National College for School Leadership

www.ncsl.org.uk

Inspiring leaders;
improving children's lives

3

Section 3 **Evaluation methodology**

Section 3 **Evaluation methodology** 15

3.1 Introduction

In this chapter, we present our evaluation methodology and the development of data collection instrumentation and sampling.

This project required the development of an evaluation strategy, because the end results of the BDP are to be evaluated in terms of their impact on individual SBMs in their schools. This presents the evaluator with a number of issues, arising from:

- the various roles of NCSL as programme leader
- multiple contracted course providers
- multiple course material reviewers
- multiple facilitators for the programme's blended-learning element
- a third-party evaluation over which providers and leaders have influence but no control

This chapter explains how we went about this task. It is divided into four sections:

- the working framework, that of realistic evaluation
- evaluation instrumentation and data-gathering techniques
- sampling approaches and the extent of the evidence base
- analysis methods and software packages used

3.2 Realistic evaluation

Choice of methodology for any evaluation project presents an important issue for the evaluator. The evaluation strategy must be sensitive to variations in conditions and contexts; what might work in some situations would not work in others. Working on a similarly complex subject (policing and criminology), Pawson and Tilley developed an approach that they termed 'realistic evaluation' (Pawson & Tilley, 1997).

We adopted realistic evaluation because we were evaluating a professional development programme designed to impact not only on participants, both personally and professionally, but through them on their schools. For a more detailed look at the appropriateness of this approach, see appendix 1.

Central to realistic evaluation is the question of what works for whom and in what circumstances. History and contextuality are seen as key influences on programme outputs and effects. In this respect, realistic evaluation differs from traditional experimental designs where such sources of variation have to be controlled.

Within a realistic-evaluation framework, the key to understanding how a programme works lies in identifying its context outcome mechanism configuration. Specific impacts, outputs and outcomes arise only in specific circumstances. Learning more about these circumstances, and the causal influences within them, may lead to useful knowledge for the fine-tuning and scaling-up of programme initiatives.

3.3 Evaluation instrumentation

Evaluation of a programme requires the collection of data and evidence upon which to base judgements about how the programme is proceeding, its outcomes and its impacts. Our brief to evaluate impact has required the development of a number of data-gathering instruments, both quantitative and qualitative. Given the number of course participants, survey instrumentation has taken the form of self-report questionnaires for:

- residential session satisfaction surveys
- post-course impact surveys
- pre-course surveys
- assessment of knowledge and confidence

In order to gain a fuller understanding of reactions in specific situations we also used:

- interviews
- professional observations

Table 3.1 summarises the data collection tools.

Table 3.1: Evaluation instrumentation

Tool	Nature of tool	Data gathered	What the data tells us
Residential session satisfaction survey	Short questionnaire completed by all candidates after each residential session	Numerical data indicating satisfaction with materials, tutoring and other aspects of the session	How candidates felt about materials, tutoring and other aspects of the session
Post-course impact survey	Detailed questionnaire completed by a sample of graduates after course completion	Statistical evidence on course impacts on candidates and their schools	How courses are impacting upon participants and their schools
Pre-course survey	Questionnaire to establish candidates' knowledge and understanding before starting a course	Numerical data indicating candidates' perceptions of their knowledge and understanding	Information for pre-course baseline on candidates' knowledge and understanding
Assessment of knowledge and confidence (KnoCon)	Questionnaire to assess candidates' knowledge and confidence on a range of key topics	Statistical data on candidates' perceived knowledge and confidence in key CSBM areas	Match between knowledge and confidence across 26 items; extent to which course is meeting candidates' needs
Interviews	Professional conversations based on a broad schedule of areas for exploration	Transcripts of candidates' comments and post-course understandings	In-depth knowledge of candidates' perceptions of the course and its impact on their schools
Professional observations	Observation of teaching sessions; establishment of contacts for future work	Contextual data, in form of evaluators' notes, on candidates' views about sessions and tutors	How materials, teaching, work in schools, and talk2learn are appreciated

The following sections provide more detailed descriptions of the instrumentation we have employed.

3.3.1 Residential session satisfaction survey (CSBM)

A key component of course delivery is face-to-face teaching in a residential format. We inherited a basic satisfaction survey (appendix 2) which all candidates are invited to complete after each of the course's three two-day residential sessions. These surveys have been identical for each residential session.

