



Integrating Process with Outcome Data in a Randomized Controlled Trial of Sex Education

VICKI STRANGE

University of London Institute of Education, UK

ELIZABETH ALLEN

Royal Free and University College London Medical School, UK

ANN OAKLEY

University of London Institute of Education, UK

CHRIS BONELL

London School of Hygiene and Tropical Medicine, UK

ANNE JOHNSON

Royal Free and University College London Medical School, UK

JUDITH STEPHENSON

Royal Free and University College London Medical School, UK

THE RIPPLE STUDY TEAM

Integration of process with outcome evaluations has been recommended as a way of addressing some of the criticisms of Randomized Controlled Trials (RCTs) for evaluating social interventions. This is the second of two articles (the first was published in issue 10(4) of this journal) discussing the design, implementation and analysis of a process evaluation as an integral part of the RIPPLE (Randomized Intervention of Pupil Peer-Led sex Education) Study. The methods for integrating process and outcome data are outlined and findings are presented to illustrate how integration of process and outcome data enabled exploration of: 1) variation between schools in the extent and quality of the intervention implemented, and any impact of this on outcomes; 2) the processes by which the intervention might affect outcomes;

and 3) the extent to which different subgroups of students or schools may have responded differently to the intervention. The article concludes with a discussion of the methodological issues arising from this attempt at integrating process and outcome evaluations within the design of an RCT.

KEYWORDS: evaluation methods, peer education, process evaluation, randomized controlled trial, sex education

Introduction and Aims

This is the second of two articles (the first was published in issue 10(4) of this journal) discussing the design, implementation and analysis of a process evaluation as an integral part of the RIPPLE (**R**andomized **I**ntervention of **P**upil **P**eer-Led sex **E**ducation) Study, a randomized controlled trial (RCT) of peer-led sex education conducted in English schools in 1997–2002. The first article (Oakley et al., 2004) described the design and implementation of the process evaluation. This article focuses on the integration of process and outcome data in the trial analysis.

Definitions of, and approaches used to carry out, process and outcome evaluations differ. In this article outcome evaluations are defined as those examining questions about the impact of interventions, while process evaluations examine how interventions are delivered and received, and how and why an intervention might (or might not) be effective (Tones and Tilford, 1994). Within the terminology of evaluation research, ‘process evaluation’ is a term which has some overlap with the notion of ‘implementation’ research, popular among North American evaluators during the 1960s–1980s period of social policy experimentation (Rossi et al., 2004; Yeaton, 1985). Most definitions of process evaluation include describing the implementation of programmes as a specific aim (see e.g. McGraw et al., 1989), as we did previously (Oakley et al., 2004). It has been argued that process evaluations have a particular role to play within RCTs of social interventions (Gueron, 2002).

The integration of process with outcome evaluations has been particularly recommended as a way of addressing some of the criticisms of RCTs for evaluating social interventions (Campbell et al., 2000; Oakley, 2000; Wight and Obasi, 2003). The practice of including the collection of process data in trials may be increasing in line with these recommendations. Of RCTs or controlled trials on a database of health promotion and public health trials maintained by the EPPI-Centre at the University of London Institute of Education, 58 percent (788/1351) included some element of process evaluation. However, the process element of most of these studies is limited and does not include the integration of process and outcome data. Notable exceptions include Hollister et al. (1979), McGraw et al. (1989), Oakley (1992), Wiggins et al. (2004), and Wight and Obasi (2003).

This article has three aims: to outline the methods developed to link process and outcome data in the RIPPLE trial; to present the findings of this analysis; and to explore some of the methodological issues that arose, especially in using the process data to explain the impact on trial outcomes.

The RIPPLE Study

The RIPPLE Study is a cluster RCT involving 27 coeducational comprehensive schools in England in order to evaluate the effectiveness of peer-led sex education in decreasing risky sexual behaviour. Full descriptions of the design of the study, methods used to collect and analyse data and findings from the process evaluation can be found elsewhere (e.g. Forrest et al., 2002; Stephenson et al., 2003; Strange et al., 2002).

Fourteen of the 27 schools that took part in the RIPPLE study were randomly assigned to receive the intervention comprising a programme of peer-led sex education. Training was provided by a team of external trainers to volunteer Year 12 students (aged 16/17 years) who delivered a short programme of three classroom sessions focusing on relationships, sexually transmitted infections (STIs) and contraception to two successive cohorts of Year 9 students (13/14 year olds). Schools, randomized to the control group, continued with their usual teacher-led sex education. Follow-up surveys, involving self-administered questionnaires, were carried out in both intervention and control schools 6 and 18 months post-intervention. Further follow-up of these students is ongoing and outcomes at aged 18–20 will be reported at a later date.

A total of 9508 students were eligible to take part in the study at baseline. Of these, 7770 (81.7 percent) students completed questionnaires at 6 months and 6656 (70.0 percent) at 18 months post-intervention. Findings at age 16 indicated that, compared with teacher-led sex education, peer-led sex education increased satisfaction with sex education and knowledge about STIs, but did not influence condom or contraception use at first or last sex or otherwise impact on the quality of sexual relationships or sexual experiences among boys or girls. Girls reported increased confidence in using condoms but lower confidence in refusing unwanted sexual activity. Peer-led sex education also reduced the proportion of girls reporting having experienced heterosexual intercourse by age 16. The proportion of girls reporting unintended pregnancy was smaller in the peer-led arm, although the study was not powered to find significant differences in this outcome for girls at this age (Stephenson et al., 2004).

The Process Evaluation

The process evaluation was designed to: document how the peer-led intervention, as well as the provision of teacher-led sex education, was implemented in intervention and control schools respectively; compare the two approaches to sex education; and collect information about the experiences of the different groups of research participants. These topics were explored using a wide range of methods: self-administered questionnaires; focus groups; interviews; observations; and researcher field-notes. These were used in order to explore the perspectives of the three key groups participating in the research: teachers and other school staff; peer educators; and other students. Full details of these methods and the data collected have previously been summarized (Oakley et al., 2004).

In integrating our process and outcome data, we decided to focus on three specific questions:

1. What impact might variation in the extent and quality of the intervention implemented have on outcomes?
2. What processes might explain the relationship between intervention and outcomes?
3. How did different subgroups of students (e.g. female and male students, those more or less engaged with school) or schools (e.g. those serving more or less 'at risk' populations) respond to the peer- and teacher-led sex education and how might the effectiveness of peer-led sex education differ according to these subgroups?

These three questions were explored because of their potential to explain the findings of the outcome evaluation already referred to. Key outcome findings from the RIPPLE study include differences in the impact of the peer-led sex education on girls and boys, and the mixed pattern of results in terms of behavioural and attitudinal outcomes. Outcome findings of the RIPPLE study were analysed using an 'intention-to-treat' strategy, the main statistical approach for RCT analyses. In this, individuals/clusters are analysed according to the group (intervention or control) to which they were originally allocated, regardless of whether the intervention was actually received and despite the fact that there may be no, or less, impact on those who did not receive the intervention. Our process data allowed us to explore the extent to which the peer-led sex education programme was implemented, and therefore enabled us to examine the extent to which 'on-treatment' analyses, whereby outcomes are analysed on the basis of receiving the intervention, rather than allocation, might yield different results.

