Using process data to understand outcomes in sexual health promotion: an example from a review of school-based programmes to prevent sexually transmitted infections

J. Shepherd1*, A. Harden2, E. Barnett-Page3, J. Kavanagh3, J. Picot1, G. K. Frampton1, K. Cooper1, D. Hartwell1 and A. Clegg1

1Southampton Health Technology Assessments Centre (SHTAC), University of Southampton, Southampton SO16 7NS, UK, 2Institute for Health and Human Development, University of East London, London E15 4LZ, UK and 3Social Science Research Unit, University of London, London WC1H 0AL, UK

*Correspondence to: J. Shepherd. E-mail: Jonny.Shepherd@soton.ac.uk

Received on June 10, 2013; accepted on December 2, 2013

Abstract

This article discusses how process indicators can complement outcomes as part of a comprehensive explanatory evaluation framework, using the example of skills-based behavioural interventions to prevent sexually transmitted infections and promote sexual health among young people in schools. A systematic review was conducted, yielding 12 eligible outcome evaluations, 9 of which included a process evaluation. There were few statistically significant effects in terms of changes in sexual behaviour outcomes, but statistically significant effects were more common for knowledge and self-efficacy. Synthesis of the findings of the process evaluations identified a range of factors that might explain outcomes, and these were organized into two overarching categories: the implementation of interventions, and student engagement and intervention acceptability. Factors which supported implementation and engagement and acceptability included good quality teacher training, involvement and motivation of key school stakeholders and relevance and appeal to young people. Factors which had a negative impact included teachers’ failure to comprehend the theoretical basis for behaviour change, school logistical problems and omission of topics that young people considered important.

It is recommended that process indicators such as these be assessed in future evaluations of school-based sexual health behavioural interventions, as part of a logic model.

Introduction

In the last three decades, there have been a number of published evaluations of sexual health promotion, featuring a range of types of intervention and reporting varied degrees of effectiveness [1–5]. However, with high prevalence of sexually transmitted infections (STIs), particularly in younger age groups, and continued high rates of unintended pregnancy in many countries, strong evidence of effective sexual health promotion interventions remains a priority. In addition to interventions addressing the wider determinants of sexual health, effective interventions at the level of the individual, such as those designed to encourage safer sexual behaviour, are still warranted.

Generally, effective sexual health promotion has been judged in terms of its ability to influence outcomes such as safer sexual behaviour and prevention of transmission of STIs, thereby potentially reducing morbidity and mortality. For many decision makers these outcomes are important for demonstrating impact and judicious use of resources.
For example, the Sexual Health Framework for England [6] states that safer sexual behaviour is a priority for health promotion, and proposes three indicators of success relating to the health improvement domain of the Public Health Outcomes Framework [7], including reducing under 18 conceptions and reducing chlamydia diagnoses in 15- to 24-year olds. However, there are challenges in demonstrating health-related behaviour change and maintenance [8, 9], and assessing longer term outcomes such as infection and its sequelae.

Health outcomes may be classified on a scale ranging from proximal and intermediate to distal [10]. Proximal outcomes in sexual health promotion may be varied, and include knowledge of infection transmission and methods of prevention, attitudes and behavioural intentions to practice safer sex, beliefs around susceptibility to STIs, self-efficacy and communication skills to negotiate safer sex. Such outcomes can be theoretically hypothesized as being causally linked to intermediate outcomes (e.g. sexual behaviour) [11–14] and, in turn, to distal outcomes (e.g. STIs and associated health complications, morbidity and mortality) and therefore in the interim give an indication of how successful an intervention is ultimately likely to be, and the reasons for success or failure. It is important that evaluators not only assess all relevant outcomes along the scale but also that they evaluate process [15]. Process evaluations explore how an intervention was implemented and how it was received, aiding the interpretation of outcomes and helping to unpick the ‘black box’ between inputs and outcomes [16, 17]. Process evaluation may occur with or without outcome evaluation and may include qualitative and/or quantitative data collection. In particular, the assessment of the level of implementation of an intervention in an evaluation has been linked to larger effect sizes, underlining the importance of this processes measure [18].

Process indicators associated with an intervention can be thought of as complementary to proximal and intermediate outcomes in helping to predict and explain distal outcomes to provide a comprehensive perspective on effectiveness. However, few published systematic reviews of sexual health have assessed outcomes within the context of processes, and there is a lack of guidance on which process indicators are of most relevance in sexual health.

