many PYD approaches are drawn	Selection, optimisation and compensation model: Baltes and Baltes ¹⁸⁰ Theory of planned behaviour: Ajzen and Fishbein ¹⁸¹	Causal

P.A.T.H.S., Positive Adolescent Training through Holistic Social programmes.
Dashes indicate that no existing theory was cited.

TABLE 15 Quality assessment of theory studies

Paper	Type of theory	Clarity of constructs		Clarity of relationship between constructs		Testability		Parsimony		Generalisability		Total score	
		CB/JT	KH	CB/JT	KH	CB/JT	KH	CB/JT	KH	CB/JT	KH	CB/JT	KH
Benson <i>et al.</i> ¹¹⁴	Normative	0	0.5	0	0.5	0	0	0	0	1	1	1	2
Benson ¹¹³	Normative	0	0.5	1	0	1	0	0	0	1	0	3	0.5
Benson and Scales ¹¹⁵	Normative	1	0	0.5	0	1	0	0	0	1	1	3.5	1
Benson <i>et al.</i> ¹¹⁶	Normative	0	1	0	0.5	1	1	0	0	1	1	2	3.5
Berg <i>et al.</i> ⁸⁶	Normative/causal	0	1	0	1	0	1	0	1	1	1	1	5
Busseri and Rose-Krasnor ¹¹⁷	Causal	1	1	0.5	1	1	1	1	1	1	1	4.5	5
Catalano <i>et al.</i> ¹¹⁸	Causal	0.5	0.5	0.5	0.5	1	0	1	0	1	1	4	2
Ginwright and Cammarota ¹¹⁹	Normative	0	0	1	0	0	0	1	0	1	0	3	0
Kia-Keating <i>et al.</i> ¹²⁰	Causal	0	1	0	1	1	0	0	1	1	1	2	4
Kim <i>et al.</i> ¹²¹	Causal	1	1	1	1	1	1	1	1	1	1	5	5
Lee ¹²²	Normative	0	0	0	0.5	0	0	0	0	1	1	1	1.5
Lerner and Lerner ¹²³	Normative	1	0.5	0.5	1	1	1	1	1	1	1	4.5	4.5
Lerner <i>et al.</i> ¹²⁴	Normative	1	0	1	0.5	1	0	1	0	1	1	5	1.5
Perkins <i>et al.</i> ¹²⁵	Normative	1	0	1	0	0	0	0	0	1	1	3	1
Roth and Brooks-Gunn ⁵⁰	Causal	1	1	1	0	0	0	1	1	1	1	4	3
Schwartz <i>et al.</i> ¹²⁶	Causal	1	0	1	0	1	0	1	0	1	0	5	0
These scores were agreed between KH and JT													
Ginwright and Cammarota ¹¹⁹	Normative	0	0	0	0	0	0	1	1	1	2	2	
Perkins <i>et al.</i> ¹²⁵	Normative	1	1	0	0	0	0	0	0	1	3	3	
Roth and Brooks-Gunn ⁵⁰	Causal	1	1	1	0	0	0	1	1	1	4	4	

1, yes; 0.5, partial; 0, no.

Appendix 14 Characteristics and quality assessment of process evaluations studies

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2)

Study details	Characteristics of process evaluations	
Armstrong and Armstrong⁸⁴		
Stated aim of study	To identify characteristics predicting the effective implementation of ASPs	
Methods	Data collection: in-depth interviews; non-participant observation	
	Data analysis: descriptive analysis	
Details of participants	Country: USA, 'urban southwest city'	
	Actual sample: programme staff	
	Sample number: sampled '6 Sites'	
	Age: not stated	
	Sex: mixed	
	Ethnicity: not stated	
	SES: not stated	
Details of PYD programme	Description: supervised ASP	
	Target population: 'At risk youth'	
	Theory: not stated	
	Provider(s): Parks, Recreation, and Libraries Department	
	Training: not stated	
	Setting: after-school sites	
	Content: life skills, educational support, healthy living skills, social and peer interaction, physical activity, cultural awareness and fine arts and locally relevant programme activities	
	Length/intensity: not stated	
Notes		
<i>Quality appraisal</i>		
Questions used to judge rigour and relevance	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	Yes	Randomly selected five sites from 16 sites identified as serving an 'at-risk youth' population
Were steps taken to minimise bias and error/increase rigour in data collection?	No/not stated	None provided
Were steps taken to minimise bias and error/increase rigour in data analysis?	No/not stated	None provided
		continued

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
Were the findings of the study grounded in/supported by data?	No/not stated	None provided
There was good breadth and/or depth achieved in the findings?	Yes	Sufficient breadth and depth
The perspectives of young people were privileged?	No/not stated	No data collected from young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	Low	
Usefulness of findings	High	
Baker et al.⁸⁵		
Stated aim of study	To evaluate the testing and refining of an empirically based model of effective anti-violence programmes for youth	
Methods	Data collection: unstructured interviews and focus groups Data analysis: ethnography; no further details provided	
Details of participants	Country: USA, South Baltimore Actual sample: programme staff; participating youth Sample number: unclear Age: not stated Sex: mixed Ethnicity: not stated SES: not stated	
Details of PYD programme	Description: a violence, delinquency and substance-abuse prevention programme Target population: inner city youth 'at risk' of delinquency, violence and substance abuse Theory: Scientific evidence that utilising youth advocates to guide youth development can have an impact Provider(s): a South Baltimore youth centre Training: not stated Setting: after school youth centre Content: safe haven to engage in positive social activities; empowering youth in decision-making; providing job training via computers; case management, mentoring; street worker for community outreach; tutoring academic skills; community involvement and collaboration with other agencies Length/intensity: not stated	

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
Notes		
<i>Quality appraisal</i>		
Questions used to judge the rigour and usefulness	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	No/not stated	None provided
Were steps taken to minimise bias and error/increase rigour in data collection?	Cannot tell	None provided
Were steps taken to minimise bias and error/increase rigour in data analysis?	No/not stated	None provided
Were the findings of the study grounded in/supported by data?	No/not stated	No direct quotes provided
There was good breadth and/or depth achieved in the findings?	No/not stated	Lack of depth
The perspectives of young people were privileged?	Cannot tell	No direct quotes from young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	Low	
Usefulness of findings	Low	
Berg et al.⁸⁶		
Stated aim of study	To describe and assess the way in which YARP empowered individual youth and groups of youth to engage in social action in their schools, communities and at the policy level	
Methods	Data collection: interviews with staff, ethnographic observation, youth focus groups and youth self-reflection	
	Data analysis: participant observation	
Details of participants	Country: USA, Hartford, CT	
	Actual sample: participating youth	
	Sample number: 114	
	Age: 14- to 16-year-olds	
	Sex: mixed	
	Ethnicity: African Caribbean; Latino	
	SES: not stated	

continued

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
Details of PYD programme	<p>Description: YARP</p> <p>Target population: youth aged 14 to 16 years</p> <p>Provider(s): prevention research educators</p> <p>Theory: prevention, cognitive and critical theory including: (a) ecological theory; (b) identity theory, (c) learning and instructional theories (including social learning, social construction, and multiple intelligence), and (d) critical, transformative theories including voice, empowerment and action research</p> <p>Training: educators are trained in action research methods and empowerment-oriented, social construction facilitation skills</p> <p>Setting: community-based after-school and summer programme</p> <p>Content: participatory action research involving formative community ethnography (e.g. training youth to identify adolescent risk behaviours, develop a collective action plan and carry out activities as a group, including using research to understand their community better)</p> <p>Length/intensity: 4 hours/day for 8 weeks, 8-month implementation period for action plan</p>	
Notes		
<i>Quality appraisal</i>		
Questions used to judge rigour and relevance	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	No/not stated	None provided
Were steps taken to minimise bias and error/increase rigour in data collection?	No/not stated	None provided
Were steps taken to minimise bias and error/increase rigour in data analysis?	Yes	Systematic observation was used to gather data to assess intervention fidelity
Were the findings of the study grounded in/supported by data?	Yes	Direct quotes provided
There was good breadth and/or depth achieved in the findings?	Yes	Breadth and depth, but limited negative case examples
The perspectives of young people were privileged?	Yes	Data collected from participating young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	Medium	
Usefulness of findings	Medium	
Bloomberg et al.⁸⁷		
Stated aim of study	To describe the theory of action that links the Chicano Latino Youth Leadership Institute with PYD and to highlight promising practices	
Methods	<p>Data collection: focus groups; reflection form completed by youth</p> <p>Data analysis: not stated</p>	