The next section of the article looks at each of the three questions listed in turn, describes the methods we used to address the integration of process and outcome data, and discusses the results of doing this. The statistical approaches used in these analyses raise a number of issues. These include the possibility of bias resulting from the use of on-treatment analysis: any conclusions drawn must be treated with caution, since the schools and young people receiving peer-led sex education who were included in these analyses may differ systematically from those in the comparison group. Similar caution should be applied to findings from tests of interactions involving comparisons between non-randomly assigned groups of individuals. There is also the danger of over-interpreting 'chance' findings arising from multiple interaction tests and subgroup analysis. A final section of the article summarizes our main findings and the methodological challenges of the approaches used.

What Impact might Variation in the Extent and Quality of the Intervention Implemented have on Outcomes?

The extent to which the peer-led intervention was implemented as planned has been reported in detail (Oakley et al., 2004). To summarize: in one of the 14 schools randomized to receive the intervention, the peer-led programme was not implemented with either cohort of students; in two other schools, some classes in

one cohort did not receive any peer-led sex education; and in 11 schools, peer-led sex education was delivered to all Year 9 classes. Overall 84 percent of students in intervention schools reported having received some peer-led sex education and the proportion in the 13 schools in which at least some peer-led sex education was delivered varied between 51 and 97 percent.

Based on analysis of our process data, we hypothesized that the peer-led sex education might be more effective in schools in which the programme was most fully implemented and for students who reported receiving the peer-led sex education than for those who did not. We were also interested in exploring whether the inclusion, in our intention-to-treat analysis, of schools in which the peer-led intervention had not been implemented, and of students who reported not having received the sex education, might have resulted in a failure to detect the impact on some outcomes in other schools/students.

To explore this, process evaluation data were combined with outcome data in two on-treatment analyses and compared with the results of the intention-to-treat analysis (Table 1). The first on-treatment analysis, at the student level, looked at outcomes for intervention school students, who reported having actually received peer-led sex education, compared to outcomes for all control group students. The second on-treatment analysis, at the school level, examined outcomes for all students in the 11 intervention schools where peer-led classes were delivered to all Year 9 classes, again compared to outcomes for all control school students. As in the intention-to-treat analysis, clustering of individuals within schools was taken account of by using the generalized equations methodology of Liang and Zeger (1986) or by using the survey analysis methodology of STATA 7 (see Stephenson et al., 2003). Eleven key outcomes were examined and these are listed in Table 1. Odds ratios (ORs)/coefficients for outcomes in the on-treatment analysis were compared to those for the intention-to-treat analyses. Where ORs for both on-treatment analyses were higher or lower, thereby suggesting that the peer-led programme, if more consistently implemented, might have an even greater impact on these outcomes, these are highlighted in bold in Table 1 and summarized subsequently. All analyses were adjusted for the students' socioeconomic status and attitudes to school as these are key markers of social exclusion (Bonell et al., 2003) and were likely to be important confounders. A question about family housing tenure was used as a measure of students' socioeconomic status, with students reporting that they lived in non-privately owned accommodation considered socially disadvantaged. Students' responses to a question about whether they liked school or not were used to indicate those who were dissatisfied with school.

Table 1 shows higher odds ratios for both on-treatment analyses than for the intention-to-treat analyses for the outcomes knowledge of STIs, methods to prevent STIs and the timing of emergency contraception. This suggests that more consistent implementation of the peer-led programme might have a greater impact on these outcomes. However, as with the intention-to-treat analysis, the on-treatment analysis found that the peer-led intervention had no effect on condom use at first heterosexual intercourse by the age of 16, and little effect

on condom or contraceptive use at last sex. This provides no evidence that more consistent implementation of the peer-led sex education would result in impact on these outcomes.

Gender differences were found for some outcomes. The intention-to-treat analysis showed that peer-led sex education reduced the likelihood of girls reporting having had heterosexual sex by second follow-up, reduced girls' confidence in refusing sex (borderline significance) and increased their confidence about negotiating condom use. However, the same analysis found that peer-led sex education had no effect on these outcomes for boys. For most of these outcomes, the ORs from both on-treatment analyses were similar to those from the intention-to-treat analyses, providing no evidence that more consistent implementation of the peer-led education would increase the impact on these outcomes. However, ORs for the on-treatment analyses were higher than those for the intention-to-treat analysis for the outcome 'no regret at first sex', and lower for the outcome 'had heterosexual sex', raising the possibility that more consistent implementation of the peer-led sex education may reduce the likelihood of boys regretting their first experience of sex and reporting having had heterosexual sex by second follow-up.

Our process data also provided information regarding the extent to which the whole peer-led programme was implemented in the different schools. In addition to the delivery of sex education classes to the Year 9 students, the intervention involved the training of peer educators. Full implementation was defined as involving delivery of: all the elements of the pre-training sessions (preparation for needs assessment, needs assessment with Year 9 students, and analysis of needs assessment/training needs; two days of training in the local community; and at least three peer-led sex education sessions for all groups of Year 9 students). Analyses of observation and researcher field-notes suggested the following as factors associated with programme implementation: the size of the sixth form; Year 12 students' willingness to volunteer as peer educators; the enthusiasm of the main contact teachers about the study, and communication between these teachers and other staff about the programme; continuity of staffing in schools during the period of the study programme; and the quality of communication between the external training team and the main contact teachers.

To explore whether more extensive implementation of the peer-led programme might result in greater impact on outcomes, we used information from researcher field-notes to rank intervention schools in terms of the extent of programme implementation. We then ranked schools for each outcome and calculated the correlation between the rank of the outcome and the rank of the fidelity of programme implementation.

The knowledge outcomes were positively correlated with the fidelity of programme implementation, but the correlations were only significant for knowledge of methods to prevent STI acquisition (girls: $r=0.66$, $p=0.01$; boys: $r=0.68$, $p=0.01$). This supports the hypothesis that better implementation of the peer-led programme might increase students' knowledge. However, it is possible that both knowledge and fidelity of programme implementation are related to a third confounding factor. One potential confounder is the proportion of five or