Logic models are being increasingly used to conceptualize the mechanisms by which interventions are thought to operate. A logic model is a tool which programme planners can use to systematically plan the resources needed to deliver an intervention, the intervention activities and the intended outputs and outcomes [19, 20]. Logic models are also used in the monitoring and evaluation of interventions [21, 22], particularly complex interventions, and can also be a valuable framework for systematic reviews of the literature [23, 24]. A visual schematic is often used to illustrate the relationships between the components of a logic model (e.g. inputs, activities, outputs and outcomes).

Logic models could be particularly useful for planning and explaining the relationships between interventions, processes and outcomes in an evaluation. Guidance is needed to assist programme planners and evaluators in the design and planning of interventions, to ensure that as many of the possible mediators of effect are assessed and accounted for. The aim of this article, therefore, is to describe a systematic synthesis of process evaluations and to discuss how these could complement outcomes in the overall assessment of effectiveness, using as an example skills-based behavioural interventions to prevent STIs and promote sexual health among young people in schools.

### Methods

The research reported in this article was one component of a health technology assessment investigating the effectiveness and cost effectiveness of behavioural interventions to prevent STIs in young people [25, 26]. A systematic review of outcome evaluations, including outcome evaluations with integrated process evaluations, was conducted using standard methodology for evidence synthesis [27]. This article focuses on the process evaluations included.
To be included in the review studies had to: include young people aged 13–19 years; evaluate behavioural interventions based in (but not restricted to) schools, in which an element of the intervention included the development of sexual behavioural skills (e.g. how to use a condom, to negotiate safer sex with partners) and use a randomized controlled trial (RCT) design. The primary outcome measure of interest was self-reported sexual behaviour (e.g. condom use, number of partners, etc.), however, studies reporting other outcomes could be included providing behaviour was also measured.

Data were extracted from included studies independently by two reviewers using specialist review software [28]. The methodological quality of each study was assessed by two reviewers independently with differences in opinion resolved by discussion. The quality of the process evaluations was assessed according to a set of criteria developed specifically for this review, based on our own, and others’, previous work assessing the quality of process evaluations and qualitative research [29–32]. The criteria assess the rigour of sampling, data collection and analysis methods and the depth and breadth of the findings, and required reviewers to make a judgment in relation to the overall weight of evidence they would place on a study in terms of (i) the trustworthiness of its findings and (ii) the usefulness of the findings (see Table I).

Narrative methods were used to synthesize findings from the process evaluations, based on those we have developed in previous work on the synthesis of process evaluations and qualitative research [29, 33, 34] and from other groups working on methods for synthesizing process evaluations [35–37].

Two reviewers independently read and re-read the tabulated study details and noted down their initial thoughts on the main themes to arise from the findings. The two reviewers met to discuss and compare their individual themes and compiled a jointly agreed list. A narrative was written to describe and elaborate on these themes, and this was reviewed and discussed by the review team to consider how the insights from process evaluations might explain the findings from the outcome evaluations. The themes and associated narrative were interrogated and re-organized to address two questions, based upon the two most commonly studied processes within the process evaluations. First, what factors facilitate or hinder the implementation of skills-based behavioural interventions in schools? Second, what factors impact on student engagement and intervention acceptability? In this article, the findings of the synthesis of process evaluations are presented in response to these questions. A narrative is provided for each theme (factor) relating to implementation or engagement and acceptability, accompanied by a summary of the effects for each of the outcomes measured in the studies contributing data to the theme. The purpose was to illustrate the association between the process factors and study outcomes.

Results

A total of 8037 references were identified through literature searching. Following screening of titles and abstracts 355 references were retrieved for further screening. From these a total of 136 separate evaluations were included in a descriptive map of the evidence base, and of these, 15 RCTs met the inclusion criteria for the systematic review. Twelve of the 15 were judged to be methodologically sound following quality assessment. Studies reported on a range of behavioural outcomes, including initiation of sexual intercourse and condom use. Rates of STIs were not reported by any studies. There were few statistically significant differences between behavioural interventions and comparators in terms of changes in sexual behaviour outcomes. Statistically significant effects were reported for some, but not all, of the self-efficacy measures assessed, and statistically significant effects in favour of the intervention group over the comparison group were found by all but two of the studies for knowledge of STIs (further detail on outcomes can be found in separate publications [25, 26]). Nine of the 12 sound outcome evaluations included a process evaluation, and these are the focus of this article (Table II) [38–46].

