The benefits of combined (mixed) methods research: The large scale introduction of parenting programmes

Geoff Lindsay

Abstract: In this paper I argue for the importance of combined (mixed) methods studies when evaluating complex initiatives such as interventions. These studies combine quantitative and qualitative methods purposively, in order to capture a broader range of important variables than can be examined by either quantitative or qualitative methods alone. In essence, a combined methods study can address outcomes and outputs, and also provide evidence on the processes that are important to optimising success. A large scale Government funded initiative is used as an example. The Parenting Early Intervention Pathfinder introduced parenting programmes into 18 local authorities. Quantitative data showed the scale of the support, and the significant improvements made in parenting skills and mental well-being, and in child behaviour. Qualitative data explored the factors optimising successful implementation. The evidence from the study was used to inform the decision to roll out evidence-based parenting programmes across the whole of England.

Keywords: combined (mixed) methods research; parenting programmes; evidence based interventions

1. Professor and Director, CEDAR (Centre for Educational Development, Appraisal and Research, University of Warwick

Address for correspondence: Professor Geoff Lindsay, CEDAR, University of Warwick, Coventry CV4 7AL. Geoff.lindsay@warwick.ac.uk
Introduction

In this paper I shall argue for the importance of combined (mixed) methods designs for the evaluation of complex studies. I shall focus on large scale interventions, using the example of the introduction of parenting programmes. Combined methods designs are applicable to many research topics but are particularly useful for large scale initiatives as the complexity cannot be adequately addressed by more simple designs. For example, a quantitative study may produce high quality data that indicate important findings, but the reasons for the results will not have been examined. In terms of actions that follow from research we may need to have answers to the ‘why’ and ‘how’ questions as well. This is particularly true in public policy where implementation of an initiative that has been shown to be effective in rigorous, but typically small scale research may not work so effectively when applied in a new setting, especially when scaled up. This may be a function of associated implementation factors rather than of the substantive intervention.

The focus of this paper, therefore, is on the use of a combined methods research design to investigate the introduction of parenting programmes as a means of reducing behavioural difficulties in children. This example draws upon a study undertaken for the United Kingdom Government’s Department for Children, Schools and Families (DCSF) – see Lindsay, Strand & Davis, 2011, and Lindsay et al., 2008) now the DfE. The project was on a large scale over three years and produced findings which led the government to fund a roll out across all local authorities (LAs) in England.

Combined methods research: An example

A fundamental question to ask about any intervention is, does it work? This may appear obvious but it is the case that many interventions are in use in social work, education and speech and language therapy, for example, which have evidence that it at best limited and at worse non-existent (Glogowska, 2011); see, for example, a recent study of the interventions used by speech and language therapists and the level of evidence available for their effectiveness (Law, Lee, Roulstone, Zeng, & Lindsay, 2012). This is also the case with parenting programmes.

Programmes designed to improve parenting skills have become very popular as a means of helping parents of children with behavioural difficulties. There are now many such programmes available in the UK, US and many other countries. They are popular because evidence indicates that children exhibiting behavioural difficulties are at greater risk during adolescence and when they become adults of engaging in crime, having mental health problems and parenting difficulties (Patterson, Mockford, Barlow, Pyper & Stewart-Brown, 2002). Furthermore, the scale of behavioural difficulties is substantial with rates of 6 per cent among 5-16
year olds in the UK (Green, McGinnity, Meltzer, Ford, & Goodman, 2004).

The main, and indeed generally accepted ‘gold standard’ to examining the effectiveness of interventions is the randomized controlled trial (RCT). This is a carefully, rigorously designed and executed study which explores the effects of the intervention in one group of participants compared with a non-intervention control (comparison) group of participants. The non-intervention group may receive a placebo (in a drug trial), be placed on a waiting list before receiving the intervention later, or receive an alternative intervention, for example one that is known from previous research to be effective. Participants are randomly allocated to one or other group (intervention or control) as a means of avoiding bias in favour of one or other group.

There are now a number of parenting programmes that have one or more RCTs that show their efficiency; some now have large numbers (for example, Nowack & Henrichs, 2008). These studies have been reviewed by organisations to advise which parenting programmes can be considered ‘evidence based’, for example the United Nations Organisation for Drugs and Crime (UNODC, 2010) and in the UK the National Academy for Parenting Research (http://www.kcl.ac.uk/iop/depts/cap/research/napr/parentingprogrammesevaluationtool.aspx – see also O’Mara et al, 2011). Such evidence has been used in recent reports to argue for the importance of a sound evidence base, for example, the Allen report on early intervention on behalf of the UK government (Allen, 2011) and nongovernmental agencies such as Save the Children (Lindsay, Cullen & Wellings, 2011).