Table 1. Comparison of Intention-to-Treat with On-Treatment Analysis

	Control		Intention-to-treat analysis		On-treatment (pupil level)*		On-treatment (school level)**	
	%	(n)	Intervention % (n)	Adjusted effect for intervention OR/coefficient (95% CI)	Intervention % (n)	Adjusted effect for intervention OR/coefficient (95% CI)	Intervention % (n)	Adjusted effect for intervention OR/coefficient (95% CI)
Satisfaction with sex education ^{ab}								
Girls	N=1540	N=1985	N=1985	0.16 (0.06, 0.25)	N=1755	0.17 (0.07, 0.26)	N=1686	0.17 (0.08, 0.27)
	3.30	3.47	3.47		3.47		3.47	
Boys	N=1623	N=2058	N=2058	0.17 (0.10, 0.25)	N=1730	0.20 (0.12, 0.28)	N=1722	0.18 (0.10, 0.26)
	3.33	3.51	3.51		3.53		3.51	
Knowledge of methods to prevent STIs								
Girls	N=1372	N=1758	N=1758	1.34 (0.97, 1.84)	N=1590	1.45 (1.08, 1.95)	N=1512	1.56 (1.17, 2.08)
	77.84 (1068)	82.31 (1447)	82.31 (1447)		83.71 (1331)		84.79 (1282)	
Boys	N=1419	N=1841	N=1841	1.31 (1.02, 1.68)	N=1588	1.42 (1.13, 1.79)	N=1547	1.43 (1.14, 1.78)
	64.06 (909)	68.71 (1265)	68.71 (1265)		70.72 (1123)		70.78 (1095)	
Girls	N=1377	N=1766	N=1766	1.25 (0.91, 1.71)	N=1602	1.30 (0.95, 1.78)	N=1521	1.39 (1.04, 1.87)
	60.42 (832 ^c)	65.07 (1149 ^c)	65.07 (1149 ^c)		66.29 (1062 ^c)		67.98 (1034 ^c)	
Boys	N=1383	N=1809	N=1809	1.12 (0.81, 1.55)	N=1567	1.16 (0.85, 1.59)	N=1520	1.20 (0.87, 1.66)
	49.75 (688 ^c)	43.95 (795 ^c)	43.95 (795 ^c)		55.07 (863 ^c)		55.86 (849 ^c)	
Knowledge of timing of contraception								
Girls	N=1391	N=1777	N=1777	1.08 (0.78, 1.50)	N=1609	1.16 (0.86, 1.56)	N=1529	1.18 (0.88, 1.58)
	65.06 (905)	67.25 (1195)	67.25 (1195)		68.43 (1101)		68.74 (1051)	
Boys	N=1397	N=1824	N=1824	1.25 (0.89, 1.75)	N=1576	1.32 (0.94, 1.86)	N=1529	1.31 (0.95, 1.83)
	32.64 (456)	36.95 (674)	36.95 (674)		37.94 (598)		37.87 (579)	
Girls	N=1292	N=1715	N=1715	0.05 (0.01, 0.09)	N=1553	0.05 (0.02, 0.09)	N=1479	0.05 (0.01, 0.09)
	4.14	4.14	4.14		4.15		4.17	
Boys	N=1278	N=1760	N=1760	0.03 (-0.04, 0.06)	N=1527	0.03 (-0.03, 0.09)	N=1487	0.04 (-0.03, 0.10)
	4.36	4.38	4.38		4.38		4.39	

Confidence about refusing sex	Girls	N=1293	N=1716	0.86 (0.74, 1.00)	N=1554	0.83 (0.71, 0.96)	N=1480	0.86 (0.73, 1.01)
	Boys	83.68 (1082 ^d)	79.66 (1367 ^d)		80.05 (1244 ^d)		80.14 (1186 ^d)	
		N=1271	N=1757	1.01 (0.84, -1.21)	N=1525	1.02 (0.85, 1.22)	N=1484	1.03 (0.85, 1.24)
No regret at first sex	Girls	68.61 (872 ^d)	67.33 (1183 ^d)	0.93 (0.70, 1.24)	68.00 (1037 ^d)	0.98 (0.77, 1.25)	67.59 (1003 ^d)	1.00 (0.76, 1.33)
	Boys	N=508	N=572		N=507		N=487	
		64.37 (327)	63.99 (366)	1.03 (0.68, 1.56)	65.09 (330)	1.10 (0.75, 1.62)	65.30 (318)	1.22 (0.89, 1.84)
Had sex	Girls	80.65 (321)	82.05 (393)	0.82 (0.68, 0.98)	82.91 (325)	0.79 (0.65, 0.95)	84.11 (323)	0.81 (0.67, 0.98)
	Boys	N=1297	N=1614		N=1459		N=1388	
		43.33 (562)	37.77 (610)	0.92 (0.65, 1.28)	36.87 (538)	0.84 (0.60, 1.67)	37.68 (523)	0.86 (0.62, 1.20)
Condom use at first sex	Girls	N=1300	N=1700	0.86 (0.67, 1.12)	N=1470	0.91 (0.73, 1.14)	N=1411	0.89 (0.68, 1.16)
	Boys	34.15 (444)	31.94 (543)		30.27 (445)		30.62 (432)	
		N=550	N=602	0.95 (0.62, 1.43)	N=531	0.99 (0.64, 1.55)	N=516	1.00 (0.64, 1.55)
		76.36 (420)	75.42 (454)		76.46 (406)		76.16 (393)	
Contraception use at last sex	Girls	N=432	N=528	1.06 (0.70, 1.63)	80.65 (350)	1.13 (0.74, 1.71)	80.00 (336)	1.00 (0.66, 1.52)
	Boys	81.48 (352)	79.55 (420)		N=434		N=420	
		83.37 (341)	84.68 (431)	1.06 (0.70, 1.63)	85.21 (386)	0.91 (0.51, 1.63)	83.94 (371)	0.96 (0.53, 1.77)
		N=264	N=352		N=290		N=300	
Condom use at last sex	Girls	86.74 (299)	85.51 (301)	0.96 (0.74, 1.25)	85.86 (249)	0.98 (0.73, 1.31)	86.00 (258)	0.93 (0.71, 1.23)
	Boys	N=409	N=509		N=453		N=442	
		62.35 (255)	62.48 (318)	0.69 (0.47, 1.01)	62.69 (284)	0.66 (0.45, 0.97)	61.76 (273)	0.71 (0.47, 1.09)
		N=264	N=352		N=290		N=300	
		76.89 (203)	71.31 (251)		71.03 (206)		72.00 (216)	

Notes: (All adjusted analyses have been adjusted for tenure and dissatisfaction with school.)

*The effect of the peer-led intervention on those students who reported receiving peer-led sex education in Year 9, compared to all control students.

**The effect of the intervention on pupils in those schools where the peer-led sessions were delivered to all groups of Year 9 students (fully implemented), compared to all students in all control schools.

^aAll outcomes are measured at 18 months post-intervention except for satisfaction with education which is measured at 6 months post-intervention.

^bAverage of 5-point likert scales where 1='very unsatisfied' and 5='very satisfied' with sex education. For confidence about condom use 1=low confidence and 5=high confidence. Effect of intervention is coefficient.

^cThe number of students who answered three or four knowledge questions correctly.

^dThe number of students who answered 'easy' or 'very easy'.

more GCSE A–C grades achieved by students in a school, which was positively correlated with knowledge measures. Adjusting, in a regression analysis, for this measure of school effectiveness reduced the correlation between the knowledge measures and the fidelity of programme implementation. The correlation was no longer statistically significant. There was also a significant negative correlation between fidelity of programme implementation and the likelihood of boys reporting heterosexual sex by the second follow-up. These analyses raise the possibility that the peer-led programme, if fully implemented, might reduce the likelihood of boys as well as girls reporting having had heterosexual sex by age 16.

One limitation to the approaches described arises from the lack of data relating to the implementation of sex education in the control schools. In order to assess the impact on outcomes of the one experimental school which did not implement the peer-led sex education programme, an approach suggested by Sommer and Zeger (1991) was employed (see Bloom, 1984, for a similar approach). This identifies a subgroup of young people/schools in the control arm who are equivalent with regard to outcomes to those in the experimental arm who did not receive the peer-led sex education. These equivalent subgroups from both arms are then excluded from the analyses. The analysis was carried out for the behavioural outcomes only. It found little evidence that the peer-led sex education had any effect, even when findings were adjusted for the lack of programme implementation.

What Processes might Explain the Relationship between Intervention and Outcomes?

In order to develop hypotheses about those processes that might be important in determining the effectiveness of peer-led sex education, we drew on our previous analyses of focus group data from students and peer educators and researchers' observations of classes. These analyses highlighted factors that seemed important in determining the acceptability of sex education, both with respect to the peer-led approach and in relation to sex education in general. Students identified the personal characteristics of sex educators and their ways of interacting with students as crucial. They perceived peer educators to have greater expertise and respect for students than teachers, described peer educators as more confident, empathetic, caring and trustworthy, less moralistic and patronizing, as holding similar sexual values and as using humour in the sessions, making these more fun. Students felt that peer educators provided more relevant, detailed information and that they were learning something new. Weaknesses of peer educators included some being 'shy and embarrassed' and some struggling to manage the disruptive behaviour of some students. Students also identified the use of skill-based activities as important. These involved practising putting on a condom, and activities that involved moving around, working in small groups and discussion, all of which were elements in the peer-led programme (Forrest et al., 2002).