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
Details of participants	Country: USA, Minnesota	
	Actual sample: participating youth	
	Sample number: 18	
	Age: 12–17 years	
	Sex: mixed	
	Ethnicity: 'mostly' Mexican Americans and 'a few' Cubans	
	SES: not stated	
Details of PYD programme	Description: Chicano Latino Youth Leadership Institute	
	Target population: Chicano Latino youth age 12–17 years residing in the Region Nine area	
	Theory: not stated	
	Provider(s): prevention and health community division	
	Training: not stated	
	Setting: community	
	Content: support youth participation in planning and implementing a community service project. Provide technical assistance to guide youth graduating and leadership opportunities through conference presentations and specific youth service projects	
	Length/intensity: not stated	
Notes		
<i>Quality appraisal</i>		
Questions used to judge the rigour and usefulness	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	No/not stated	None provided
Were steps taken to minimise bias and error/increase rigour in data collection?	No/not stated	None provided
Were steps taken to minimise bias and error/increase rigour in data analysis?	No/not stated	None provided
Were the findings of the study grounded in/supported by data?	No/not stated	None provided
There was good breadth and/or depth achieved in the findings?	No/not stated	Limited in breadth and depth to answer the RQ
The perspectives of young people were privileged?	Yes	Minimal quotes from young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	Low	
Usefulness of findings	Low	

continued

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations
<i>Bulanda and McCrea⁸⁸</i>	
Stated aim of study	To addresses three central questions: (1) What do disadvantaged African American youth find most valuable about ASP services? (2) How can we understand, given previous research and youths' feedback, the nature of the constructive relationship skills that an ASP can develop in disadvantaged youth? (3) What does the process of developing those constructive relationship skills look like from the youths' perspectives?
Methods	Data collection: (1) youth provide three reasons for joining the programme; (2) roundtable discussion with all students; (3) interviews with pairs of students; (4) participant observations Data analysis: qualitatively to provide an in-depth understanding of the adolescents, their context and their experiences of the programme
Details of participants	Country: USA, Chicago Actual sample: participating youth Sample number: 32 Age: 14–16 years Sex: mixed Ethnicity: African American SES: not stated, but youth residing in poverty-level communities
Details of PYD programme	Description: Stand Up Help Out, a leadership development ASP Target population: African American youth residing in socioeconomically disadvantaged neighbourhoods Provider(s): school social workers and/or graduate students in social work Theory: self-determination theory Training: not but social workers had clinical supervision from a supervisor and training in counselling skills Setting: schools and community Content: training and pay to participate in an apprenticeship in social work. Activities also included college tours, completed resumés, learning about non-violent conflict resolution, mentoring elementary-school children and planned community health and safety fairs. Counselling also available to young people Length/intensity: not stated
Notes	

Quality appraisal

<i>Questions used to judge the rigour and usefulness</i>	<i>Reviewer judgement</i>	<i>Description</i>
Were steps taken to minimise bias and error and increase rigour in sampling?	Cannot tell	Limited description of sampling methods
Were steps taken to minimise bias and error/increase rigour in data collection?	Yes	Use of more than one method; youth as interviewers; ongoing gathering of field notes by programme instructors

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
Were steps taken to minimise bias and error/increase rigour in data analysis?	Yes	Inductive and deductive hierarchy of categories developed to analyse emerging relationships between codes to highlighting the most prominent themes
Were the findings of the study grounded in/supported by data?	Yes	The data presented illuminate/illustrate the findings, although the quotes are not easily identifiable in text
There was good breadth and/or depth achieved in the findings?	Yes	Some breadth, occasional contrasting perspectives, some richness
The perspectives of young people were privileged?	Yes	Yes as interviewers and interviewees
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	High	
Usefulness of findings	Medium	
Cross et al.⁸⁹ (linked studies: Gottfredson et al.⁹¹)		
Methods	Data collection: documents – attendance records and employment records; non-participant programme and student observation, youth self-report of quality of experiences measured with the Youth Experiences Survey	
	Data analysis: site comparison analysis	
Stated aim of study	To evaluate how dimensions of implementation (e.g. level of dosage, quality of management and climate, participant responsiveness and staffing quality achieved at the five programme sites co-varied with self-reported positive experiences after school)	
Details of participants	Country: USA; urban east coast school district	
	Actual sample: programme staff, participating youth	
	Sample number: site observation included the following visits: A, C and D, $n = 16$ times; B, $n = 18$ times; E, $n = 14$ times	
	Age: not stated	
	Sex: mixed	
	Ethnicity: not stated	
	SES: not stated	
Details of PYD programme	Description: All Stars prevention curriculum: an enhanced ASP	
	Target population: pupils aged 11–14 years from five under-performing middle schools	
	Theory: not stated	
	Provider(s): a county-level government agency that specialising in providing recreation and leisure activities for youths	
	Training: 6 hours in tutoring model, 19 hours in running ASP, 3 days in All Stars	
	Setting: middle schools	
	Content: leisure activities (e.g. fitness activities, board games, arts and crafts, field trips, computer projects or computer free time, service learning, workforce skills and holiday or other special event celebrations)	
	Length/intensity: 3 days per week, for 3 hours	
		<i>continued</i>

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (continued)

Study details	Characteristics of process evaluations	
Notes		
<i>Quality appraisal</i>		
Questions used to judge the rigour and usefulness	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	Yes	They visited and observed all the sites in the programme
Were steps taken to minimise bias and error/increase rigour in data collection?	Yes	Use of multiple data collection methods
Were steps taken to minimise bias and error/increase rigour in data analysis?	Yes	Reflective account of difficulty with analysing implementation and the attempts made to do achieve this using a 'relative' approach
Were the findings of the study grounded in/supported by data?	Yes	The data presented illuminate/illustrate the findings
There was good breadth and/or depth achieved in the findings?	Yes	Yes, good breadth and depth
The perspectives of young people were privileged?	Cannot tell	They survey young people but uncertain if there is an opportunity to provide more in-depth responses via open-ended questions or if sites observers spoke directly to young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	High	
Usefulness of findings	High	
Lee et al.⁹⁴		
Methods	Data collection: interviews, document analysis, retrospective weekly diary of the Unit's co-ordinator	
	Data analysis: extracting and summarising recurrent themes by one investigator reanalysed by a second when uncertain	
Stated aim of study	To evaluate a community-driven initiative established to prevent substance misuse and increase respect for culture and their elders among young people in a group of remote Aboriginal communities in Arnhem Land, Northern Territory (NT), Australia	
Details of participants	Country: Australia	
	Actual sample: programme staff, participating youth	
	Sample number: 73 including representatives from each Aboriginal community council and the Land Council	
	Age: not stated	
	Sex: mixed	
	Ethnicity: 30 Indigenous and 43 non-Indigenous respondents	
	SES: not stated	

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
Details of PYD programme	Description: a community-driven youth initiative Target population: all young people in the community Provider(s): programme co-ordinator, case worker and Indigenous youth worker Theory: not stated Training: not stated Setting: community Content: activities include technical studies, film-making, drivers' licence programme, marine debris clean-up programme, youth leadership opportunities, recruitment, support and training for new employees in local youth recreational sector, youth and community festivals, mural painting programme, sporting carnivals, cultural knowledge programme, discos, mental health promotion poster programme, health promotion Length/intensity: not stated	
Notes		
<i>Quality appraisal</i>		
Questions used to judge the rigour and usefulness	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	Cannot tell	Unclear, other than convenience sample of a range of key informants
Were steps taken to minimise bias and error/increase rigour in data collection?	Cannot tell	Described in sufficient depth to be able to conclude that it was done rigorously
Were steps taken to minimise bias and error/increase rigour in data analysis?	Not stated	Insufficient detail to judge
Were the findings of the study grounded in/supported by data?	Not stated	Minimal reporting of participant data in the form of quotes
There was good breadth and/or depth achieved in the findings?	No	Neither breadth nor depth
The perspectives of young people were privileged?	No	No, despite interviewing young people, their 'voice' is not clearly represented in the data; none of the quotes appears to come from young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	Low	
Usefulness of findings	Low	
Maxfield et al.⁹⁹		
Methods	Data collection: observational site visits Data analysis: programme component analysis	
Stated aim of study	To evaluate programme implementation by site	

continued

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations
Details of participants	<p>Country: USA, Ohio; Texas; Tennessee; Washington, DC; Pennsylvania; Washington</p> <p>Actual sample: programme staff</p> <p>Sample number: not stated</p> <p>Age: not stated</p> <p>Sex: mixed</p> <p>Ethnicity:</p> <p>SES: not stated</p>
Details of PYD programme	<p>Description: QOP</p> <p>Target population: youth who met the following criteria: (1) began the ninth grade with a dropout rate of 40% or more; (2) grade point average below the 67th percentile of entering ninth graders; (3) not repeating the ninth grade; (4) not so physically disabled or learning disabled that participation in programme is inappropriate</p> <p>Theory: youth development model</p> <p>Provider(s): case managers and mentors, community-based organisations</p> <p>Training: unclear</p> <p>Setting: schools and community-based organisations</p> <p>Content: intensive case management, mentoring and educational, developmental, cultural and recreational, and community-based activities. <i>Education services</i>: academic assessment, an individual education plan, one-on-one tutoring and computer-assisted instruction in specific coursework and basic reading and mathematics. Visiting nearby college campuses and other activities designed to promote awareness of and planning for college. <i>Developmental activities</i>: life-skills training, employment-readiness training, cultural awareness and recreation. <i>Community service activities</i>, e.g. visiting the residents of a local nursing home or volunteering at a neighbourhood food bank. Cultural and recreational activities: movies, ice skating, bowling, swimming, sailing, golfing, mountain biking, amusement/water parks, haunted houses, board/computer games, local fairs, picnics, attending sporting events, pizza lunches, dinners in restaurants</p> <p>Length/intensity: 750 hours per year, or > 14 hours per week on average throughout the year for up to 5 years</p>
Notes	

Quality appraisal

Questions used to judge the rigour and usefulness	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	Cannot tell	Limited description of which sites were sampled and which staff were spoken to
Were steps taken to minimise bias and error/increase rigour in data collection?	Cannot tell	Limited description of data collection to judge
Were steps taken to minimise bias and error/increase rigour in data analysis?	Cannot tell	Limited description of data collection to judge
Were the findings of the study grounded in/supported by data?	No/not stated	Lack of direct quotes