Randomized controlled trials allow the analysis of quantitative data: in the case of parenting programmes they typically include measures of parenting and child behaviour. Data are statistically analysed in order to identify changes (improvements) over time in the key variables used in the study. It is less common to examine process variables, that is factors concerning how the implementation was carried out. This is also an important area of research. The RCT(s) may demonstrate a parenting programme has efficacy in the particular context of that, typically small scale study. But to implement that programme as part of a public policy initiative it is also important to explore whether it works on a larger scale. This will require larger scale studies in community settings (effectiveness trials) which will also include standardised studies of quantitative data concerning improvements in before and after programme scores on key measures, as detailed above (Society for Prevention Research, www.preventionresearch.org). However, scaling up brings new challenges and qualitative methods have an important role to play here. Other measures may include parental mental health, for example depression or mental well-being.

The quantitative data provide information about whether the intervention works and, if so, the size of the effects, but not how or why, or how the implementation may be optimised. The latter questions go beyond the specific element of the intervention, in the present case the parenting programme(s). In order to investigate these questions it is important to examine the delivery of the programme(s) in a broader sense, for
example: the recruitment processes, efficiency and cost effectiveness, participants’ perspectives, and improvement ideas. In summary, the quantitative strand is used to demonstrate effects and the qualitative strand explores the reasons for these effects and how they may be maintained, or even improved.

There are many different designs within combined methods research. Leech and Onwuegbuzie (2009) have produced a typology of possible designs comprising three dimensions. The first is the combination (or in their terms mixing) dimension: partially or fully combined methods. The second dimension is time: concurrent or sequential. The third dimension concerns emphasis, whether qualitative and quantitative methods have equal status or one has a dominant status.

In this paper I shall describe the use of both quantitative and qualitative methods in a study of parenting programmes. I refer to this as a combined methods design, although the term mixed methods is more often used. I prefer ‘combined methods’ as a key characteristic is that these methods are carefully and deliberately chosen to be planned and complementary; to my mind, the term ‘mixed’ does not capture this deliberate, analytically driven approach; rather there is a danger of appearing causal and less rigorous.

A combined methods study of parenting programmes

The Parenting Early Intervention Pathfinder (PEIP) was funded by the UK government’s Department for Children, Schools and Families (DCSF) now the Department for Education (DfE). Its aim was to examine the effectiveness of three parenting programmes when rolled out on a large scale for parents of children aged 8-13 years exhibiting or at risk of behavioural difficulties. The initiative was driven by concerns about the level of behavioural difficulties in young children and their links to later crime and other adverse outcomes, as described above (Respect Taskforce, 2006). Initially a review of existing evidence was commissioned which identified programmes with a good evidence base (Moran, Ghate & van der Werwe, 2004).

The DCSF selected 18 local authorities (LAs) in which to fund one of three parenting programmes (6 per LA) selected on the basis of the review by Moran et al. (2004), namely: Triple P (Sanders, 1999; website), Incredible Years (Webster-Stratton & Reid, 2003; website) and Strengthening Families Strengthening Communities (SFSC: Steele, Marigna, Tello, & Johnston, 2000). Of these, Triple P and Incredible Years had evidence from RCTs but the evidence for SFSC did not include an RCT. We were commissioned to carry out a study of the effectiveness of the PEIP over its implementation during 2006-08.
Design

The PEIP was based on the already existing evidence that the three programmes had *efficacy*. The purpose of the PEIP was to examine *effectiveness* of the programmes when implemented across 18 LAs in community settings, on a large scale. The study was designed to explore improvements in parenting and child behaviours and also factors that aided or limited improvements. The study therefore examined outputs, outcomes and processes. The design was a *parallel* quantitative and qualitative combined methods study: data from both strands were collected throughout the study in parallel.

Quantitative strand

Demographic information on the parents attending the parenting programme and their child about whom they had most concerns (within the 8-13 target age range) were collected at the first session, together with parents’ responses to four questionnaires: the Warwick Edinburgh Mental Well-being Scale (Tennant, Fishwick, Platt, Joseph, & Stewart-Brown, 2000); the Parenting Scale (Irvine, Biglan, Smolkowski, & Ary, 1999) which measures parents’ perceptions of their over-reactivity and laxness; and *Being a Parent* (Johnston & Mash, 1989) which measures parents’ perceptions of their efficacy and sense of satisfaction with being a parent. Parents also completed the Strengths and Difficulties Questionnaire (SDQ: Goodman, 1997; [http://www.sdqinfo.com](http://www.sdqinfo.com)) to provide an assessment of the behaviour of their child about whom they had most concern.