The methodological quality of the process evaluations was mixed. All but four studies were assigned a low weight of evidence for the reliability and
Table I. Criteria used to assess the quality of the process evaluations

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Were steps taken to increase rigour/minimize bias and error in the sampling for the process evaluation? Consider whether: (a) the sampling strategy was appropriate to the questions posed (b) attempts were made to include all relevant stakeholders and/or obtain a diverse sample (c) characteristics of the sample critical to the understanding of the study context and findings were presented.</td>
<td>(a) Yes, a fairly thorough attempt was made. (b) Yes, several steps were taken. (c) Yes, a few steps were taken. (d) No, not at all/not stated/unclear.</td>
</tr>
<tr>
<td>(ii) Were steps taken to increase rigour/minimize bias and error in the data collected for the process evaluation? Consider whether: (a) data collection tools were piloted/(if quantitative) validated (b) data collection was comprehensive, flexible and/or sensitive enough to provide a complete and/or vivid and rich description/evaluation of the processes involved in the intervention.</td>
<td>(a) Yes, a fairly thorough attempt was made. (b) Yes, several steps were taken. (c) Yes, a few steps were taken. (d) No, not at all/not stated/unclear.</td>
</tr>
<tr>
<td>(iii) Were steps taken to increase rigour/minimize bias and error in the analysis of the process data? Consider whether: (a) data analysis methods were systematic (b) diversity in perspective was explored (c) the analysis was balanced in the extent to which it was guided by preconceptions or by the data (d) the analysis sought to rule out alternative explanations for findings.</td>
<td>(a) Yes, a fairly thorough attempt was made. (b) Yes, several steps were taken. (c) Yes, a few steps were taken. (d) No, not at all/not stated/unclear.</td>
</tr>
<tr>
<td>(iv) Were the findings of the process evaluation grounded in/supported by the data? Consider whether: (a) enough data are presented to show how the author’s arrived at their findings (b) the data presented fit the interpretation/support claims about patterns in data (c) the data presented illuminate/illustrate the findings (d) (for qualitative studies) quotes are numbered or otherwise identified so that the reader can see that they do not just come from one or two people.</td>
<td>(a) Very well grounded/supported. (b) Fairly well grounded/supported. (c) Limited grounding/support.</td>
</tr>
<tr>
<td>(v) Please rate the findings of the process evaluation in terms of their breadth (extent of description) and depth (extent of data transformation/analysis) Consider whether: (a) a range of processes/issues were covered in the evaluation (b) the perspectives of participants are fully explored (c) both the strengths and weaknesses of the intervention are described/explored (d) the context of the intervention has been fully described/explored (e) richness and complexity has been portrayed (f) there has been theoretical/conceptual development.</td>
<td>(a) Limited breadth or depth. (b) Good/fair breadth but very little depth. (c) Good /fair depth but very little breadth. (d) Good/fair breadth and depth.</td>
</tr>
</tbody>
</table>

(continued)
trustworthiness of the findings, and all but three were assigned a low weight of evidence for the usefulness of the findings. Three studies stand out overall as they were judged to be medium or high weight of evidence for both trustworthiness and usefulness (Stephenson et al. [44], Wight et al. [45] and Zimmerman et al. [46]).

All but two studies [42, 43] collected data from intervention providers, and all but two [38, 42] collected data from intervention recipients themselves. A range of processes were evaluated (Table III).

### Factors facilitating or hindering the implementation of skills-based behavioural interventions in schools

The fidelity of intervention implementation (i.e. the extent to which the intervention was delivered as intended) was assessed by all studies, but only seven reported findings on this [38, 41–46]. As a number of these were multi-site cluster RCTs they involved the implementation of a standardized programme across a number of schools. From the synthesis, a hierarchy emerged (see Fig. 1), incorporating a number of factors that may impact on fidelity of implementation.

The respective elements of this hierarchy are described below in terms of the factors that appeared to work for or against the effectiveness of the intervention (Table IV).