In order to explore the relationship between study outcomes and students' experiences of sex education, we drew on the findings of our analysis of focus groups and observation data and, where possible, identified relevant questions relating to the factors arising from this analysis, in the questionnaires. These factors were:

use of participative/active methods; practising key skills; being provided with key information; feeling positive about the sex educator; and feeling satisfied with the coverage of key topics. There were differences between experimental and control school students on all these dimensions of sex education in the direction of more satisfaction in the experimental schools (Table 2).

Table 3 provides the results of regression analyses used to examine the relationship between these key dimensions of sex education and whether sex education was delivered by peers or teachers and study outcomes. We explored which dimensions of sex education were most important in determining the effectiveness of peer-led sex education by, first, comparing ORs for the effect of the intervention on outcomes, and, second, with ORs after adjustment for each dimension of sex education (see Victora et al., 1997, for a similar approach to exploring the effect of mediating factors). These analyses indicated that adjustment for participatory/active methods changed the effect of the intervention on confidence with using condoms and use of condoms and other contraception at last sex. Adjustment for practising key skills changed the effect of the intervention on confidence with using condoms. Adjusting for students' evaluation of the sex educators, and their satisfaction with the coverage of key topics made less of a difference to impact. This suggests that the impact of peer-led sex education on sexual health outcomes was most consistently 'mediated' by students' taking part in participative/active methods and practising key skills.

We took this analysis further to explore whether the various different dimensions of sex education impacted on sex education differently depending on whether it was delivered by peers or teachers. To do this, we tested for interactions between trial arm, students' reports about the different dimensions of sex education and outcomes. A *p*-value of less than 5 percent was used as a guide to significance (highlighted in bold in Table 3), but with large numbers of tests for interactions being carried out, it was likely that significant interactions would be found by chance. To counter this, we looked for consistent patterns of interactions. Nine of the 20 interactions tested for between trial arm, participatory/active methods or practising key skills and outcomes were found to be significant. Significant interactions were found between trial arm and participatory/active

Table 2. Key Aspects of Sex Education Reported by Students in Control and Experimental Schools

<i>N</i> of students reporting	Control (<i>N</i> =3559) % (<i>n</i>)	Experimental (<i>N</i> =4211) % (<i>n</i>)
Participatory/active methods	25.0 (892)	62.0 (2611)
Practising key skills	43.4 (1270)	80.5 (3012)
Receiving key information	79.5 (2329)	89.8 (3357)
Positive about sex educator	58.1 (2070)	64.1 (2700)
Satisfaction with coverage of key topics	20.4 (728)	41.4 (1243)

Table 3. Effect of Sex Education on Outcomes with Adjustment for Key Dimensions of Sex Education

	Control % (n)	Intervention % (n)	Effect of intervention on outcome OR (95% CI)	Effect adjusted for participatory/ active methods OR (95% CI)	Effect adjusted for practising key skills OR (95% CI)	Effect adjusted for receiving key information OR (95% CI)	Effects adjusted for positive about sex educator OR (95% CI)	Effect adjusted for satisfaction with coverage of key topics OR (95% CI)	Effect adjusted for all aspects OR (95 CI)
Knowledge of STI prevention	70.88 (1984)	75.38 (2716)	1.32 (1.05, 1.66)	1.28 (0.98, 1.67)	1.25 (0.96, 1.61)	1.26 (1.01, 1.58)	1.33 (1.07, 1.67)	1.28 (1.02, 1.61)	1.21 (0.92, 1.60)
Knowledge of STIs	55.09 (1525)	59.40 (2126)	1.17 (0.87, 1.61)	1.09 (0.79, 1.51)	1.07 (0.77, 1.49)	1.14 (0.84, 1.54)	1.20 (0.88, 1.64)	1.20 (0.87, 1.64)	1.02 (0.72, 1.43)
Knowledge of timing of ECP	48.82 (1365)	51.96 (1873)	1.13 (0.84, 1.53)	1.08 (0.79, 1.47)	1.15 (0.86, 1.54)	1.12 (0.82, 1.52)	1.12 (0.83, 1.50)	1.16 (0.86, 1.55)	1.08 (0.83, 1.41)
Confidence about condom use	78.24 (2017)	79.02 (2749)	0.01 (-0.05, 0.07)	-0.05 (-0.10, 0.01)	-0.05 (-0.11, 0.02)	-0.003 (-0.06, 0.06)	0.01 (-0.04, 0.07)	-0.01 (-0.06, 0.05)	-0.10 (-0.16, 0.04)
Confidence about saying no	76.21 (1960)	73.43 (2553)	0.92 (0.82, 1.04)	0.87 (0.77, 0.99)	0.86 (0.76, 0.97)	0.91 (0.81, 1.02)	0.94 (0.83, 1.05)	0.91 (0.80, 1.04)	0.81 (0.71, 0.92)
Quality of sexual experience	71.55 (649)	72.22 (759)	1.04 (0.81, 1.34)	0.92 (0.72, 1.18)	1.08 (0.81, 1.42)	1.02 (0.77, 1.34)	1.02 (0.79, 1.31)	1.03 (0.80, 1.34)	1.00 (0.74, 1.34)
Had sex by follow-up 2	38.72 (1008)	37.78 (1153)	0.82 (0.64, 1.03)	0.77 (0.59, 0.99)	0.82 (0.67, 1.00)	0.90 (0.78, 1.07)	0.87 (0.68, 1.10)	0.87 (0.68, 1.10)	0.78 (0.62, 0.98)
Used a condom at first sex	78.66 (774)	77.35 (874)	0.92 (0.72, 1.18)	0.84 (0.64, 1.10)	0.82 (0.62, 1.10)	0.88 (0.68, 1.15)	0.87 (0.68, 1.12)	0.85 (0.65, 1.11)	0.75 (0.53, 1.06)
Used a condom at last sex	70.13 (432)	67.94 (517)	0.90 (0.71, 1.15)	0.74 (0.56, 0.97)	0.86 (0.64, 1.15)	0.88 (0.69, 1.13)	0.81 (0.63, 1.04)	0.79 (0.62, 1.02)	0.70 (0.50, 0.96)
Used contraception at last sex	86.36 (532)	85.81 (653)	1.03 (0.75, 1.41)	0.80 (0.54, 1.18)	1.02 (0.72, 1.43)	0.99 (0.72, 1.36)	0.91 (0.63, 1.30)	0.98 (0.66, 1.43)	0.85 (0.55, 1.34)

Note: Bold cells are those with significant interactions between trial arm, dimension of sex education and outcomes.

methods and practising key skills for the outcomes knowledge of STI prevention, knowledge of STIs, using a condom at last sex and using contraception at last sex. A significant interaction was also found between trial arm and participatory/active methods for the outcome using a condom at first sex (see cells in bold in Table 3). Two of the ten interactions tested for between trial arm, receiving key information and outcomes, were significant: knowledge of STI prevention and using contraception at last sex. Three of the ten interactions tested for between trial arm and satisfaction with coverage of key topics and outcomes were significant: knowledge of timing of the emergency contraceptive pill; using a condom at last sex; and using contraception at last sex. All these interactions suggest that, when students reported that the sex education involved the particular dimension highlighted, the peer-led sex education intervention had a positive effect on the outcome. However, when the sex education was not reported as involving the particular dimension, the effect of the intervention was negative. For example, where participatory/active methods were used, the peer-led intervention increased the likelihood, compared to teacher-led sex education, of students using contraception at last sex (OR 1.49 (0.95,2.34)). Where participative/active methods were not used, the peer-led intervention reduced the likelihood of students using contraception at last sex (OR 0.54 (0.36,0.80)) (<0.001).