Working with the 18 LAs we arranged for these questionnaires (presented in a booklet) to be completed at the first session or prior to the first session (pre-test). The second booklet containing the same questionnaires was then sent out in time for the final session of the programme post-rest: this time period varied as the programmes had different lengths. These booklets also included a questionnaire *How was your group?* which replaced the demographic questionnaire. We collected these data throughout the project as the groups were held. We analysed data and produced interim findings and provided initial feedback to the DCSF and also to the LAs and general researcher/practitioner community (Lindsay, et al., 2007a, b).

Qualitative strand

*Interviews*

We undertook four rounds of interviews to capture the setting up of the PEIP; the initial period of implementation; engagement of PEIP with schools and extended
schools; and the final period when PEIP was well established. We interviewed key practitioners over these four phases: each LA's strategic lead (total of 74 interviews), operational lead (39), and samples of group facilitators (205) and head teachers (24); we also conducted 81 parent interviews. These interviews were specifically designed for each participant group but all examined the operation and experience of the PEIP. For example, interviews with the strategic leads focused on the place of parenting programmes in the LA's strategy for parenting support and the management of the parenting programmes, including recruitment. Parent interviews explored their views on the content and delivery of their programme, their experience of their group, whether they considered they had benefitted, and whether this had led to their child's behaviour improving.

Document analysis
We also examined each LA's proposal to the DCSF describing how they planned to implement the PEIP, including how it fitted into their developing parenting strategy, and the parenting policies as these developed.

Cost effectiveness
We also used the data from each LA on numbers of parents that had started their programme together with DCSF data on the grant received by each LA, to examine the cost effectiveness of each LA's implementation of PEIP over 2006-08.

Results
The results of the PEIP have been published in a series of journal articles (Cullen et al., 2012; Lindsay, Strand and Davis, 2011) and in the original reports to the DCSF. The focus of this paper is to use the results to indicate how the combined methods design allowed complementary findings to provide a comprehensive overview of the outcomes of PEIP and the factors that assisted or limited LAs' effective implementation.

Quantitative analysis
The qualitative data showed that the PEIP was effective on all measures. First, the PEIP was delivered to the target population. The parents came from a wide range of ethnic backgrounds and parents as a whole were more likely to be socially disadvantaged than the population as a whole. For example, 47% had left school at 16 years or younger and had high levels of need: 59% had seen a general practitioner in the previous six months and had low levels of mental well-being before they started their programme: 76% were below the national mean.
Second, the PEIP targeted parents who had a child with significant behavioural, emotional and social difficulties compared with the general population. The prevalence of conduct problems was about six times greater than expectation with respect to children who had statements of special educational needs, and those receiving additional support in school.

Third, the PEIP delivered support to a large number of parents: 3375 parents began one of 425 parenting classes across the 18 LAs. We collected data on 2207 parents as they started a course. However, we only received post-course booklets for 51% of those parents. This raised two important interrelated points regarding large scale studies like this. A response rate of 51% may be interpreted as a drop-out rate of 49%, but this did not tell the whole story. From our monitoring data, as well as informal discussions with LA staff, we found that there were various administrative errors, for example, some groups did not receive their post-course booklets, others returned completed booklets to the LA to post to the research team, but these were not posted back to us. Approximately half the missing post-group data were due to administrative errors at the LA. This left about a quarter of parents (27%) who did not complete their programme: true drop outs. This compares well with data from smaller scale studies, especially those in community settings.

Fourth, the PEIP funded about 1100 new facilitators who were each able to implement one of the three parenting programmes. These represent a major LA investment in the delivery of parenting programmes during the PEIP and for sustainability after Government funding ended.

Fifth, there were substantial improvements on all measures, but particularly the parenting measures. We used the statistical measure of effect size as small improvements may be statistically significant when sample size is large (effect size (Cohen’s $d$) greater than 0.8). Reductions in parents’ over-reactivity to their children showed a large effect size (0.8) and both parents’ mental well-being and parental laxness were close to showing a large effect size (0.71): in each case: see Table 1, overleaf. Improvements in sense of efficacy as a parent and satisfaction with being a parent also improved, although with lower (medium) effect sizes.

There were also improvements in parents’ ratings of their child’s behaviour, but with smaller effect sizes than for the parent measures (Table 2). However, improvements in conduct problems and total difficulties showed medium effect sizes.

Sixth, we found that all three programmes were successful, in terms of our parent and child measures, although there were some differences between them (See Lindsay, Strand, Cullen et al., 2011 for details). However, we also found that the costs of the three programmes varied greatly, so indicating differential cost effectiveness. This was largely due to the different lengths of each programme, but other factors including group size were relevant factors.
Qualitative analysis

Our qualitative strand provided important information about the processes of implementing PEIP. We examined two main issues: the nature of the implementation process and the reasons for the actions taken; and the views of parents and professionals about the PEIP implementation.