#### School culture and administration

School culture was identified as an overarching factor impacting on fidelity of implementation. School culture was determined by a range of key elements, and most clearly emerged as being significant from the Scottish SHARE trial (Sexual Health and Relationships: Safe Happy and Responsible) [45]. One element identified from this study was the involvement of key stakeholders at a number of levels, from the Health Education
Table II. Characteristics of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes measured</th>
<th>Weight of evidence for process evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study Intervention Comparator Outcomes measured</td>
<td>Weight of evidence for process evaluation</td>
<td>Trustworthiness of findings</td>
<td>Usefulness of findings</td>
</tr>
<tr>
<td>Borgia et al. [38] Country: Italy. Ethnicity: NR Sex: Mixed, SES: Representative of general population</td>
<td>Not named. Peer-led: five sessions, 10 hours in total. Teacher-led: same total length suggested over ~3 months.</td>
<td>Behavioural: • Condom use Self-efficacy Knowledge Attitudes</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>1295</td>
<td>613</td>
<td>682</td>
<td>18.2 (1.1)</td>
</tr>
<tr>
<td>Agea</td>
<td>18.3 (1.1)</td>
<td>18.3 (1.1)</td>
<td>18.2 (1.1)</td>
<td>18.2 (1.1)</td>
</tr>
<tr>
<td>N</td>
<td>496</td>
<td>269</td>
<td>227</td>
<td>NR</td>
</tr>
<tr>
<td>Age</td>
<td>13.2 (0.94)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Jemmott et al. [40] Country: United States. Ethnicity: Black Sex: Male only SES: Low</td>
<td>Not named. Single 5-hour session led by specially trained adult facilitators. Career opportunities: single 5-hour session led by specially trained adult facilitators.</td>
<td>Behavioural: • Condom use • Intercourse • Partners Knowledge Attitudes Behavioural intentions</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>157</td>
<td>85</td>
<td>72</td>
<td>NR</td>
</tr>
<tr>
<td>Age</td>
<td>14.64 (1.66)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Karnell et al. [41] Country: South Africa. Ethnicity: predominantly Zulu Sex: Mixed SES: Low</td>
<td>‘Our Times, Our Choices’: 10 units of 30 min each, delivered over ~8 weeks by teachers and peer leaders (both trained). Life orientation instruction, presumably teacher-led: Length not stated.</td>
<td>Behavioural: • Condom use • Intercourse Self-efficacy Knowledge Attitudes</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>661</td>
<td>325</td>
<td>336</td>
<td>16 (median)</td>
</tr>
<tr>
<td>Age</td>
<td>16 (median)</td>
<td>16 (median)</td>
<td>16 (median)</td>
<td>16 (median)</td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes measured</th>
<th>Weight of evidence for process evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trustworthiness of findings</td>
</tr>
<tr>
<td>Levy et al. [42] Country: United States. Ethnicity: mixed Sex: Mixed SES: Low</td>
<td>‘Youth AIDS Prevention Project (YAPP)’ delivered by trained health educators: 10 sessions over 2 weeks. Booster: five sessions over 1 week the following year.</td>
<td>Basic AIDS education (current practice) presumably delivered by teachers. Length not stated.</td>
<td>Behavioural: <em>Condom use</em> <em>Intercourse</em> <em>Partners</em> Knowledge Behavioural intentions</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>2392</td>
<td>1459 (1001 for baseline characteristics)</td>
<td>933 (668 for baseline characteristics)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>12.5</td>
<td>12.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roberto et al. [43] Country: United States. Ethnicity: predominantly European American Sex: Mixed. SES: Low</td>
<td>Not named: six computer activities for students, each ~15 min, over 7 weeks.</td>
<td>Not described (no intervention).</td>
<td>Behavioural: <em>Condom use</em> <em>Sexual initiation</em> <em>Partners</em> Self-efficacy Knowledge Attitudes</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>378</td>
<td>164 (139 for baseline characteristics)</td>
<td>214 (187 for baseline characteristics)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>15.50 (0.63)</td>
<td>15.68 (0.73)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stephenson et al. [44] Country: United Kingdom. Ethnicity: NR Sex: Mixed SES: Representative of general population</td>
<td>‘RIPPLE’: three peer-led sessions, each of 1 hour over the summer term.</td>
<td>Teacher-led SRE (current practice). Length not stated.</td>
<td>Behavioural: <em>Condom use</em> <em>Sexual initiation</em> <em>Intercourse</em> <em>Partners</em> <em>Contraception</em> <em>Pregnancy</em> Self-efficacy Knowledge Attitudes</td>
<td>High</td>
</tr>
<tr>
<td>N</td>
<td>8766</td>
<td>4516</td>
<td>4250</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>13–14</td>
<td>13–14</td>
<td>13–14</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Outcomes measured</th>
<th>Weight of evidence for process evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Trustworthiness of findings</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>7616</td>
<td>2867 (follow-up)</td>
<td>2987 (follow-up)</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>13–14</td>
<td>13–14</td>
<td>13–14</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>2647 (1944 for baseline characteristics)</td>
<td>RtR: 681</td>
<td>Modified RtR: 1149</td>
</tr>
<tr>
<td></td>
<td>Age</td>
<td>13–0.6%; 14–53.0%; 15–35.1%; ≥16–113%</td>
<td>NR</td>
<td>NR</td>
</tr>
</tbody>
</table>