In summary, these findings suggest that, when students reported that peers delivered the sex education using participatory and skills-based approaches, the sex education was more effective than when teachers employed similar methods. However, when peers and teachers did not employ these approaches, then sex education was more effective if delivered by teachers. Limitations to these analyses are outlined in the discussion section.

How Do Different Subgroups of Students or Schools Respond to the Peer- and Teacher-Led Sex Education?

Subgroups of Students

In integrating the process and outcome data to explore possible differences between subgroups of students, we began with our researcher field-notes on classroom observations. The aim was to use these data about the way that particular groups of students responded to the peer-led sex education sessions to develop hypotheses about how peer-led sex education might be more or less effective for different subgroups of students. As with the previous analysis, these hypotheses could then be tested using quantitative data collected in the questionnaire surveys. It was clear from our observation data that, for example, girls and boys responded differently to the peer-led sessions, and were more or less engaged by different aspects of the lessons. Some students, often those who were generally disengaged from school and labelled as generally 'problematic' by teachers, found it difficult to engage in activities, and in some cases were disruptive. Students who appeared either already sexually experienced or not yet interested in sex and/or relationships were also observed to be less engaged by the lessons.

Students' responses in the baseline questionnaire (completed before the intervention) were used to identify relevant subgroups. The subgroups were defined by key markers of social exclusion – housing tenure as an index of social disadvantage, dislike of school and educational aspirations. Responses to questions asking whether students had had sex and, if not, at what age they thought they might have sex, were used to indicate students' interest in sex/sexual maturity.

As a first step to testing our hypotheses about the differential responses of specific groups of students, we drew on data from questions relating to students' responses to the different types of sex education, and tested for interactions between subgroups of students, type of sex education (peer- or teacher-led) and different aspects of sex education. Table 4 shows the results of our analysis of trial arm, social exclusion and aspects of sex education. Peer-led sex education significantly increased the likelihood of socially advantaged students, and those who anticipated having sex at age 14/15 years, compared to other students, reporting having taken part in participatory/active methods; it increased the chance of students who liked school and students with positive educational aspirations, compared to other students, reporting practising key skills; and it increased the likelihood of students with positive educational aspirations reporting satisfaction with the coverage of key topics. These findings suggest that peer-led sex education may be less good at increasing the engagement of students who are most at risk of poor sexual health: those who dislike school; who have already had sex by age 13/14 years; and who have negative educational aspirations. Although we observed differences in the way that girls and boys responded to different aspects of the peer-led classes, the statistical analyses found no significant interactions between gender and type of sex education.

In order to explore whether peer-led sex education itself was more or less effective for these different subgroups of students, we went on to test for interactions between outcomes and subgroups defined by gender, attitude to school, housing tenure, educational aspirations and actual/anticipated age for first sex. There were no significant interactions between gender and type of sex education for any of the outcomes. This suggested that the peer-led intervention was not significantly more or less effective for either boys or girls. Although the main intention-to-treat analysis found the peer-led sex education to be significantly more effective than teacher-led sex education with regard to some outcomes for girls but not boys and vice versa, these gender differences were small and the effects were in the same direction for both boys and girls. Nor were there any significant interactions between aspirations towards education and type of sex education for any of the outcomes. Just one significant interaction was found between attitude to school and the outcome 'used contraception at last sex'. The peer-led intervention was significantly more effective for those students who reported liking school (OR 1.60 (0.91,2.83)) than for those who were ambivalent or disliked it (OR=0.76 (0.46,1.25), 0.79 (0.46,1.36) respectively) ($p= 0.004$). This could reflect a chance finding.

Significant interactions were found between housing tenure and type of sex education for three of the 11 outcomes: knowledge of emergency contraception,

Table 4. Students' Reports of Key Aspects of Sex Education According to Type of Sex Education, Where a Significant Interaction ($p < 5\%$) was Found

	Control		Experimental	
	Privately owned (N=2022)	Non-privately owned (N=837)	Privately owned (N=2560)	Non-privately owned (N=846)
% Participatory/active methods	25.67 (519)	26.05 (218)	64.10 (1614)	60.52 (512)
<i>Educational aspirations</i>				
% Practising key skills	Negative 49.59 (60/121)	Positive 42.83 (765/1786)	Negative 76.98 (97/126)	Positive 81.64 (1979/2424)
% Satisfaction with coverage of key topics	22.92 (33/144)	19.95 (427/2140)	32.45 (49/151)	42.25 (1139/2696)
<i>Attitude to school</i>				
% Practising key skills	Like (N=1265) 42.29 (535)	Ambivalent (N=958) 43.84 (420)	Like (N=1724) 81.00 (1412)	Ambivalent (N=1190) 80.84 (962)
		Dislike (N=483) 45.13 (958)		Dislike (N=614) 77.85 (478)
<i>Timing of sex</i>				
% Participatory/active methods	16+ yrs (N=1903) 25.22 (480)	14/15 yrs (N=508) 24.21 (123)	16+ yrs (N=2353) 62.01 (1459)	14/15 yrs (N=629) 70.43 (433)
		Had sex (N=211) 31.75 (67)		Had sex (N=244) 69.26 (169)

regret at first heterosexual sex and condom use at last heterosexual sex. The peer-led intervention was significantly less effective among socially disadvantaged students in increasing knowledge of emergency contraception (OR=0.98 (0.73,1.31) vs 1.20 (0.86,1.86) $p=0.021$), more effective in reducing regret following first heterosexual sex (1.26 (0.94,1.69) vs 0.89 (0.66,1.21) $p=0.005$) and increasing condom use at last heterosexual sex (1.22 (0.92,1.63) vs 0.75 (0.56,1.00) $p=0.032$).

Subgroups of Schools

As well as questions about different groups of students, there are questions about how differences between types of schools may be contributing to the outcome findings of the RIPPLE study. We used information about the proportions of students on the school roll eligible for free school meals (an indicator of poverty¹) to categorize all the study schools into three groups: those with 'high', 'medium' and 'low' risks of adverse sexual health outcomes. Schools with 9 percent or less of students eligible for free school meals were considered to serve populations with a 'low' risk of poor sexual health; those with 10–20 percent of students eligible for free school meals as at 'medium' risk, and those with more than 20 percent of students eligible for free school meals as 'high' risk.

We carried out tests for interactions between 'free school meals' and type of sex education (peer- or teacher-led) for the 11 key outcomes (Table 5). However, the only 'high-risk' school in the experimental arm did not implement the peer-led intervention at all; this meant that no comparison between the two arms of the trial was possible for the 'high-risk' group.

Of the 11 interactions tested, we found significant interactions for three outcomes: satisfaction with sex education, and knowledge of STIs and emergency contraception. The peer-led intervention was more effective in increasing knowledge for schools serving 'medium-risk' populations, and students in these schools were more satisfied with their sex education (see Table 5). While knowledge was similar in 'low-' and 'medium-risk' schools following peer-led sex education, it was much lower in 'medium-risk' compared to 'low-risk' schools in the teacher-led sex education schools. The peer-led sex education intervention was not more or less effective in influencing confidence, regret following first heterosexual intercourse, or timing of first sex or condom/contraceptive use at first or last sex in the different groups of schools.