Data gathered

This instrument gathered candidate responses to questions about venue, teaching materials, tutor support and pre-course administration.

Objective

The data enabled the BDP team to know how well received the course materials, session and support were, and has provided a baseline for candidate reactions. We have suggested and devised instruments specific to each of the course's three residential sessions; at the time of reporting these await approval.

3.3.2 Residential session satisfaction surveys (DSBM)

After consultation with the BDP team, we have taken a more focused approach to satisfaction surveys for the DSBM course.

Data gathered

There are three survey instruments (see appendix 8), each specific to one of the course's three sessions. Our intention was to enable key aspects of each session to be scrutinised. Content was drawn from each session's objectives and accompanying materials.

Objective

This data provides a more detailed understanding of how each residential session has progressed, and permits a clearer identification of problematic issues.

3.3.3 Post-course impact survey

As part of our evaluation we devised two post-course impact surveys. The first was used in 2005 with a sample of CSBM graduates. The second, modified from a 2005 pilot course, was first used in 2006 with DSBM graduates. These surveys provided a means of assessing impact across a large sample. In view of the illuminating data gathered by the 2005 CSBM survey, we felt it important to develop something similar for DSBM. NCSL subsequently took the view that all candidates should, as a course completion requirement, complete an end-of-course survey. Additionally it took the view that all candidates should be surveyed in a similar fashion before they joined either course.

Data gathered

These instruments sought to collect demographic data about SBMs and their schools, thus providing some insight into the context. This included information about their phase of schooling, time in post, membership of the senior management team (SMT; also referred to as senior leadership team or SLT), highest educational qualification, salary bracket and the number of support staff they managed.

Objective

This contextual data enabled us to identify certain issues within SBM groups and also to confound some popular misconceptions.

We were particularly interested in SBMs' estimates of how they spent their work time. This was subsequently used in assessing their perceptions of changes resulting from their participation in the BDP. Using ideas similar to those in the baseline study by the University of Lincoln (2004), we asked SBMs to classify their time over a week in terms of administration, management and leadership. Further information was then sought about how they perceived different aspects of course provision, whether they felt they had learnt new things and whether they felt this was having some impact in their schools. As a key part of a school business manager's work is of a financial nature, we also asked about the school budget and whether they had made savings and brought in additional money for their school.

3.3.4 Pre-course survey

Starting from CSBM cohort 6 (2005) and DSBM cohort 3 (2005) , we introduced a survey to gather data about SBMs before they commence either course.

Data gathered

These instruments gather data about SBMs' prior knowledge and experience.

Objective

This data will permit comparisons with exit data, and give the BDP team a clearer idea of the initial impact of the courses.

3.3.5 Assessment of SBMs' knowledge and confidence

As an early measure of how appropriate candidates were finding CSBM materials and content, we surveyed a sample from cohort 2 in 2004.

Data gathered

Here we were interested in how candidates rated themselves on 26 items taken from aspects of the course content. For each item they were asked to rate their knowledge of the topic or issue and their level of confidence in dealing with it. Following Cains and Brown (1996; 1998a; 1988b), we developed a knowledge and confidence questionnaire, referred to as 'KnoCon', with replies sought on parallel six-point scales (for details, see chapter 4 and appendix 4).

Objective

Comparisons of mean scores for each of the 26 items allowed us to identify:

- areas where SBMs felt knowledgeable and those where they felt less secure
- particular topics they felt more confident or less confident about

In this way it would be possible to identify areas where the course might need to be strengthened.

3.3.6 Professional observations

Face-to-face teaching is a key element in the delivery of both courses. For CSBM cohort 2, an evaluator attended one full residential session in each of the nine regions (course locations map onto Government Office regions). Both evaluators attended every DSBM pilot session except the first (in pilot cohort 1a). For cohort 1 of the nationally rolled-out DSBM course, an evaluator attended at least one residential session in each region.

Our attendance at these sessions was very important for a number of reasons, noted below.

Course delivery

We needed to see the delivery of the course as it actually happened, and to observe delivery by various tutors working for three different providers. Our brief did not involve quality assurance per se, because each provider had its own internal arrangements. We felt it important, however, that sessions for a national programme such as the BDP be observed across the range of providers. Our observations were primarily related to how the candidates found the experience, how they reacted to the materials, the tutoring and the online elements, and their preparation for assessment.