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
There was good breadth and/or depth achieved in the findings?	Yes	Breadth and depth
The perspectives of young people were privileged?	No	Young people not included in the design or collection of data
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	Low	
Usefulness of findings	High	
Schwartz et al.⁷⁷ (linked studies: Bloom et al.¹⁰⁶)		
Methods	Data collection: semistructured interviews	
	Data analysis: thematic analysis using NVivo	
Stated aim of study	To explore the nature of the relationships and the process through which enduring mentoring relationships may have influenced youth outcomes	
Details of participants	Country: USA (sites across the country)	
	Actual sample: participating youth	
	Sample number: 30 participants from 3 of the 10 sites	
	Age: 20–22 years	
	Sex: 90% male	
	Ethnicity: white 60%; Latino 20%; white and Latino 7%; black 7%; white and Native American 3%; unknown 3%	
	SES: not stated	
Details of PYD programme	Description: NGYCP YIM programme	
	Target population: youth ages 16–18 years who have dropped out or been expelled from school, who are drug-free at the time of entry into the programme and not currently on probation or parole for anything beyond juvenile status offences, not serving time or awaiting sentencing, not under indictment or charged and not convicted of a felony or capital offence	
	Theory: military training model that believes incorporating caring relationships with non-parental adults can contribute to a range of PYD outcomes	
	Provider(s): National Guard	
	Training: unclear	
	Setting: community military-style boot camp	
	Content: the Pre-ChalleNGe Phase is a 2-week period of orientation and assessment in which young people adjust an intensive, structured lifestyle required at the programme site. The residential phase is a 20-week period during which youth are working towards their high school diploma or GED and take classes on life skills, health and job skills, while also participating in other activities such as physical training, sports, leadership and citizenship activities and community service. The post-residential phase is characterised by a post-residential action plan in which youth identify specific post-residential activities (e.g. GED programme, community college, vocational training, a job, or military service)	
	Length/intensity: 5 months' full-time residential, 1 year's post-residential with job placement and structured mentoring	
continued		

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (continued)

Study details	Characteristics of process evaluations	
Notes		
<i>Quality appraisal</i>		
Questions used to judge the rigour and usefulness	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	Yes	Well-reasoned and justified with attempts made to obtain a diverse sample of the mentors
Were steps taken to minimise bias and error/increase rigour in data collection?	Yes	Semi-structured, with participants given the opportunity to ask questions and open-ended questions used to draw out experiences and stories from participants, with follow-up questions tailored to participants' individual
Were steps taken to minimise bias and error/increase rigour in data analysis?	Yes	Line-by-line coding to allow as part of thematic analysis
Were the findings of the study grounded in/supported by data?	Yes	Sufficient quotes to support analysis
There was good breadth and/or depth achieved in the findings?	Yes	Sufficient breadth and depth
The perspectives of young people were privileged?	Yes	Interviews with young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	High	
Usefulness of findings	High	
Wiggins et al.¹¹²		
Methods	Data collection: monthly project monitoring records for all 27 YPDP projects; questionnaires with YPDP young people; YPDP staff questionnaires (annual) and focus groups (once); comparison site staff questionnaires (twice); observation of training sessions, meetings and conferences	
	Data analysis: thematic content analysis	
Stated aim of study	To evaluate was the processes of planning and providing PYD services to young people	
Details of participants	Country: UK	
	Actual sample: programme staff, participating youth	
	Sample number: all pilot projects, a subsample of 7 case study sites and 4 comparison group case-study sites	
	Age: not stated (for process evaluation)	
	Sex: mixed	
	Ethnicity: not stated (for process evaluation)	
	SES: not stated (for process evaluation)	

TABLE 16 Characteristics of process evaluations (studies included in the review to answer RQ2) (*continued*)

Study details	Characteristics of process evaluations	
Details of PYD programme	<p>Description: YPDP</p> <p>Target population: young people at risk of teenage pregnancy, substance use or school drop out</p> <p>Provider(s): youth service providers</p> <p>Theory: youth development model</p> <p>Training: training of volunteers on programme requirements, recognising child abuse, working with youth; monthly supervision</p> <p>Setting: community youth services</p> <p>Content: activities focusing on young people's health and education as well as their broader social development with specific programme content determined by the individual projects delivering services. These could include: education (literacy, numeracy, IT, vocational skills). training/employment opportunities life skills (e.g. communication, decision-making, goal-setting, relationships, negotiation, anger-management), mentoring (weekly one-to-one sessions with staff), volunteering (both career-oriented and community-based), health education (particularly sexual health, substance misuse) arts and sports, advice on accessing services (health, contraceptive, drug and alcohol services, welfare, benefits advice, counselling and advice, housing)</p> <p>Length/intensity: 6–10 hours' weekly provision for 1 year</p>	
Notes		
<i>Quality appraisal</i>		
Questions used to judge the rigour and usefulness	Reviewer judgement	Description
Were steps taken to minimise bias and error/increase rigour in sampling?	Yes	Representative sample of programme sites
Were steps taken to minimise bias and error/increase rigour in data collection?	Yes	Multimethod approach taken
Were steps taken to minimise bias and error/increase rigour in data analysis?	Yes	Clear, established and transparent methods for data analysis (e.g. thematic content analysis)
Were the findings of the study grounded in/supported by data?	Yes	Quotes to support analysis
There was good breadth and/or depth achieved in the findings?	Yes	Breadth and depth
The perspectives of young people were privileged?	Yes	Data drawn from young people
<i>Overall reliability and usefulness of findings</i>		
Reliability of findings	High	
Usefulness of findings	High	
YIM, youth-initiated mentoring.		

TABLE 17 Quality assessment of process evaluations

Study ^a	Sample ^b		Data collection ^c		Data analysis ^d		Findings grounded ^e		Breadth/Depth ^f		Privilege ^g		Overall reliability		Overall usefulness	
	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD
Armstrong and Armstrong ⁸⁴	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baker <i>et al.</i> ⁸⁵	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Berg <i>et al.</i> ⁸⁶	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bloomberg <i>et al.</i> ⁸⁷	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Bulanda and McCrea ⁸⁸	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Cross <i>et al.</i> ⁸⁹	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lee <i>et al.</i> ⁹⁴	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Maxfield <i>et al.</i> ⁹⁹	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Schwartz <i>et al.</i> ⁷⁷	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Wiggins <i>et al.</i> ¹¹²	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

✓, no; ✓, yes; A, high; B, medium; C, low.

a All studies were coded between CB and KD except for Wiggins *et al.*,¹¹² which was coded by SM and KD.

b Refers to the rigour of sampling.

c Refers to the rigour of the data collection.

d Refers to the rigour of the data analysis.

e Refers to the extent to which the study findings are grounded in the data.

f Refers to the breadth and depth of findings.

g Refers to whether or not the study privileges the perspectives of participants.

Appendix 15 Application of theory synthesis to descriptions of Positive Youth Development interventions included in the process and outcome evaluations

TABLE 18 Application of theory synthesis to descriptions of Positive Youth Development interventions included in the process and outcome evaluations

Area of theoretical literature from which characteristics derived	Developmental assets				Social Learning Theory	Socioecological theory	Expectation state theory	Critical theories
	Is there a focus on thriving?	Is there reference to multiple, specific assets?	Is there a breadth of activities offered?	Can participants pursue some activities in greater depth?	Do participants learn skills from observing adults and/or peers?	Is there intervention to improve participants' wider environments?	Do intervention activities explicitly address risk behaviours	Is there explicit reference to transforming society or challenging inequalities?
Outcome evaluations								
All Stars (Cross <i>et al.</i> , ⁹⁰ Gottfredson <i>et al.</i> ^{91,92})	0	0	1	0	0	0	1	0
MAP (Gottfredson <i>et al.</i> ⁹³)	0	0	1	0	0	0	1	0
Cool Girls, Inc. (Kuperminc <i>et al.</i> ⁹⁵)	1	1	1	0	1	0	1	0
BBBS (Tierney ⁹⁸)	0	0	0	0	0	0	1	0
QOP (Rodriguez-Planas ⁷⁶)	1	0	0	0	1	0	0	0
Stay SMART (St Pierre and Kaltreider ¹¹¹)	1	0	0	0	1	0	1	0
PYDC (Tebes <i>et al.</i> ⁷⁹)	1	0	1	0	0	0	1	0
Process and outcome evaluations								
YARP (Berg <i>et al.</i> ⁸⁶)	1	0	0	1	1	1	1	1
NGYCP (Millenky <i>et al.</i> ^{107,108})	1	0	1	0	0	1	1	0
YPDP (Wiggins <i>et al.</i> ^{46,112})	1	1	1	1	1	1	1	0

0, not reported as present; 1, reported as present.