These findings from the analysis exploring the differential effects of peer-led sex education on subgroups of schools appear to contradict the findings from the analysis exploring differential effects in subgroups of students. While, overall, peer-led sex education appeared less effective in increasing knowledge among socially disadvantaged compared to advantaged students, it was more effective in increasing knowledge in schools serving 'medium-risk' rather than 'low-risk' populations. Closer examination of the knowledge levels for different groups of students in schools serving different populations showed that in 'medium-risk' schools, peer-led sex education was significantly more effective than teacher-led sex education in increasing knowledge for socially disadvantaged students (e.g. 26.7 percent vs 22.0

percent reported good knowledge of emergency contraception), whereas in ‘low-risk’ schools, peer-led sex education was significantly less effective at increasing knowledge for disadvantaged students (29.6 vs 37.2 percent). This suggests that peer-led sex education may be more effective in increasing knowledge among the most disadvantaged students in schools serving ‘medium-risk’ populations compared with those serving ‘low-risk’ populations.

Schools in both the control and experimental arm of the study were also categorized into three groups according to the proportion of students achieving five or more GCSEs grade A–C. Again we tested for interactions between ‘academic achievement’ and type of sex education for the 11 key outcomes. There were no significant interactions between academic achievement and type of sex education for any of these outcomes.

We went on to carry out further analyses examining interactions between trial arm, type/subgroup of school and students’ reports of the sex education, with the aim of understanding why peer-led sex education might be more effective in ‘medium-risk’ as opposed to ‘low-risk’ schools. These analyses suggest that the increased effect of peer-led sex education in ‘medium-’ compared to ‘low-risk’ schools results from students’ greater satisfaction with peers than teachers in these schools, and the increased likelihood that peers were able to provide students with key information. Students from ‘medium-risk’ schools in the experimental arm were more likely than those from ‘low-risk’ schools to be positive about their sex educators (OR 1.36 (1.17,1.59)), and more likely to report receiving key information (OR 1.37 (0.81,2.31)). Those from ‘medium-risk’ control schools were less likely than those from ‘low-risk’ schools to be positive about their sex educators (OR 0.62 (0.38,1.01)) and less likely to report receiving key information (0.43 (0.26,0.71)).

Table 5. Relationship between Types of School, Type of Sex Education and Sexual Health Outcomes

Outcome	Control % (n)	Intervention % (n)	Odds ratios/coefficients for effect of intervention by subgroup (95% CI)
<i>Satisfaction with sex education^a</i>			
‘Low risk’	3.32 (0.53)	3.43 (0.53)	0.09 (–0.05, 0.22)
‘Medium risk’	3.29 (0.50)	3.57 (0.51)	0.26 (0.16, 0.35)
<i>Knowledge of STIs</i>			
‘Low risk’	61.4 (718/1170)	60.7 (1197/1971)	0.96 (0.64, 1.44)
‘Medium risk’	48.3 (565/1169)	60.8 (916/1506)	1.64 (1.14, 2.38)
<i>Knowledge of timing of emergency contraception</i>			
‘Low risk’	58.2 (690/1186)	53.4 (1059/1982)	0.81 (0.59, 1.10)
‘Medium risk’	40.6 (481/1186)	52.6 (798/1518)	1.58 (1.19, 2.10)

^a Average of 5-point likert scales where 1= ‘very unsatisfied’ and 5=‘very satisfied’ with sex education. For confidence about condom use, 1=low confidence and 5=high confidence. Effect of intervention is coefficient (Standard Deviation).

Discussion

A number of conclusions can be drawn from the attempt, described in this article, to integrate process and outcome data in the analysis of findings from the RIPPLE Study. We will first summarize our substantive findings and then discuss some of the methodological limitations and challenges encountered.

By integrating process with outcome data, we have been able to extend the outcome analysis to examine three particular issues. The first is the extent to which variations in the implementation of the peer-led programme across the experimental schools impacted on the study outcomes. This enabled us to explore the possibility that the overall conclusions about the effectiveness of the peer-led sex education arising from the original intention-to-treat analysis may have underestimated the impact of peer-led sex education or failed to detect the effect of such education on some outcomes as a result of poor or inadequate implementation of the programme in some schools or to some students. Findings from the 'on-treatment' analyses were generally consistent with those from the original intention-to-treat analysis. ORs for a few outcomes in the on-treatment analysis raised the possibility that, when fully implemented, peer-led sex education may have greater impact on increasing knowledge. There was little evidence that more consistent implementation of the peer-led programme would have increased the impact on most other outcomes. Two outcomes that did not show differences between control and experimental arms in the original intention-to-treat analysis were the likelihood of boys reporting having heterosexual sex by age 16 and boys regretting their first experience of sexual intercourse. In the 'on-treatment' analysis, ORs suggested that, if more consistently implemented, peer-led sex education might reduce the likelihood of boys reporting having sex by age 16 and reduce the chance of regret. Further analysis examining correlations between the fidelity of programme implementation and outcomes provided additional support for the conclusion that better implementation of the peer-led programme might reduce the likelihood of boys reporting having sex by age 16.

Second, we have explored some of the processes associated with the impact of the intervention on outcomes. This analysis is important as it illustrates the way in which the integration of process with outcome data can contribute to our understanding of the mechanisms by which an intervention might impact on outcomes. It thus provides evidence to challenge arguments about the inappropriateness of the RCT design for evaluating social interventions (e.g. Pawson and Tilley, 1997) which have suggested that RCTs cannot provide information about such mechanisms. Our analysis suggested that the use of participatory/active methods and activities to practise key skills may be particularly important in determining the effectiveness of peer-led sex education. Sex education delivered by peers was more effective with regard to some knowledge and behavioural outcomes (e.g. at increasing knowledge and condom/contraception use at last sex) than that delivered by teachers, where students reported the educators employing participatory/active methods and practising key skills. However, when sex education did not involve these key processes, sex education was more effective if delivered by teachers. This could be an important finding for policy and practice in school-based sex education, both in the UK and elsewhere.

Third, we have looked at the extent to which particular subgroups of students and schools may have been differently influenced by the peer-led sex education. Again this analysis goes some way to examining the mechanisms by which the intervention might impact on outcomes and challenges arguments that RCTs cannot consider the way in which a particular intervention may impact differently on different participants. When we looked at the way that different subgroups of students responded to the peer-led intervention, we found that more students receiving peer-led than teacher-led sex education reported this education as having involved participatory/active methods, the opportunity to practise key skills, and the provision of key information, and more expressed satisfaction with this information and with their sex educators generally. However, students who were most socioeconomically disadvantaged, most dissatisfied with school, had lowest educational expectations and were most likely to report sex, were significantly less likely than others to report experience of many of these key dimensions of sex education or to report satisfaction with sex educators or key information. These data suggest that the peer-led programme may have failed sufficiently to increase engagement among those most at risk of poor sexual health.

Few differences were found between different groups of students with regard to the impact of the peer-led sex education on outcomes. Peer-led sex education was not significantly more or less effective for boys than girls or for students with different expectations of education, although differences were found for students who were more or less socially disadvantaged. Overall, peer-led sex education was less effective at increasing knowledge among those who were most disadvantaged. Analysis exploring the effectiveness of peer-led sex education in different subgroups of schools found that students in schools serving populations of students with a 'medium risk' of poor sexual health outcomes were more likely than students from 'low-risk' schools to report greater knowledge. Further analysis indicated that peer-led sex education might be more effective at increasing the knowledge of disadvantaged students when delivered in schools serving medium-risk populations than in schools serving low-risk populations.