We identified the importance of the LAs’ strategic and organisational approaches. The outcomes of the PEIP were ultimately dependent on the quality of the programmes. However, the delivery of the PEIP also required that LAs be efficient and cost effective. The quantitative data showed that there was substantial variation in the number of groups run and therefore of parents who could be supported, from...
42 (1.9% of the total parents) to 366 (16.6% of all parents). We therefore explored the reasons for this. Effectiveness and cost effectiveness were dependent upon both LA efficiency and the programmes as Triple P was shorter and cheaper to implement, and LAs could provide courses to more parents at lower cost per parent: the numbers of courses totalled 185 by Triple P, 83 by SFSC and 70 by Incredible Years.

However, this did not tell the whole story. We also explored important factors known to be successful in delivering parenting programmes. One of these is fidelity, the extent to which delivery stays close to that which was examined in the RCTs and as set out in the manual. The key factors here are the manual itself, initial training of facilitators by the programme staff and the supervision received by facilitators once they started implement the programme. These programmes all had manuals and all facilitators had been trained by the programme staff, that is trainers themselves trained to deliver the programme, optimising the likelihood of fidelity compared with non-manualised programmes and where non-programme staff are used as trainers. We also explored the nature of supervision offered by LAs. Here we found variations across LAs as supervision, as conceptualised in the helping professions, was often interpreted as line management. This important distinction was one of the points we stressed in our reports and later guidance.

We explored the nature and reasons for variations in the set-up period (phase 1) when LAs differed in getting their systems in place and fully operational. This was a key period and it was clear that LAs varied greatly, for different reasons. For example, time taken to appoint senior staff and have them in place varied, with direct knock on effects on the implementation.

We explored why fathers were so rarely engaged: only 12% of all parents were male (Cullen et al. 2011). We also explored the LAs’ approaches to creating a parenting policy and the role of parenting programmes within it. Our interviews revealed the importance of the developing parenting strategy in longer-term sustainability of this support. We found that the PEIP had, reciprocally, informed the parenting policy. Its success reinforced the inclusion of evidence based parenting programmes.

We examined the role of schools and extended schools and services in the delivery of the parenting classes (Lindsay et al., 2007a). The latter were being developed at that time so the DCSF commissioned this as additional research, as LAs and schools developed their core offers for parents. Extended schools had wide responsibilities and parenting support was welcomed as part of the new role.

We also explored the views of parents regarding their experiences of their parenting programme, both the benefits that occurred as well as the challenges they had had in attending. Our quantitative data from parents showed very positive views about attending the programme but the interviews enabled these to be explored in more detail. For example, we explored the parents’ views of the content of the programmes. Two programmes had been developed in the US and one in Australia so parents’ perceptions of match and relevance were important, both the aims and general approach of the programme and the specific content of the materials. In fact parents
had no difficulties adjusting to American and Australian examples and were positive about programme content and implementation, including the skills and sensitivity of the facilitators.

We also explored with parents how and why they had been recruited, the changes that had occurred as a result of attending, with respect to both themselves and their children, and the aspects of their attending the programme that they judged to be relevant. These parents were generally very positive about their experiences and the beneficial effects that the change in their parenting styles had had on their relationships with their children. They saw a direct relationship with the children’s behavioural improvement.

Conclusions

This paper has presented an example of a large scale research project that used a combined methods design. There are many possible variants of designs within combined methods approach (Leech & Onwuegbuzie, 2009). Our design was selected as appropriate to the specific study: a DCSF-funded initiative to develop evidence-based parenting programmes in 18 LAs. We collected substantial quantitative data from parents attending 425 courses; interviews with LA staff, parents, group facilitators and head teachers; and from analysis of the programmes themselves and policy documents. The combined methods design produced findings that were complementary. In brief, the quantitative data showed that the PEIP had been successful in delivering parenting programmes to a large number of parents; and that these parents had improved their parenting skills and their children’s behaviour, as perceived by the parent, had also improved. The qualitative strand provided information concerning how and why the PEIP had been successful; it also identified limitations and room for improvement, for example, the wide range of performance by the LAs and the consequent impact on costs.

Combined (mixed) methods research is not only appropriate but also, I would argue, essential in order to investigate the effectiveness of implementations thoroughly. The two major strands, quantitative and qualitative, must be planned to provide complementary evidence, addressing different research questions. If this is achieved then combined methods studies can provide rich data. In this case the study’s evidence was used by the DCSF when it decided to extend the funding of evidence based programmes to all English LAs, an initiative that was itself evaluated (Lindsay, Strand, Cullen et al., 2011).
References

Leech, N.L. and Onwuegbuzie (2009) A typology of mixed methods research designs. Qualitative and Quantitative, 43, 265-275