SES, socio-economic status; NR, not reported.
*Age in years reported as mean (SD) unless otherwise stated.
Board of Scotland, to Guidance Teams, sex education co-ordinators and senior management. Higher levels of involvement tended to facilitate implementation. Motivation of these key players was also important, as was communication between teachers and importantly, communication style: in one school in the SHARE trial, the senior management imposed the programme (in all other schools there had been consultation with staff) and teachers felt unhappy with this ‘imposition’ which impacted negatively on implementation.

Factors relating to school organization also impacted on implementation. Staff absence and turnover, time-tableing issues and time shortages generally, led to lessons being missed or cut short [44, 45]. In the SHARE trial schools found it hard to set aside 20 lessons for sex education and, even when they felt sex education was a priority, they were all too aware that other topics competed with the programme [45]. Although the lessons had been designed to last for 40 min each, this was found to be

### Table III. Processes evaluated by included studies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility/programme reach</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Collaboration and partnerships</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Content</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Implementation</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Acceptability</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Quality of materials</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Skills and training of providers</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Other</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Total: 3 2 2 9 8 1 4 5
### Table IV. Themes identified from the analysis of the process evaluations, and the study outcomes (effects) for each theme

<table>
<thead>
<tr>
<th>Overarching theme</th>
<th>Study outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Implementation</strong></td>
<td></td>
</tr>
</tbody>
</table>
| School culture and administration | Intermediate: Sexual behaviour: no significant effects [45]  
                                Proximal: Knowledge: significant effects [45] |
| Teachers: enthusiasm, expertise and autonomy | Intermediate: Sexual behaviour: no significant effects [38, 45]  
                                Proximal: Self-efficacy: no significant effects [38]  
                                Knowledge: significant effects [38, 45]  
                                Attitudes: no significant effects [38] |
| Peers | Intermediate: Sexual behaviour: no significant effects [38, 46]; no significant effects on most outcomes [44]  
                                Proximal: Self-efficacy: no significant effects [38, 46]; no significant effects on most outcomes [44]  
                                Attitudes: no significant effects [38, 44, 46]  
                                Behavioural intentions: no significant effects [46]  
                                Knowledge: significant effects [38, 44, 46] |
| Student engagement |                |
| Appeal of content | Intermediate: Sexual behaviour: no significant effects [46]; no significant effects (expect for sub-group) [41]  
                                Proximal: Self-efficacy: no significant effects [46]; no significant effects (except for sub-group) [41]  
                                Knowledge: significant effects [46]  
                                Attitudes: no significant effects [41, 46]  
                                Behavioural intentions: no significant effects [46]; no significant effects (expect for sub-group) [41] |
| Qualities of providers |                |
| Meeting needs | Intermediate: Sexual behaviour: no significant effects on most outcomes [44]  
                                Proximal: Self-efficacy: no significant effects on most outcomes [44]  
                                Knowledge: significant effects [44]  
                                Attitudes: no significant effects [44] |
| Gendered norms | Intermediate: Sexual behaviour: no significant effects [45]; no significant effects on most outcomes [44]  
                                Proximal: Self-efficacy: no significant effects on most outcomes [44]  
                                Knowledge: significant effects [44, 45]  
                                Attitudes: no significant effects [44] |
| Age and timing | Intermediate: Sexual behaviour: no significant effects [38, 45]; no significant effects on most outcomes [44]  
                                Proximal: Self-efficacy: no significant effects [38]; no significant effects on most outcomes [44]  
                                Knowledge: significant effects [38, 44, 45]  
                                Attitudes: no significant effects [38, 44] |
optimistic and around a third of the sessions could not be completed.