On the methodological front, it is clear that the analyses reported in this article would not have been possible without a detailed process evaluation as an integral part of the RIPPLE trial. Questions about variations in the implementation and impact of peer-led sex education, and about processes mediating the relationship between sex education and outcomes for students, would not have been apparent as a result of the analysis of the outcome data, and we would not have been able to draw on data from the trial to throw any light on these. Process evaluations are, in this sense, part of a more theory-based approach to evaluation (Patton, 1989; Stame, 2004), which acknowledges that social interventions 'are complex systems thrust amidst complex systems' (Pawson et al., 2004: iv). Within such a model, qualitative and observational data are used to generate theories about intervention implementation and efficacy which can then be applied in a combined statistical analysis. This goes beyond a traditional quantitative analysis of mediator effects since it interposes a theoretical explanatory model derived from the process evaluation itself.

Our approach to integrating process with outcome data followed this route. This use of qualitative data also has an important function in reducing the likelihood of bias resulting from 'data dredging' in the interpretation of qualitative process data that might arise if such data are analysed with the aim of explaining previously identified outcome findings (Wight and Obasi, 2003).

In combining process with outcome data, we have used data collected from observations and in focus groups to develop hypotheses which have then been tested using data collected in questionnaires. This approach relies on extensive pilot work to ensure that important aspects of the process are adequately captured in questionnaire measures. Some aspects of the peer-led sex education sessions highlighted by students in focus groups failed on this score, and are therefore not reflected in the statistical analysis. These included the greater empathy and trustworthiness of the peer educators as described by students, the use of fun and humour in the sessions, and the sharing of values between the students and their peer educators. Some aspects of school culture and ethos which may be important in determining the success or otherwise of peer-led sex education are also not reflected in our analysis. Although we did collect this information and attempted to code observation and focus group data in order to combine these with outcome data, this proved difficult; many such data were not collected systematically enough to use in the statistical analysis.

The statistical analysis we carried out is complex. Several key methodological issues should be highlighted. These include the value of 'on-treatment' analysis; the use of regression analysis; the procedure of testing for interactions, and the use of subgroup analyses. The most fundamental of these issues concerns the use of 'on-treatment' analysis as distinct from the conventional 'intention-to-treat' approach used in the analysis of RCTs. These terms reflect what might fairly be called the 'medical model' of RCT design, in which interventions are conceived in terms of clinical 'treatment'. The appropriateness of the terms, the advantages and disadvantages of the use of the two approaches and the value of different approaches to on-treatment analysis have, so far as we are aware, been little explored in the analysis of social intervention trials. The value of some approaches to on-treatment analysis may also differ for cluster as opposed to individual RCTs. In cluster trials one might expect a diffused impact of an intervention on those who do not directly receive the intervention. Therefore, on-treatment analyses that exclude individuals may be less likely to show different findings from intention-to-treat analysis in cluster RCTs than in RCTs that randomize individuals. It is also important to recognize that while 'on-treatment' analyses can be illuminating, they lose the benefits of the original random assignment, since the comparisons are no longer between the original randomly assigned groups, and there is therefore the potential for bias. For this reason, findings from any 'on-treatment' analysis need to be interpreted with caution (Pocock and Abdalla, 1998).

Our integration of process and outcome data also involved a number of subgroup analyses and tests for interactions. While the appropriateness of drawing conclusions from these types of analyses is controversial (Early Breast Cancer Trialists' Collaborative Group, 1990; Yusuf et al., 1984), care was taken to follow guidelines suggested by Oxman et al. (1992). Specifically, the number of

subgroup analyses and interactions has been clearly indicated and in all cases the analysis was based on prior hypotheses. In addition all interactions between the trial arm and the subgroups were formally tested in suitable regression analyses (Pocock et al., 1987). Any subgroup-specific effects and confidence intervals were used to help interpret the differential effects after the formal test of interaction. Although a 5 percent level of significance was used as a guide to both the strength of association and to which results to report, we also looked for consistency in the patterns of interactions found, and used this as evidence of a genuine as opposed to chance finding. However, it should also be remembered that the trial was not powered with interactions in mind and therefore any lack of effect should also be interpreted with caution. Where tests for interactions involved comparisons between non-randomly assigned groups (as with our analyses of the processes involved in peer- and teacher-led sex education and their relationships to outcomes), the potential for bias means that findings should be interpreted with additional care.

Conclusion

The analysis has addressed some of the limitations of RCTs highlighted by those who criticize this design for its tendency to omit process data, and to ignore the sensitivity of interventions to personal and contextual factors. In particular, we have been able to unpack some of the 'black box' in which causal mechanisms are commonly trapped, and we have examined some of the contextual and subgroup-specific factors which may underlie the capacity of specific interventions to affect outcomes. One value of the approach which combines process and outcome data is in suggesting ways in which more appropriate and therefore more effective interventions may be developed. Some of the findings in this article suggest that this type of peer-led intervention may be most effective in schools where the programme can be fully implemented. Ensuring full implementation is a complex business, requiring continuity of staffing during the recruitment and training of peer educators and over the period in which sex education sessions are delivered, as well as good lines of communication between school staff and trainers. Where schools have small sixth forms and uneven ratios of boys/girls in younger classes and sixth forms, then alternative approaches to recruiting peer educators (e.g. from other schools/colleges) may be needed.

With regard to methods of sex education, the use of participatory and skill-based activities by peer educators appears to be important. Finally, while this type of peer-led programme may be more effective in schools serving more disadvantaged populations of students, the overall effectiveness of sex education across all types of schools may be enhanced by developing ways of engaging the most socially excluded students who are at most risk of poor sexual health outcomes. More research eliciting the views of particular groups of students about sex education is needed in order to understand how peer-led programmes could be amended to increase the engagement of socially disadvantaged students.

Our analysis also points to ways in which the design of process evaluations within trials may be improved, so as to increase the potential for integrating process

and outcome data. For example, thorough pilot work is crucial in order to develop questionnaire measures for capturing processes, and to design appropriate coding procedures for the systematic collection of observational, interview and focus group data. In the RIPPLE Study, collection of information about the presence of particular students in sessions observed or in focus groups would have enabled the linkage of process and outcome data for individual students. It would also have been useful to have developed a tool for asking peer-educators to collect basic information about each of their classes. These issues point to the general desirability of giving as much prior thought to the design of process as of outcome evaluations, especially when the two types of data are to be combined statistically in the analysis of findings.

Acknowledgement

The RIPPLE Study was funded by the Medical Research Council.

Note

1. In UK schools, children are eligible for free school meals if their parents/carers receive income support or income-based job seekers' allowance. In this context, the proportion of students receiving free school meals is an accepted and accessible indicator of the demographic profile of a school's intake.