Social control theory									
Multiple theories									
Is there reference to conformity, or contribution to society as it stands?	Are there leadership opportunities?	Do participants have some choice in which activities to pursue?	Do adult providers aim to develop affective rather than merely instrumental relationships with participants?	Do participants develop enduring (at least 1 year) relationships with adult providers?	Are positive behaviours rewarded or celebrated?	Is there explicit reference to hope or optimism in a young person's potential?	Is there explicit reference to family like atmosphere or safe space?	Do participants carry out service in the community?	
0	0	1	0	0	0	0	0	1	4
1	0	1	0	0	0	0	0	0	4
0	0	1	0	1	0	1	0	0	7
1	0	1	1	1	0	1	0	1	7
0	0	1	1	1	0	0	0	0	5
0	1	0	1	1	0	0	0	1	7
0	1	0	0	0	0	0	0	1	5
0	1	1	0	0	0	1	0	1	10
0	1	1	1	1	0	0	0	1	7
0	0	1	1	0	0	0	0	1	10

Appendix 16 Characteristics and quality risk of bias of outcome evaluations

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3)

Study details	Characteristics of outcome evaluations
Berg et al.⁸⁶	
Methods	Non-randomised trial: matched comparison group
Unit of allocation	Individual
Participants' details	Country: USA Sample number: 114 (IG); 202 (CG) Sex: 51% female, 49% male in intervention group Ethnicity: 47% African American/black; 41% Latino; 12% mixed/white/newly arrived SES: not stated
Intervention details	Description: YARP Targeted/universal: targeted Target population: youth aged 14 to 16 years Provider(s): prevention research educators Theory: prevention, cognitive and critical theory including: (1) ecological theory; (2) identity theory; (3) learning and instructional theories (including social learning, social construction and multiple intelligence); and (4) critical, transformative theories including voice, empowerment and action research Training: educators are trained in action research methods and empowerment-oriented, social construction facilitation skills Setting: community-based after-school and summer programme Content: participatory action research involving formative community ethnography (e.g. training youth to identify adolescent risk behaviours, develop a collective action plan and carry out activities) as a group, including using research to understand their community better Length/intensity: 4 hours/day for 8 weeks, 8-month implementation period for action plan Control: measurement
Outcomes	At post intervention (less than unity favours intervention): Frequency of marijuana use in past 30 days: $r = -0.12$; p -value = 0.053 Frequency of alcohol use in past 30 days: decrease but not significant

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (continued)

Study details	Characteristics of outcome evaluations	
<i>Risk of bias</i>		
Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Not applicable	This was a non-randomised trial
Allocation concealment: was the allocation adequately concealed?	Not applicable	This was a non-randomised trial
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	This was a non-randomised trial
Incomplete outcome data: were complete data for each outcome reported, and, if not were adequate reasons for incomplete outcome data provided?	No	By outcome: reasons for dropout not reported, and all drug-use outcomes not reported 17.4% attrition overall Reasons for dropout not reported; however, authors report no differences between groups on dropout characteristics
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Yes	The evaluation discussed collecting outcome data on a variety of substance use measures, but only one is presented in sufficient detail to calculate an effect size
Key confounders: Were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Yes	Intervention and control samples matched, and differences between groups used as covariates in analyses
Clustering: was clustering of participants accounted for in the analysis?	No	No evidence of accounting for clustering
Other source of bias: were there any other sources of bias that might affect the results of the study?	No	
Cross et al.⁹⁰ (linked studies: Gottfredson et al.^{91,92})		
Methods	RCT	
Unit of allocation	Individual	
Participants' details	Country: USA	
		Sample number: 224 (IG); 223 (CG); analysed: 211 (IG); 205 (CG)
		Sex: 48% female, 52% male
		Ethnicity: 71% African American/black; 17% white; 8% multiracial
		SES: 58% received subsidised school meals

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
Intervention details	<p>Description: the All Stars prevention curriculum: an enhanced ASP</p> <p>Targeted/universal: universal</p> <p>Target population: pupils aged 11–14 years from five under-performing middle schools</p> <p>Theory: not stated</p> <p>Provider(s): a county-level government agency that specialising in providing recreation and leisure activities for youths</p> <p>Training: 6 hours in tutoring model, 19 hours in running ASP, 3 days in All Stars</p> <p>Setting: middle schools</p> <p>Content: leisure activities (e.g. fitness activities, board games, arts and crafts, field trips, computer projects or computer free time, service learning, workforce skills and holiday or other special event celebrations)</p> <p>Length/intensity: 3 days per week, for 3 hours</p> <p>Control: monthly 'fun activity' at programme</p>	
Outcomes	<p>At post intervention (less than unity favours treatment):</p> <p>Any use of alcohol, cigarettes or marijuana in the past month: OR 1.03; <i>p</i>-value = 0.91</p> <p>Smoking initiation: OR 0.95; <i>p</i>-value = 0.90</p> <p>Drinking initiation: OR 1.10; <i>p</i>-value = 0.75</p> <p>Marijuana initiation: OR 0.72; <i>p</i>-value = 0.47</p> <p>Inhalant initiation: OR 0.73; <i>p</i>-value = 0.17</p> <p>Other drug initiation: OR 1.73; <i>p</i>-value = 0.40</p>	
<i>Risk of bias</i>		
Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Yes	Random number generator used
Allocation concealment: was the allocation adequately concealed?	Unclear	No information provided
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	Unclear	No information provided
Incomplete outcome data: were complete data for each outcome reported, and, if not were adequate reasons for incomplete outcome data provided?	No	Sample restricted to post-test questionnaire completers; dropout characteristics analysed but missing data on post-test questions as well 7% attrition overall
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Unclear	We could not assess this with the given information
		continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (continued)

Study details	Characteristics of outcome evaluations
Key confounders: Were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Not applicable RCT
Clustering: was clustering of participants accounted for in the analysis?	Yes Analysis controlled for site
Other source of bias: were there any other sources of bias that might affect the results of the study?	Yes Unbalanced missing and differential follow-up patterns are present
Gottfredson et al.⁹³	
Methods	RCT/non-randomised trial
Unit of allocation	Individual and group
Participants' details	Country: USA Sample number: 239 in IG?, 201 in CG? (only older children) Sex: 44% female, 56% male in IG; 50% female, 50% male in CG Ethnicity: 76% non-white in IG, 53% in CG SES: no information
Intervention details	Description: Maryland After School Community Grant Program Targeted/universal: universal Target population: elementary- and middle-school children Theory: not stated Provider(s): ASP providers Training: 6 hours in tutoring model, 19 hours in running ASP, 3 days in All Stars Setting: public schools, community centres Content: three activity areas: (1) academic assistance; (2) social skill; or (3) character development, and recreational/leisure activities aimed at retaining young people in the programme. Recreational activities mainly included: sports, arts and crafts, in addition to specialty activities, such as entrepreneurial activities, karate, sailing, or soccer Length/intensity: 3 days/week for 90 sessions (middle school), 4 days/week for 120 sessions (elementary school); average of 3 hours/session Control: waiting list or community-recruited comparison
Outcomes	At post intervention: Past year variety of drug use: IG: pre-test 0.036 (SD 0.119, $n = 234$) post-test 0.038 (SD 0.127, $n = 222$) CG: pre-test 0.053 (SD 0.146, $n = 199$) post-test 0.086 (SD 0.198, $n = 173$)

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
<i>Risk of bias</i>		
Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Not applicable	Non-randomised trial
Allocation concealment: was the allocation adequately concealed?	Not applicable	Non-randomised trial
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	Non-randomised trial
Incomplete outcome data: were complete data for each outcome reported, and, if not were adequate reasons for incomplete outcome data provided?	No	Dropout characteristics analysed, uneven attrition and inadequate treatment of missingness (e.g. imputed means were not reported)
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Unclear	We could not assess this using the information provided
Key confounders: Were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Yes	Differences between conditions tested across randomisation and non-random assignment
Clustering: was clustering of participants accounted for in the analysis?	No	No evidence of accounting for clustering
Other source of bias: were there any other sources of bias that might affect the results of the study?	No	
Kuperminc et al.⁹⁵		
Methods	Non-randomised trial	
Unit of allocation	Individual	
Participants' details	Country: USA	
	Sample number: 86 (IG), 89 (CG) (additionally, 29 girls assigned to mentors)	
	Sex: 100% female	
	Ethnicity: 76% non-white in IG, 53% in CG	
	SES: 80–96% of students attending programme schools lived in economic disadvantage	

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
Intervention details	<p>Description: Cool Girls, Inc.</p> <p>Targeted/universal: universal</p> <p>Target population: young girls aged 9 to 15 years from disadvantaged and low-income communities</p> <p>Theory: not stated</p> <p>Provider(s): community-based organisation</p> <p>Training: not stated</p> <p>Setting: public schools</p> <p>Content: three main components: (1) Girls Club; a comprehensive life skills curriculum covering a range of topics (e.g. positive sexual health, hygiene, conflict resolution, self-esteem and cultural awareness); (2) Cool Scholars: providing homework assistance, individual tutoring and supporting girls to complete projects (e.g. making presentations, journaling and participating in academic tournaments). Participants are also eligible for (3) 'Cool Sisters', a one-to-one mentoring programme, after actively taking part in the programme for 1 year. Additional components include: weekend workshops (e.g. computer skills, financial literacy and career development); field trips (e.g. tours of the Cable News Network, the Weather Channel and visits to local museums); special events; and summer programmes</p> <p>Length/intensity: weekly meetings during the school year</p> <p>Control: comparators nominated by participants and schools</p>	
Outcomes	<p>At post intervention (greater than unity favours intervention):</p> <p>Avoidance of drug use (cigarettes, alcohol, marijuana, cocaine, or inhalants) in past 6 months: OR 0.62 (ns) for full sample; OR 1.53 (ns) for mentored</p>	
<i>Risk of bias</i>		
Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Not applicable	Non-randomised trial
Allocation concealment: was the allocation adequately concealed?	Not applicable	Non-randomised trial
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	Non-randomised trial
Incomplete outcome data: were complete data for each outcome reported, and, if not were adequate reasons for incomplete outcome data provided?	Yes	Dropout characteristics analysed, uneven attrition and inadequate treatment of missingness but multiple imputation carried out