**Teachers: enthusiasm, expertise and autonomy**

Most of the findings on teachers as intervention providers come from Wight *et al.* [45] who found that the process of actively cultivating teacher expertise and enthusiasm was vital in the delivery of SHARE. The training teachers received was particularly important in this regard and was viewed very positively by teachers. The study authors note that the perceived success of the training was largely attributed to the trainer, who was an experienced sexual health educator, who was felt to be clear and in control and able to put teachers at their ease.

Despite high levels of acceptability of the teacher training, a number of problems with the teachers' implementation of the intervention persisted. Some failed to engage with the theoretical basis for the intervention, in particular the mechanism for behavioural change, via the modelling and practising of skills. Wight and Buston [47] report that, in interviews, teachers seldom referred to skills development unprompted, and when interviewers raised the issue, they only talked about this briefly. In the interviews, the teachers never mentioned the theoretical aspects of the training which the authors had hypothesized were fundamental to achieving behavioural change and which distinguished the course from traditional sex education. The authors concluded that the teachers regarded the skills required to apply the theoretical principles of the intervention to be too far removed from their established repertoires. In practice, this meant that many teachers failed to persevere with skills development and role play activities.

This failure to engage with the theory behind an intervention was experienced in a different way by Borgia *et al.* [38] who suspected teachers of selecting the peer leaders who delivered their intervention for their academic skills rather than for the qualities that are hypothesized to make a good peer leader.

There was also some conflict between issues of fidelity and issues of professional autonomy: some teachers made amendments to the programme without consulting the intervention developers, as they felt they had the skills and experience to do so. Wight *et al.* [48] assessed the extent to which teachers delivered the intervention as intended and found that for 71% of sessions, teachers reported having followed the pack ‘very closely’, for 23% teachers reported modifying the session ‘slightly’ and for 6% they reported making ‘considerable’ modifications. ‘Considerable’ modifications included missing out sessions or key exercises, amalgamating sessions, abandoning exercises when pupil resistance was experienced and modifying teaching methods. Ten teachers reported making modifications which they viewed as ‘slight’ but which, in the opinions of the pack authors, would compromise the intervention in important ways, e.g. missing out or not completing key exercises.

**Peers**

Four of the studies used peer educators. Karnell *et al.* [41] and Zimmerman *et al.* [46] used peer educators alongside teachers; and Borgia *et al.* [38] and Stephenson *et al.* [44] employed only peers to deliver their intervention. The process evaluation data showed that selection of peer leaders was clearly important. As seen earlier, Borgia *et al.* [38] hypothesized that there were problems with the criteria by which the teachers in their intervention selected peer leaders (i.e. they were selected for academic skills rather than for the qualities that make an effective peer leader). This, they believed, compromised the ‘trustfulness and communication’ (p. 514) between educators and pupils. Stephenson *et al.* [44] had other problems with selection: one school was unable to recruit enough peer leaders to implement the intervention.

Like Borgia *et al.* [38], Stephenson *et al.* [44] questioned the aptitude of some of the recruited peer educators in the ‘Randomized Intervention of PuPil-Led sex Education’ (RIPPLE) trial. They report that in two schools some classes failed to receive peer-led sex education due to the disorganization and lack of enthusiasm of the peer educators for the programme. Zimmerman *et al.* [46] noted that
although the peer educators performed the tasks assigned to them, they did not achieve the level of involvement hoped for, ‘making this component of the modified curriculum less than ideal’ (p. 49). Some peer leaders in the RIPPLE trial were hampered by structural factors, including long gaps between training and delivery of sex education and timetable clashes for peer educators taking examinations.

Student engagement and intervention acceptability

The interventions evaluated by the studies included in this review shared two common features. First, all interventions were designed to engage young people actively in their own learning through interactive exercises such as role plays, discussions and small group work. Second, all interventions were designed to be relevant and appealing to young people in general or particular groups of young people. Although there was evidence from the process evaluations to confirm that interventions did engage and appeal to many of the young people involved, this was not always the case.

Six of the process evaluations contributed findings which illuminated issues of student engagement and intervention acceptability [38, 41, 43–46]. A number of factors emerged as influences on this issue.