References

- Bloom, H. (1984) 'Accounting for No-Shows in Experimental Evaluation Designs', *Evaluation Review* 8(2): 225–46.
- Bonell, C., V. Strange, J. Stephenson, S. Forrest, A. Copas, A. Oakley, A. Johnson and S. Black (2003) 'The Effect of Various Dimensions of Social Exclusion on Young People's Risk of Teenage Pregnancy: Development of Hypotheses from Analysis of Baseline Data Arising from a Randomized Trial of Sex Education', *Journal of Epidemiology and Community Health* 57: 871–6.
- Campbell, M., R. Fitzpatrick, A. Haines, A.L. Kinmonth, P. Sandercock, D. Spiegelhalter and P. Tyrer (2000) 'Framework for Design and Evaluation of Complex Interventions to Improve Health', *British Medical Journal* 321: 694–6.
- Early Breast Cancer Trialists' Collaborative Group (EBCTCG) (1990) *Treatment of Early Breast Cancer*, vol. 1, *Worldwide Evidence 1985–1990*. New York: Oxford University Press.
- Forrest, S., V. Strange, A. Oakley and the RIPPLE Study Team (2002) 'A Comparison of Students' Evaluations of a Peer-Delivered Sex Education Programme and Teacher-Led Provision', *Sex Education* 2(3): 195–214.
- Gueron, J.M. (2002) 'The Politics of Random Assignment: Implementing Studies and Affecting Policy', in F. Mosteller and R. Boruch (eds) *Evidence Matters: Randomized Trials in Education Research*, pp. 15–49. Washington, DC: Brookings Institution Press.

- Hollister, R.G., P. Kemper and J. Wooldridge (1979) 'Linking Process and Impact Analysis: The Case of Supported Work', in T.D. Cook and C.S. Reichardt (eds) *Qualitative and Quantitative Methods in Evaluation Research*, pp. 140–9. Beverly Hills, CA: SAGE.
- Liang, K.Y. and S.L. Zeger (1986) 'Longitudinal Data Analysis Using Generalized Linear Models', *Biometrika* 73(1): 13–22.
- McGraw, S.A., S.M. McKinlay, L. McClements, T.M. Lasater, A. Assaf and R.A. Carleton (1989) 'Methods in Program Evaluation: The Process Evaluation System of the Pawtucket Heart Health Program', *Evaluation Review* 13(5): 459–83.
- Oakley, A. (1992) *Social Support and Motherhood: The Natural History of a Research Project*. Oxford: Basil Blackwell.
- Oakley, A. (2000) *Experiments in Knowing: Gender and Method in the Social Sciences*. Cambridge: Polity Press.
- Oakley, A., V. Strange, J. Stephenson, S. Forrest, H. Monteiro and the RIPPLE Study Team (2004) 'Evaluating Processes: A Case Study of a Randomized Controlled Trial of Sex Education', *Evaluation* 10(4): 440–62.
- Oxman, A.D., H. Gordon and A. Guyatt (1992) 'Consumer's Guide to Subgroup Analyses', *Annals of Internal Medicine* 116: 78–84.
- Patton, M.Q. (1989) 'A Context and Boundaries for a Theory-Driven Approach to Validity', *Evaluation and Program Planning* 4: 375–7.
- Pawson, R. and N. Tilley (1997) *Realistic Evaluation*. London: SAGE.
- Pawson, R., T. Greenhalgh, G. Harvey and K. Walshe (2004) *Realist Synthesis: An Introduction*. ESRC Research Methods Programme, RMP Methods Paper 2/2004. Manchester: University of Manchester.
- Pocock, S.J. and M. Abdalla (1998) 'The Hope and Hazards of Using Compliance Data in Randomized Controlled Trials', *Statistics in Medicine* 17: 303–17.
- Pocock, S.J., M.D. Hughes and R.J. Lee (1987) 'Statistical Problems in the Reporting of Clinical Trials: A Survey of Three Medical Journals', *New England Journal of Medicine* 317: 426–32.
- Rossi, P.H., M.W. Lipsey and H.E. Freeman (2004) *Evaluation: A Systematic Approach*, 7th edn. Thousand Oaks, CA: SAGE.
- Sommer, A. and S. Zeger (1991) 'On Estimating Efficacy from Clinical Trials', *Statistics in Medicine* 10: 45–52.
- Stame, N. (2004) 'Theory-Based Evaluation and Varieties of Complexity', *Evaluation* 10(1): 58–76.
- Stephenson, J.M., A. Oakley, A.M. Johnson, S. Forrest, V. Strange, S. Charleston, S. Black, A. Copas, A. Petrukevitch and A. Babiker (2003) 'A School-Based Randomized Controlled Trial of Peer-Led Sex Education in England', *Controlled Clinical Trials* 24: 643–57.
- Stephenson, J.M., V. Strange, S. Forrest, A. Oakley, A. Copas, E. Allen, A. Babiker, S. Black, M. Ali, H. Monteiro and A.M. Johnson (2004) 'A Randomized Intervention Trial of PuPil-Led Sex Education in England (RIPPLE)', *The Lancet* 36: 338–46.
- Strange, V., S. Forrest, A. Oakley and the RIPPLE Study Team (2002) 'What Influences Peer-led Sex Education in the Classroom? A View from the Peer Educators', *Health Education Research: Theory and Practice* 17(3): 327–38.
- Tones, K. and S. Tilford (1994) *Health Education: Effectiveness, Efficiency and Equity*. London: Chapman Hall.
- Victora, C.G., S.R. Huttly, S.C. Fuchs and M.T.A. Olinto (1997) 'The Role of Conceptual Frameworks in Epidemiological Analysis: A Hierarchical Approach', *International Journal of Epidemiology* 26(1): 224–7.
- Wiggins, M., A. Oakley, I. Roberts, H. Turner, L. Rajan, H. Austerberry, R. Mujica and M. Mugford (2004) 'The Social Support and Family Health Study: A Randomised

Evaluation 12(3)

- Controlled Trial and Economic Evaluation of Two Alternative Forms of Postnatal Support for Mothers Living in Disadvantaged Inner City Areas', *Health Technology Assessment Monograph* 8(32).
- Wight, D. and A. Obasi (2003) 'Unpacking the Black Box: The Importance of Process Data to Explain Outcomes', in J.M. Stephenson, J. Imrie and C. Bonell (eds) *Effective Sexual Health Interventions*. Oxford: Oxford University Press.
- Yeaton, W.H. (1985) 'Using Measures of Treatment Strength and Integrity in Planning Research', in D.A. Cordray (ed.) *Utilizing Prior Research in Evaluation Planning*. San Francisco, CA: Jossey-Bass.
- Yusuf, S., R. Collins and R. Peto (1984) 'Why Do We Need Some Large, Simple, Randomized Trials?', *Statistics in Medicine* 3: 409–20.

VICKI STRANGE is a Research Officer at the Social Science Research Unit, University of London Institute of Education. Please address correspondence to: Social Science Research Unit, Institute of Education, 18 Woburn Square, London WC1H 0NR, UK.
[email: v.strange@ioe.ac.uk]

ELIZABETH ALLEN is a Research Fellow in the Department of Sexually Transmitted Diseases at the Royal Free and University College London Medical School.
[email: eallen@gum.ucl.ac.uk]

ANN OAKLEY is Professor of Sociology and Social Policy and Founding Director of the Social Science Research Unit at the University of London Institute of Education.
[email: a.oakley@ioe.ac.uk]

CHRIS BONELL is a Senior Lecturer in Sociology and Epidemiology at the London School of Hygiene and Tropical Medicine. [email: chris.bonell@lshtm.ac.uk]

ANNE JOHNSON is Professor of Infectious Diseases and Head of Primary Care and Population Sciences at the Royal Free and University College London Medical School.
[email: a.johnson@pcps.ucl.ac.uk]

JUDITH STEPHENSON is Head of the Epidemiology and Social Science Group in the Department of Sexually Transmitted Diseases at the Royal Free and University College London Medical School. [email: jstephen@gum.ucl.ac.uk]