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Unclear	Unable to assess based on provided information
Key confounders: were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Yes	Differences between conditions tested using covariates in regression
Clustering: was clustering of participants accounted for in the analysis?	Yes	Cluster effects were small and not statistically significant
Other source of bias: were there any other sources of bias that might affect the results of the study?	No	
Rhodes et al.⁹⁶ (linked studies: Grossman and Tierney;⁹⁷ Tierney⁹⁸)		
Methods	Randomised trial	
Unit of allocation	Individual	
Participants' details	Country: USA	
	Sample number: 487 (IG), 472 (CG)	
	Sex: about 25% female, 75% male	
	Ethnicity: 45% white; 42% black; 14% Hispanic	
	SES: 37% welfare recipients, 36% incomes below poverty level	
Intervention details	Description: BBBS	
	Targeted/universal: targeted	
	Target population: young people between 5 and 18 years of age, with minimal social skills, who live in an agency catchment area, with priority given to young people who only have one parent engaged in their life	
	Theory: matched youth mentoring; based on sex and various other factors such as shared interest, reasonable geographic proximity and same ethnicity	
	Provider(s): trained community volunteers	
	Training: initial training on abuse and programme rules, additional training often provided on youth development, monthly supervision	
	Setting: community	
	Content: 1 : 1 weekly mentoring aimed at developing the 'whole person', with the mentoring relationship is seen is the mechanism that enables the mentor to support the mentee transition from childhood and/or adolescence into adulthood	
	Length/intensity: weekly meetings with mentor	
	Control: unmatched children	

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
Outcomes	At post intervention:	
	Initiation of drug abuse: CG mean 11.47%, IG less by 45.8% ($p < 0.05$)	
	Initiating alcohol use: CG mean 26.72%, IG less by 27.4% ($p < 0.10$)	
	Cigarette smoking: CG mean 17.2%, IG less by 19.7% ($p > 0.10$)	
	Number of times hit someone: CG mean 2.68, IG less by 0.85 ($p < 0.05$)	
	Number of times involved in a fight: CG mean 1.54, IG less by 0.02 ($p > 0.10$)	
<i>Risk of bias</i>		
Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Unclear	Not enough information was provided
Allocation concealment: was the allocation adequately concealed?	Yes	Allocation done by an external survey subcontractor
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	Not possible in this intervention
Incomplete outcome data: were complete data for each outcome reported and, if not, were adequate reasons for incomplete outcome data provided?	No	Only people with all measurement occasion included
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Unclear	Could not assess based on provided information
Key confounders: Were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Not applicable	Randomised trial
Clustering: was clustering of participants accounted for in the analysis?	No	No evidence of accounting for clustering
Other source of bias: were there any other sources of bias that might affect the results of the study?	No	

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations
	Rodriguez-Planas⁷⁸ (linked studies: Maxfield,^{99,100} Rodriguez-Planas;¹⁰¹ Schirm et al.;¹⁰³ Schirm and Rodriguez-Planas;¹⁰⁴ Schirm and McKie¹⁰⁵)
Methods	RCT
Unit of allocation	Individual
Participants' details	Country: USA Sample number: 580 (IG), 489 (CG) Sex: IG: 52.2% male, 47.8% female; CG: 55.8% male, 44.2% female Ethnicity: IG: 26.2% Hispanic, 68.3% Black, CG: 25.7% Hispanic, 67.9% Black SES: not reported
Intervention details	Description: QOP Targeted/universal: targeted Target population: youth who met the following criteria: (1) began the ninth grade with a dropout rate of 40% or more; (2) grade point average below the 67th percentile of entering ninth graders; (3) not repeating the ninth grade; (4) not so physically disabled or learning disabled that participation in programme is inappropriate Theory: youth development model Provider(s): case managers and mentors, community-based organisations Training: unclear Setting: schools and community-based organisations Content: intensive case management, mentoring and educational, developmental, cultural and recreational and community-based activities. <i>Education services</i> : academic assessment, an individual education plan, one-on-one tutoring and computer-assisted instruction in specific coursework and basic reading and mathematics. Visiting nearby college campuses and other activities designed to promote awareness of and planning for college. <i>Developmental activities</i> : life-skills training, employment-readiness training, cultural awareness and recreation. <i>Community service activities</i> (e.g. visiting the residents of a local nursing home or volunteering at a neighbourhood food bank). Cultural and recreational activities: movies, ice skating, bowling, swimming, sailing, golfing, mountain biking, amusement/water parks, haunted houses, board/computer games, local fairs, picnics, attending sporting events, pizza lunches, dinners in restaurants Length/intensity: 750 hours per year, or over 14 hours per week on average throughout the year for up to 5 years Control: no treatment

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations
Outcomes	<p>In the fourth year of the demonstration programme (i.e. post intervention):</p> <p>Near the end of the fourth academic year:</p> <p>Drinking in past 30 days: 40% (IG) vs. 33% (CG)</p> <p>Frequent drinking in past 30 days: 11% (IG) vs. 11% (CG)</p> <p>Binge drinking in past 30 days: 24% (IG) vs. 20% (CG)</p> <p>Frequent binge drinking in past 30 days: 7% (IG) vs. 5% (CG)</p> <p>Drunk or high at school in past 12 months: 20% (IG) vs. 20% (CG)</p> <p>Used an illegal drug in past 30 days: 34% (IG) vs. 28% (CG)</p> <p>Involved in gang fight in past 12 months: 16% (IG) vs. 14% (CG)</p> <p>Telephone survey 7 months later:</p> <p>Binge drinking in past 30 days: 19% (IG) vs. 23% (CG)</p> <p>Frequent binge drinking in past 30 days: 5% (IG) vs. 4% (CG)</p> <p>Used an illegal drug in past 30 days: 16% (IG) vs. 19% (CG)</p> <p>3 to 4 years post intervention:</p> <p>Binge drinking in past 30 days: 25% (IG) vs. 31% (CG)</p> <p>Binge drinking on eight or more days in past 30 days: 7% (IG) vs. 5% (CG)</p> <p>Used an illegal drug in past 30 days: 12% (IG) vs. 18% (CG)</p> <p>Approximately 6 years post intervention:</p> <p>Smoked cigarettes or used tobacco in past 30 days: 34% (IG) vs. 34% (CG)</p> <p>Smoked cigarettes or used tobacco daily in past 30 days: 22% (IG) vs. 24% (CG)</p> <p>Binge drinking in past 30 days: 31% (IG) vs. 31% (CG)</p> <p>Binge drinking on eight or more days in past 30 days: 8% (IG) vs. 6% (CG)</p> <p>Used an illegal drug in past 30 days: 12% (IG) vs. 13% (CG)</p>

Risk of bias

Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Unclear	No information provided
Allocation concealment: was the allocation adequately concealed?	Unclear	No information provided
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	Impossible given the intervention
Incomplete outcome data: were complete data for each outcome reported and, if not, were adequate reasons for incomplete outcome data provided?	Yes	Weighted analysis partially accounts for missingness
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Unclear	Could not assess given information

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations
Key confounders: Were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Not applicable Randomised trial
Clustering: was clustering of participants accounted for in the analysis?	Yes Analysis controlled for site
Other source of bias: were there any other sources of bias that might affect the results of the study?	No
Schwartz et al.⁷⁷ (linked studies: Millenky et al.,¹⁰⁷⁻¹⁰⁹ Perez-Arce et al.¹¹⁰)	
Methods	RCT
Unit of allocation	Individual
Participants' details	Country: USA Sample number: 722 (IG), 451 (CG) Sex: 88.0% male, 12.0% female Ethnicity: 18.1% Hispanic; 42.3% white; 33.8% black; 5.7% other SES: 26.4% of families receive public assistance
Intervention details	Description: NGYCP YIM programme Targeted/universal: targeted Target population: youth 16 to 18 years old who have dropped out or been expelled from school, who are drug free at the time of entry into the programme and not currently on probation or parole for anything beyond juvenile status offences, not serving time or awaiting sentencing, not under indictment or charged and not convicted of a felony or capital offence Theory: military training model that believes incorporating caring relationships with non-parental adults can contribute to a range of PYD outcomes Provider(s): National Guard Training: unclear Setting: community military-style boot camp Content; The Pre-ChalleNGe Phase is a 2-week period of orientation and assessment in which young people adjust to an intensive, structured lifestyle required at the programme site. The residential phase is a 20-week period during which youth are working towards their high school diploma or GED and take classes on life skills, health and job skills, while participating in other activities such as physical training, sports, leadership and citizenship activities and community service. The post-residential phase is characterised by a post-residential action plan in which youth identify specific post-residential activities (e.g. GED programme, community college, vocational training, a job, or military service) Length/intensity: 5 months' full-time residential, 1 year's post-residential with job placement and structured mentoring Control: waitlist