Appeal of intervention content

Attempts to design interventions that were appealing to young people were met with some success in the three studies with relevant findings [41, 43, 46]. The interventions were judged to be relevant, clear and enjoyable. For example, three quarters of the participants rated four animated characters who modelled skill development in the South African-based intervention delivered to (predominantly) Zulu youth evaluated by Karnell et al. [41] as seeming ‘very’ or ‘extremely’ real to them, and 74% found the curriculum, delivered by peers and teachers, ‘very’ or ‘extremely’ interesting. However, some aspects of interventions were evaluated negatively, such as animated PowerPoint presentations used to introduce the curriculum in the US based ‘Reducing the Risk’ intervention evaluated by Zimmerman et al. [46] an intervention delivered by teachers and peers to ethnically diverse ‘high sensation seeking youth’. The authors concluded that it may be difficult to ‘yield a high level of sensation value in a classroom intervention’ (p. 49).

Qualities of intervention providers

The qualities and expertise of intervention providers was another factor that impacted upon student engagement and acceptability. Focus group data collected by Stephenson et al. [44] in the RIPPLE trial revealed that students receiving peer-led sex education were considerably more positive about their experience than those in control schools who received the teacher-led sex education that was usually delivered in their schools. Peer educators were perceived to be more relevant to pupils, using similar language, sharing similar values about sex and being less moralistic. However, data from questionnaires show that around a third of students receiving peer-led sex education found the peer-led component less acceptable, for example, when peer educators were not able to engage boys or manage their behaviour.

Meeting needs

Stephenson et al. [44] also argue that dissatisfaction or lack of engagement may be related to the possibility that the intervention, despite being designed to appeal to young people, did not in fact meet their own self-identifed needs. Their process evaluation examined this issue directly and found that overall participants felt that topics such as sexual feelings, emotions and relationships were not covered well by either teachers or peers. Researchers’ observations in the RIPPLE trial revealed that some important topics may not have been addressed in many of the peer-led sessions (e.g. emergency contraception). The issues discussed below around the format and timing of interventions also highlight how the failure of school-based interventions may be explained by a failure to respond to/acknowledge young people’s self-identified needs.
Gendered norms and mixed versus single sex groups

The interventions varied according to whether activities were delivered in mixed or single sex groups. The RIPPLE trial [44] was delivered in mixed sex groups but some of the participants said they would have preferred single sex sessions. However, observation data from the evaluation of the SHARE trial [49] revealed problems with the single sex classes. Boys tended to conform to ‘macho’ stereotypes or practiced self-censoring. It was observed that boys tended to work better in mixed sex groups and that this could build their confidence for talking about sex to young women.

Age and timing

Three studies raised the issue of the age appropriateness of the interventions. Wight and Abraham [49] concluded that their content was too advanced for the pupils in their intervention, since pupils’ lack of sexual experience at age 13–14 years meant that they failed to identify with the vignettes presented to them, which were designed to make pupils more aware of gendered interaction and power dynamics in sexual relationships, and/or found them alien. Borgia et al. [38] reported that the ‘work-groups’ that evaluated their programme judged it to be more suitable for younger populations (the participants in this study were 17–18 years of age). However, if content can be successfully matched to the age of the participants, a related issue is appropriate timing: at what age should school-based sex education start? Stephenson et al. [44] reported that more than half of students in both arms of their trial—who were the same age as the pupils in the SHARE trial [45]—would have liked their sex education earlier, although it is not reported exactly when.

One final reason why the interactive, skills-building exercises in the SHARE intervention failed was the evident discomfort felt by pupils in engaging with issues relating to sex in a classroom setting. This discomfort expressed itself either in disruptive hilarity or embarrassment [49]. The very element of the intervention that was considered by the developers to be its ‘active ingredient’—the interactive nature of many of the sessions—combined with the sensitive subject matter, in fact worked against its success in a classroom context.

Discussion

Our systematic review identified some statistically significant effects for the proximal outcomes considered to be mediators of health-related behaviour (e.g. knowledge, self-efficacy). However, few statistically significant effects in terms of changes in sexual behaviour outcomes—a key outcome of behavioural intervention—were found [25, 26]). There are a number of potential explanations for this finding, such as a lack of an external supportive environment (e.g. community norms), insensitive and/or inappropriate outcome measures and study methodological limitations (e.g. relatively short follow-up times inadequate for capturing changes in behaviour).