Tutor contact

We wanted an opportunity to talk with tutors about teaching in a blended-learning context, and about issues in the delivery of the residential sessions. We needed to hear at first hand how tutors felt SBMs were handling the courses. For many SBMs this was their first excursion into professional development.

SBM contact

We needed to make personal contact with SBMs to facilitate follow-up work. To that end we viewed our involvement at residential sessions in terms of participant observation. We participated in sessions and worked with SBMs as a way of listening to them and building relationships which would enable us to make contact later and gain access to them in their schools.

3.3.7 Interviews with graduates and headteachers

As a result of our approach to the residential sessions, we subsequently interviewed samples of CSBM and DSBM graduates in their own schools and, where possible, also interviewed their headteachers. These interviews took the form of professional conversations.

Interviews were undertaken within the framework of an interview schedule, but so as not to constrain what the SBM or headteacher wanted to tell us. In research terms these could be loosely described as semi-structured interviews. The approach meant that the number of questions and prompts which might normally be needed in an interview format was considerably reduced. We were particularly anxious not to put words into the mouths of respondents, preferring to let them tell us about their jobs in relation to the BDP course. This has meant that some topics, which we might have expected respondents to talk about, were sometimes absent from the conversation, because some respondents clearly didn't think them important enough to mention. In some cases, for example, it meant meeting with a respondent over lunch as that was the only time available.

Where possible, the interviews were recorded and the contents transcribed for analysis. Transcripts were cleaned up before analysis. We have used a *grounded* approach to analysis, determining data codes on the basis of themes emerging from what respondents said, rather than using a hard and fast pre-existing schedule of codes.

3.4 Overview of evidence base

As set out in section 3.3, we have sampled SBMs for self-report questionnaires.

We attended residential sessions in at least one centre in each region, giving a distribution across the country. No work was specifically limited to one geographical area and sessions were attended across England. Each residential session was attended by one of the evaluation team, except for those in the DSBM pilot, where both evaluators were present.

Data for the CSBM impact survey was gathered by telephone in 2005. Candidates were contacted using details supplied by NCSL. There was an element of opportunity sampling involved with this group as our telephone interviewer was only able to speak with those SBMs who happened to be available at the time she called. In total 283 SBMs responded to this part of the data collection.

Data for the DSBM impact survey was compiled from electronic responses which candidates left as part of NCSL's programme exit requirements. This data had been gathered by NCSL and was forwarded to the evaluation team for analysis. This sample too was opportunistic in that we only had data from those candidates who were prepared to complete the form. We received 118 valid responses for this.

Candidates from CSBM cohort 2 who completed the KnoCon questionnaire also formed an opportunity sample: those SBMs prepared to devote time to the completion of the instrument and return it to us via NCSL. A total of 282 SBMs completed this questionnaire.

We visited and interviewed 33 SBMs and 10 headteachers in their schools. This data must also be seen as essentially opportunistic in that we could only interview SBMs who were prepared to meet us. It was even more difficult to meet headteachers as they were often not on the premises at a time when it was convenient to interview their SBMs.

Overall, we received and analysed data from 715 SBMs, 10 headteachers, 40 tutors and the managers of each of the 3 contractor-providers and of the NCSL project team. In addition, through the residential sessions we met with over 450 SBMs.

We undertook special visits to evaluate other pilot activities: an environmental module for CSBM, a local provider initiative (Chesil), a pilot external candidate workshop, a DSBM pilot debriefing conference, a two-day briefing event prior to national roll-out of the DSBM, a tutor training and updating conference and at least two graduation events.

Finally, in line with our commitment to a 'no surprises' evaluation and to enable the project team to respond to emerging findings, we met regularly – usually not less than once every six weeks – with the BDP team at NCSL.

3.5 Note on analysis methods and software

Quantitative data from questionnaires was analysed using SPSS, the principal statistical package used in higher education for such analysis.

Qualitative data from interviews was cleaned, anonymised and converted to text-only format for entry into Atlas.ti, a software package for analysis of qualitative data.

3.6 Summary

In this chapter, we have provided an overview of the evaluation framework, namely realistic evaluation, employed for this project. We have provided a synopsis of the development and use of a range of data-gathering instruments, both quantitative and qualitative. Finally, we have provided an account of the nature, geographical spread and size of the evidence base which underpins the findings reported in the remainder of this document.