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
Outcomes	<p>At post intervention:</p> <p>Charged with a violent crime: 3.4% (IG) vs. 3.6% (CG)</p> <p>Convicted of a violent crime: 1.4% (IG), 1.2% (CG)</p> <p>Any violent incidents: 54.0% (IG) vs. 57.3% (CG)</p> <p>Number of violent incidents: 0.9 (IG) vs. 1.3 (CG)</p> <p>Binge drinking in past 14 days: 2.8% (IG) vs. 4.7% (CG)</p> <p>Frequent marijuana use in past 12 months: 22.5% (IG) vs. 25.2% (CG)</p> <p>Ever used other illegal drugs (LSD, cocaine, methamphetamine, amphetamine sulphate, barbiturates, heroin, PEDs): 24.1% (IG) vs. 23.1% (CG)</p> <p>Frequent illegal drug use in past 12 months: 5.6% (IG) vs. 4.4% (CG)</p> <p>Charged with a drug crime: 2.9% (IG) vs. 5.3% (CG)</p> <p>Convicted of a drug crime: 1.4% (IG) vs. 1.9% (CG)</p> <p>At 18 months post intervention:</p> <p>Convicted of a violent crime: 2.1% (IG) vs. 2.3% (CG)</p> <p>Any violent incidents: 48.7% (IG) vs. 44.5% (CG)</p> <p>Number of violent incidents: 0.9% (IG) vs. 0.8% (CG)</p> <p>Binge drinking in past 14 days: 26.1% (IG), 30.2% (CG)</p> <p>Frequent marijuana use in past year: 26.0% (IG), 24.4% (CG)</p> <p>Ever used other illegal drugs (LSD, cocaine, methamphetamine, amphetamine sulphate, barbiturates, heroin, PEDs): 28.2% (IG) vs. 23.2% (CG)</p> <p>Frequent illegal drug use in past 12 months: 4.7% (IG) vs. 4.2% (CG)</p> <p>Convicted of a drug crime: 8.1% (IG) vs. 5.9% (CG)</p>	
<i>Risk of bias</i>		
Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Unclear	No information provided
Allocation concealment: was the allocation adequately concealed?	Unclear	No information provided
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	Unclear	No information provided
Incomplete outcome data: were complete data for each outcome reported and, if not, were adequate reasons for incomplete outcome data provided?	Yes	Weighted analysis partially accounts for missingness, response bias tested
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Unclear	We could not assess this question given the information provided

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations
Key confounders: were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Not applicable Randomised trial
Clustering: was clustering of participants accounted for in the analysis?	Yes Analysis accounted for clustering by site
Other source of bias: were there any other sources of bias that might affect the results of the study?	No
<i>St Pierre and Kaltreider</i>¹¹	
Methods	Non-randomised trial
Unit of allocation	Group
Participants' details	Country: USA Sample number: 52 (IG), 54 (IG with booster), 55 (CG) Sex: about 25% female, 75% male Ethnicity: 45% white; 42% black; 14% Hispanic SES: 37% welfare recipients, 36% incomes below poverty level
Intervention details	Description: Stay SMART is a component of the SMART moves National prevention programme of Boys and Girls Clubs of America Targeted/universal: universal Target population: 13-year-old members of Boys and Girls Clubs of America Theory: personal and social competence approach Provider(s): Boys and Girls Clubs employees Training: compulsory training on Stay SMART and yearly training of group leaders on SMART Leaders component Setting: local 'clubhouses' for children Content: life skills training programme includes structured small group sessions on: gateway drugs; decision-making; advertising; self-image and self-improvement; coping with change; communication skills; social skills (meeting and greeting people; and boy meets girl); assertiveness; relationships; life planning skills. The booster programme SMART Leaders added leadership sessions on: (1) orientation; improving self-image; coping with stress; resisting media pressures; being assertive in pressure situations and (2): resisting alcohol; other drugs; and early sexual activity. Sessions included culturally relevant experiential activities and videos. The video format introducing the session objectives; gave background information; and discussion questions on the video. Participants were also encouraged after sessions to become involved in other prevention activities (e.g. wearing drug-free t-shirt; being drug-free role models; helping with general club activities; helping with specific drug-prevention activities) Length/intensity: 12 sessions in 3 months; 8 sessions over 2 years added in booster sessions Control: Boys and Girls Clubs without a prevention programme

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (continued)

Study details	Characteristics of outcome evaluations	
Outcomes	<p>At post intervention for binary substance-use outcomes:</p> <p>Alcohol behaviour: not reported</p> <p>Marijuana behaviour: not reported</p> <p>Cigarette behaviour: not reported</p> <p>At 12 months post intervention for binary substance-use outcomes:</p> <p>Alcohol behaviour: not reported</p> <p>Marijuana behaviour: IG2 vs. CG: $b = 0.80$; p-value = 0.19</p> <p>Cigarette behaviour: IG2 vs. CG: p-value = 0.91; IG2 vs. CG: p-value = 0.54</p> <p>At 24 months post intervention for binary substance-use outcomes:</p> <p>Alcohol behaviour: not reported</p> <p>Marijuana behaviour: IG2 vs. IG: $b = 1.09$; p-value = 0.12; IG2 vs. CG: $b = 1.23$; p-value = 0.09</p> <p>Cigarette behaviour: IG2 vs. IG: $b = 0.86$; p-value < 0.12; IG2 vs. CG: $b = 0.96$; p-value < 0.08</p> <p>Chewing tobacco at all time points: unclear and unreported</p>	
Notes	Further clarification on outcomes was unavailable	
<i>Risk of bias</i>		
<i>Item</i>	<i>Author judgement</i>	<i>Description</i>
Sequence generation: was the allocation sequence adequately generated?	Not applicable	Non-randomised trial
Allocation concealment: was the allocation adequately concealed?	Not applicable	Non-randomised trial
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	Non-randomised trial
Incomplete outcome data: were complete data for each outcome reported and, if not, were adequate reasons for incomplete outcome data provided?	No	Only people with all measurement occasions were included
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	No	All outcomes reported, though incompletely
Key confounders: were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Unclear	Covariate results inadequately presented
Clustering: was clustering of participants accounted for in the analysis?	No	No evidence of accounting for clustering
Other source of bias: were there any other sources of bias that might affect the results of the study?	Yes	Outcomes unclear and poorly presented

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations
Tebes et al.⁷⁹	
Methods	Non-randomised trial
Unit of allocation	Group: comparator chosen by intervention programmes in neighbouring cities
Participants' details	Country: USA Sample number: 149 (IG), 155 (CG) Sex: 47% female, 53% male Ethnicity: 75.7% African American, 19.7% Hispanic, 3.9% Caucasian SES: about 13% of parents less than high school educated
Intervention details	Description: PYDC Targeted/universal: universal Target population: 12- to 16-year-olds Theory: not stated Provider(s): community group leaders Training: 12 hours in curriculum and facilitation, bi-weekly supervision Setting: public schools Content: two core components: (1) The Adolescent Decision-Making for the PYDC (ADM-PYDC) a substance-use prevention curriculum which included: (a) programme introduction and overview; (b) understanding and coping with stress and learning stress-reduction strategies; (c) learning the steps of effective decision-making; (d) learning essential information about tobacco, alcohol and other drug use (two sessions); (e) applying the decision-making process to one's life through identifying positive personal attributes, dealing with job and school stressors, setting positive goals for healthy living and enhancing one's social networks and resources (four sessions); and (f) programme close and review; and (2) participation in health education and cultural heritage activities e.g. regular field trips to community agencies, civic organisations, businesses and schools to promote learning about community service and understanding one's cultural heritage. The field trips promoted after-school experiences and access to academic and vocational support, counselling services and participating in intergenerational programmes and community theatre Length/intensity: school year Control: ASPs without PYD/adolescent decision-making components

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
Outcomes	<p>At post intervention (less than unity favours intervention):</p> <p>Use of alcohol in past 30 days: OR 1.179, 95% CI 0.49 to 2.87</p> <p>Use of marijuana in past 30 days: OR 1.759, 95% CI 0.66 to 4.68</p> <p>Use of other drugs (amphetamines, cocaine, heroin, non-prescription methadone, hallucinogens, tranquillisers, inhalants) in past 30 days: OR 1.266, 95% CI 0.52 to 3.10</p> <p>Use of any drug in past 30 days: OR 1.694, 95% CI 0.74 to 3.90</p> <p>At 3 months post treatment:</p> <p>Use of alcohol in past 30 days: OR 0.365, 95% CI 0.15 to 0.90</p> <p>Use of marijuana in past 30 days: OR 0.178, 95% CI 0.08 to 0.42</p> <p>Use of other drugs in past 30 days: OR 0.188, 95% CI 0.08 to 0.44</p> <p>Use of any drug in past 30 days: OR 0.289, 95% CI 0.013 to 0.67</p> <p>Notes</p> <p>ORs as expressed in study are change indices rather than controlled for baseline values</p>	
<i>Risk of bias</i>		
Item	Author judgement	Description
Sequence generation: was the allocation sequence adequately generated?	Not applicable	Non-randomised trial
Allocation concealment: was the allocation adequately concealed?	Not applicable	Non-randomised trial
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	Non-randomised trial; research team responsible for recruiting comparison group
Incomplete outcome data: were complete data for each outcome reported and, if not, were adequate reasons for incomplete outcome data provided?	No	Dropout characteristics analysed, but missing data unimputed though missingness roughly balanced across arms
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Yes	Some outcomes (e.g. tobacco) not reported
Key confounders: Were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Yes	Propensity score matching entered into level 2 of multilevel model
Clustering: was clustering of participants accounted for in the analysis?	No	No evidence of accounting for clustering
Other source of bias: were there any other sources of bias that might affect the results of the study?	No	