Our systematic assessment of process evaluations identified a range of factors that influenced the degree to which the interventions were implemented as intended, and how they were received and accepted by the young people they were aimed at. In turn these illuminated specific factors which appear to be working for and against the effectiveness of the intervention (Table III).

Positive factors included: good quality teacher training raising teacher’s competence and enthusiasm, involvement and motivation of key school stakeholders at a number of levels and interventions that were considered enjoyable and relevant to young people. In contrast, interventions suffered from factors including: high staff turnover, and logistical problems with school timetabling; teachers’ failure to comprehend the theoretical basis for the behaviour change intervention; unsanctioned changes to the intervention by teachers; omission of topics considered important by young people such as relationships and emotions; misjudgements about the age appropriateness of intervention messages and contradictory findings about the appropriateness of delivery to mixed versus single sex
groups. Some positive factors also had disadvantages. For example, peer educators were popular and considered to be more familiar to young people, sharing similar values. On the other hand, however, they were not always able to manage pupils' behaviour which disrupted their ability to provide effective sex education. Our findings resonate with other studies of school-based health promotion, such as of Payne et al. [50], who found significant relationships between implementation intensity in schools and factors such as integration into school operations and support from principals.

Arguably, without the assessment of process evaluations reported here we would be less informed about what factors appear to influence the effectiveness of behavioural interventions with young people. The process indicators identified can be situated within a logic model showing their relationship to the inputs and intervention activities and the proximal, intermediate and distal outcomes (Fig. 2). We recommend that process indicators such as these be assessed in future evaluations of school-based sexual health behavioural interventions as appropriate, and that evaluators devise a logic model such as this or a similar framework to guide the planning, implementation and evaluation of the intervention. The process indicators themselves could be subjected to further evaluation (e.g. comparing interventions delivered to single versus mixed sex groups, or comparing different interventions in terms of age appropriateness). Further to this, and in-keeping with the need to increase public participation in research [51, 52], it would be beneficial to involve young people themselves in the planning and design of interventions and the selection of outcome measures and process indicators that are relevant and meaningful to them. Such involvement may result in more effective...
interventions, and in evaluations that are better able to demonstrate effectiveness.

Unfortunately for the study in our systematic review that was most effective in terms of encouraging safer sexual behaviour, ‘Safer Choices’ [53], we were unable to identify a published process evaluation (hence why it is not included in this article), limiting insights into the factors contributing to its success. This underlines the need for outcome evaluations to include an assessment of process. Furthermore, a limitation of our own study is that there may be other important relevant process indicators that can contribute to our understanding of effectiveness but which were not evident from this particular set of studies. We therefore recommend that future evaluations assess a comprehensive range of processes thought to be relevant to the mechanism of the intervention and its outcomes.

A further limitation of our research is that it is based on a relatively small set of studies, some of which have methodological limitations, and is specific to one type of sexual health intervention in a defined group of people. However, this research has used transparent and systematic processes to identify, appraise and analyse the available literature, and the indicators identified here may be applicable more broadly in health promotion. For example, the finding that involvement of key stakeholders at a number of levels, external to, and within a school, facilitated intervention implementation is likely to be relevant to other settings (e.g. workplaces, health services, communities). The findings also resonate with published conceptual models of intervention implementation in health, such as the Consolidated Framework for Implementation Research (CFIR) [54]. The CFIR includes a number of theory-driven constructs to understand implementation, including those relating to individual knowledge and beliefs towards changing behaviour and the level of self-efficacy to make the change. In this study, key issues relating to the individual included teachers’ self-efficacy to deliver certain aspects of behavioural interventions, as well as their beliefs about the intervention and their desire for professional autonomy. Our findings therefore can be seen as being applicable more broadly within the field of implementation science.

Acknowledgements

We would like to thank the advisory group for their input throughout the course of the project. We also acknowledge the contribution of Jeremy Jones (Principal Research Fellow) to the economic evaluation and Alison Price (Information Specialist) who undertook some of the literature searches.

Funding

This work was supported by the National Institute for Health Research (NIHR) Health Technology Assessment programme (project number 06/72/02) and is published in full in Health Technology Assessment, Vol. 14, No. 7. See the HTA programme website for further project information. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health.

Conflict of interest statement

None declared.

References

Outcomes and process indicators for sexual health promotion


31. Harden A. Does study quality matter in systematic reviews that include qualitative research? XV Cochrane Colloquium, São Paulo, Brazil, 23–27 October 2007.