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations
Wiggins et al.¹¹² (linked studies: Wiggins et al.⁴⁶)	
Methods	Non-randomised trial: group matched
Unit of allocation	Group
Participants' details	Country: UK Sample number: 1637 (IG), 1087 (CG) Sex: IG: 62% male, 38% female; CG: 56% male, 44% female Ethnicity: IG: 23% BME, CG: 20% BME SES: 73% IG, 61% CG in non-private housing; 39% and 35% in workless households
Intervention details	Description: YPDP Targeted/universal: targeted Target population: universal targeting of young people Provider(s): youth service providers Theory: youth development model Training: training of volunteers on programme requirements, recognising child abuse, working with youth; monthly supervision Setting: community youth services Content: activities focusing on young people's health and education as well as their broader social development with specific programme content determined by the individual projects delivering services. These could include: education (literacy, numeracy, IT, vocational skills). training/employment opportunities life skills (e.g. communication, decision-making, goal-setting, relationships, negotiation, anger-management), mentoring (weekly one-to-one sessions with staff), volunteering (both career-oriented and community-based), health education (particularly sexual health, substance misuse) arts and sports, advice on accessing services (health, contraceptive, drug and alcohol services, welfare, benefits advice, counselling and advice, housing) Length/intensity: 6–10 hours' weekly provision for 1 year Control: matched comparison
Outcomes	Post intervention (9-month follow-up, less than unity favours intervention): Cannabis use weekly or more in previous 6 months: OR 1.41, 95% CI 0.75 to 1.71 Cannabis use monthly or more in previous 6 months: OR 1.56, 95% CI 0.93 to 2.63 Drunkness monthly or more in previous 6 months: OR 0.98, 95% CI 0.65 to 1.47 9 months post intervention (18-month follow-up, less than unity favours intervention): Cannabis use weekly or more often in previous 6 months: OR 1.97 95% CI 0.93 to 4.17 Drunkness monthly or more often in previous 6 months: OR 1.20 95% CI 0.78 to 1.84

continued

TABLE 19 Characteristics of outcome evaluations (studies included in the review to answer RQ3) (*continued*)

Study details	Characteristics of outcome evaluations	
<i>Risk of bias</i>		
<i>Item</i>	<i>Author judgement</i>	<i>Description</i>
Sequence generation: was the allocation sequence adequately generated?	Not applicable	Non-randomised trial
Allocation concealment: was the allocation adequately concealed?	Not applicable	Non-randomised trial
Blinding: was knowledge of the allocation intervention adequately prevented during the study?	No	Non-randomised trial
Incomplete outcome data: were complete data for each outcome reported and, if not, were adequate reasons for incomplete outcome data provided?	Yes	Weighted analysis partially accounts for missingness
Selective outcome reporting: were the findings of the study subject to selective outcome reporting?	Yes	Full outcomes presented at follow-up 1 not presented at follow-up 2
Key confounders: were differences in non-random studies in key confounders (e.g. SES, sex, age) adequately controlled for?	Yes	Covariates included in model
Clustering: was clustering of participants accounted for in the analysis?	Yes	Analysis accounted for clustering
Other source of bias: were there any other sources of bias that might affect the results of the study?	No	
CG, control group; IG, intervention group; LSD, lysergic acid diethylamide; PED, performance-enhancing drug; SD, standard deviation.		

TABLE 20 Quality assessment of outcome evaluations

Study	Sequence generation		Allocation concealment		Blinding		Incomplete outcome data		Selective outcome reporting		Confounder		Other sources of bias	
	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD	CB/SM	KD
Berg <i>et al.</i> ⁸⁶	0	0	0	0	3	3	1	1	2	2	2	2	3	3
Cross <i>et al.</i> ⁹⁰	1	1	1	1	1	1	1	1	1	1	0	0	2	2
Gottfredson <i>et al.</i> ⁹³	0	0	0	0	1	1	2	2	1	2	2	1	2	1
Kuperminc <i>et al.</i> ⁹⁵	0	0	0	0	3	3	1	2	1	2	2	2	3	3
Rodriguez-Planas ⁷⁸	2	1	1	1	1	1	1	1	2	2	0	0	3	3
Rhodes <i>et al.</i> ⁹⁶	1	1	1	1	1	1	2	2	1	3	2	2	3	2
Schwartz <i>et al.</i> ⁷⁷	1	1	1	1	1	1	1	1	1	1	0	0	3	1
St Pierre <i>et al.</i> ¹¹¹	0	0	0	0	3	3	2	2	1	3	3	3	2	2
Tebes <i>et al.</i> ⁷⁹	0	0	0	0	0	3	1	1	3	1	2	2	3	3
Wiggins <i>et al.</i> ¹¹²	0	0	0	0	0	0	1	1	2	2	2	2	3	3

0, not applicable; 1, not stated/unclear; 2, yes; 3, no.
All studies were coded between CB and KD except for Wiggins *et al.*¹¹²

Appendix 17 Policy and young people's report briefs

The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes

EVIDENCE BRIEF

About this brief

This brief summarises evidence from a National Institute of Health research funded project on: The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcome.

Background and rationale

Substance use and violence are highly prevalent and damaging to young people's health. There are calls for interventions to address multiple rather than single risk behaviours because these behaviours cluster together^{31,32} and can potentially be more efficient. Positive youth development (PYD) is one such intervention to address inter-clustered risk behaviours among young people. The UK's National Youth Agency (NYA) defines such interventions as voluntary educational activities aiming to bring about generalised youth development in terms of positive assets such as skills and confidence, rather than merely remedying 'problem behaviours'. Non-systematic review of PYD effects on violence and drug use^{30,50} suggest benefits as well as variability, but must be treated with caution given these are unsystematic and quite old.

Aim and review questions

This systematic review aimed to systematically search for, appraise the quality of and synthesise evidence on PYD interventions addressing substance use or violence, asking the following review questions:

1. What **theories of change** inform PYD interventions delivered to young people aged 11-18 addressing substance use and violence?
2. What **characteristics of participants and contexts are identified as barriers and facilitators** of implementation and receipt in process evaluations of PYD?
3. What is the **effectiveness** and cost-effectiveness **of PYD** when compared to usual or no treatment in **reducing substance use** (smoking, alcohol, drugs), and **violence** (perpetration and victimization)?
4. What **characteristics of participants and contexts** appear to **moderate/are necessary and sufficient for PYD effectiveness**?

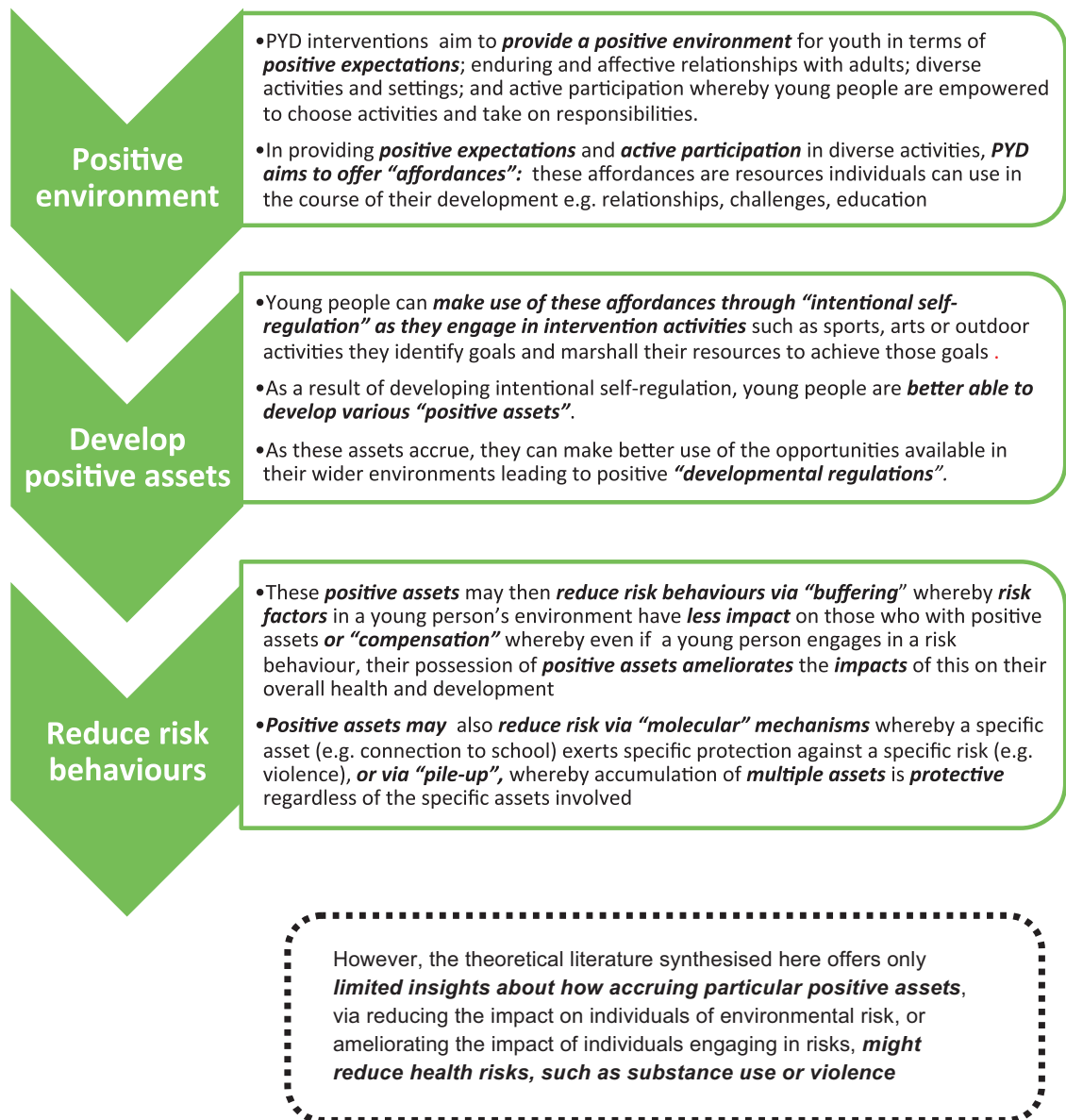
Key findings

Included studies

A total of 48 reports and **30** distinct **studies** (i.e. a distinct description of theory of change or empirical evaluation) were included in the review. **Sixteen** reports described **theories of change**; 12 reports, from **ten** distinct studies **evaluated processes**; and 26 reports, from **ten** distinct studies **evaluated outcomes**. Five reports combined theories of change, process evaluation and/or outcome evaluation. **No economic evaluations** or studies with cost data were identified.

RQ1. Theories of change for PYD effects on substance use and violence

Sixteen reports were included. We aimed to assess the quality of these theories drawing on criteria used previously but found that these were challenging to apply consistently to the PYD theoretical literature. There was insufficient information to develop a comprehensive theory of change for the effects of PYD interventions on substance use and violence. However, by filling in some of the gaps we succeeded in generating the following theory of change:



RQ2. Characteristics of participants and contexts that affect implementation and receipt of PYD

Eight of the 10 process evaluations were conducted in the USA, the remaining two were conducted in Australia and England. Study quality ranged from high reliability and usefulness (n=3) to low reliability and usefulness (n=3).

A number of themes emerged from process evaluation synthesis:

- **Community engagement** was a **key** to ensuring programmes were **culturally sensitive, accessible and appealing** to young people and their parents, and the wider community.
- **Employing community members** could be **pivotal to successful implementation** and providing role models. However, volunteers could be unreliable for example in acting as mentors.
- **Collaboration with other community agencies** could be **important** particularly in expanding the range of activities offered but could lead to drift from original approaches where other agencies e.g. schools had other goals.
- Evidence on **young people's relationships with providers and peers** suggest that: i) providers should relate to young people in a **calm, nurturing yet authoritative way** and ii) skilled providers could **bridge social differences between participants such as those involved in gangs**, but this could be undermined by poor training or retention.
- **Staff retention was challenging** where programmes could not offer full time positions.
- Providers found it **challenging to empower young people to make decisions about engagement in programme activities** while also requiring them to **engage in diverse activities** including vocational or academic activities.

RQ3. Effectiveness and cost-effectiveness in reducing substance use and violence

We included 12 study reports of 9 distinct outcome evaluations in our meta-analyses.

- **Nine** studies were conducted in the **USA**; and **one** in the **UK**.
- **Three intervention types** were identified: after-school; multi-component; and mentoring-driven interventions.
- **Four** studies were **randomised controlled trials**; **five** were **non-randomised trials with prospectively matched control groups**; and one included both randomised trial and non-randomised components
- Overall **quality of evidence** for our analyses of substance use and violence outcomes was rated '**very low**'

Meta-analysis of outcomes found:

- **Small**, statistically significant **short-term effects** (0-4 months post-intervention) for an omnibus measure of **substance use and violence**.
- However, no significant effects for alcohol, illicit drugs or smoking

We **could not** undertake meta-regression to **assess socio-demographic moderators** but narrative synthesis suggested no clear pattern of effects by sex.

RQ4. Characteristics of participants and contexts that determine effectiveness

We aimed to examine what characteristics of participants and contexts appear to moderate/are necessary and sufficient for PYD effectiveness. Synthesis of PYD theories of change and process evaluations suggested several hypotheses:

- interventions which offer and breadth of activities may be more effective for younger adolescents while those which emphasise depth may be more effective for older adolescents;
- interventions which combine prevention and positive development may be less effective than those which only focus on positive development;
- interventions of more than one year's duration may be more effective than those of shorter duration;
- interventions may be more effective for participants with low or moderate levels of baseline risk since there is more scope for stimulating "intentional self-regulation"¹²⁴
- interventions that have specific methods to engage communities will be more effective;
- projects that engage with schools will achieve better recruitment;
- interventions that are delivered by well trained staff will be more effective;
- interventions that have better staff retention will be more effective; and
- interventions that offer some choices but require some engagement with educational components will be more effective.

However the limited number of studies and very low level of statistical heterogeneity precluded examination of these hypotheses via meta-regression or qualitative comparative analyses.

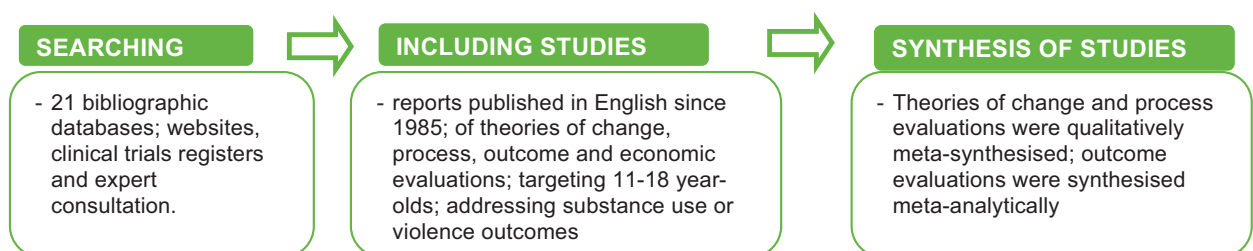
Conclusions

The health effects of PYD are currently under-theorised. Implementation can be challenging. We found no evidence that PYD interventions currently being evaluated achieve reductions in substance use or violence of public health significance. However, these may not constitute a test of the effectiveness of all aspect of the PYD model since the interventions evaluated, though meeting our inclusion criteria, may not be exemplars of PYD.

Implications

- ✓ Any investment in PYD as a strategy to reduce substance use and violence outcomes should occur only within the context of evaluation studies
- ✓ Future evaluations of the effects of PYD interventions on these outcomes must clarify intended mechanisms of action, and describe their theory of change and intervention characteristics in more detail.
- ✓ When delivering PYD programme, further attention must be given to resolving the tensions between enabling YP to choose which activities to participate in and in ensuring YP engage in sufficient breadth of intervention activities
- ✓ PYD providers need sufficient capacity in terms of staff and resources for successful implementation.

Methodology





How can we enable Positive Youth Development? The research says...

1. Provide a positive environment through.

- Positive expectations
- Real relationships with adults
- Diverse activities and settings
- Active participation

2. Positive Youth Development approach aims to offer

- Breadth, depth and duration in programmes
- Focus on positive development not just risk behaviour
- Opportunities for YP contribute to or challenge society

3. Young people change through

Reflecting on and learning from behaviour
Selecting activities that help meet goals
Changing behaviour if not successful

4. Young people gain positive assets:

- * Competence * Confidence *
- * Connection * Character *Caring *

The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes



How might PYD decrease smoking, drinking, drug use & violence? The research says...

Seeing positive behaviours; having links to adults, peers and institutions ↓peer influence on substance use

Learning how to resolve conflict from youth projects ↓violence

Feeling a part of society, connected to people and believing in yourself protects against behaviour

Young people learn to make choices based on internal values rather than due to peer influence

Does PYD decrease drug use, drinking, smoking & violence? The research says...

Positive youth development programmes:

- ❖ *Do have* a small **effect** on whether young people's **drug use**.
- ❖ *Do not* have an **effect** on whether young people drink **alcohol**.
- ❖ *Do not* influence whether young people **smoke**.
- ❖ *Do have* a beneficial **influence** on **violence** in

When considering the findings Young people

Agreed that PYD programmes

- ❖ *can* have an **effect** on their **drug use**
- ❖ *and not* on their **alcohol consumption** unless programmes provide an alternative social venue as drinking

Agreed and disagreed that PYD programmes

- ❖ *can* influence **smoking** and
- ❖ have a **beneficial influence** on **violence**



Young people views on 5 themes from the synthesis on delivering PYD

- **Theme 1: Collaboration with the community to ensure cultural sensitivity; appeal to YP and employing community members. Young people said:**
 - Mix of external & community members is best as external people bring other experiences agreed that cultural sensitivity is important;
- **Theme 2: Collaborate with other community agencies to expand range of activities and liaising with schools critical Young people said**
 - Focusing on one public health topic is better than having different sessions
 - Good to be on-site (school and youth club) but needs to be separate otherwise they will worry about confidentiality.
- **Theme 3: Focus on YP's relationship with PYD providers and peers**
 - Agreed that providers need to relate in a calm and nurturing but authoritative way and for skilled providers to bridge social differences with peers
 - Relationship with workers need to balance between friend and teacher and leader
- **Theme 4: Retain programme staff**
 - Staff retention need to increase trust and to build relationship, their loss affects the attendance of young people and their focus when they do attend
 - Agree that job instability causes workers to move on, and that this is not their fault.
 - Need to ensure that there is always senior staff and younger staff who are reliable.
- **Theme 5: Youth-led empowerment**
 - Need to give YP lots of options to choose between, but also impose certain limits within budget possibilities while taking into consideration what YP want to learn.
 - And allow for some recreational time to talk to friends
 - Some schools put on youth activities after school, which are different to youth clubs. However the YP normally choose to go to youth services.

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health 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

Published by the NIHR Journals Library