This article presents a summary review of the design and results of an independently conducted evaluation of a national New Zealand (NZ) Ministry of Education funded contract for professional development of staff of students with special education needs in 49 schools.

The evaluation was conducted as a mixed method design (Johnson & Onwuegbuzie 2004) in the following three phases: a broad questionnaire issued to all participants associated with the development; eight focus groups with a sample of participants; and success case studies (Brinkerhoff, 2003) with four schools.

The most significant overall finding was that regardless of the approach to development engaged in (either action research, AR, or action learning, AL), there was an outstanding recurring characteristic of staff and supporters wanting to see the students excel. Other key participant self-report impacts from the small-scale projects on adaptation of the curriculum fell under the headings of improved social interaction and academic achievement for students, changes in values and attitudes (for students and teachers), and changes in teaching practice. Participants referred to the importance of school context factors (inclusive planning, management support) and internal and external experts as enablers towards the effectiveness of development. Barriers to effectiveness were noted as associated with initial national contract administration, the number of development initiatives involved in and lack of alignment between these varied initiatives. Maintaining and sustaining the effective impact of projects was seen as dependent on: ongoing commitment and follow-through by school management, governors and program teachers; having ongoing funding and support (internal and external); and bringing other staff on board.

The Phase Three success case evaluation revealed an important element that distinguished projects perceived to be highly successful by both the participants and Ministry of Education personnel. In this small proportion of projects the participant action researchers/learners utilised ‘informed’ decision-making. Although many participants in Phases One and Two justified their limited use of informed decision-making by noting that it was either too early to validate project outcome changes, or it was difficult to show causal effect (changes could be attributed to the development program), a hallmark of the four success cases was the use of strong data in the reconnaissance and evaluation phases of the AR and AL and improvement initiatives that were informed by both this data and relevant previous literature.
**Introduction**

This article initially outlines the background and purpose associated with the evaluation contract for the teacher development program, including the types of schools involved, the professional development providers, the aims of the evaluation and the key questions associated with those aims. The methodological considerations are outlined next, that is, the mixed method (triangulation) design of the evaluation, the data collection methods, the sampling processes, and ethical considerations. A brief discussion of the results of each phase follows. Finally, overall conclusions and limitations are noted.

**Background and purpose**

The program was designed to:

- develop teacher knowledge and share ideas on how to support learners who require significant adaptation to the curriculum content ... one of the strategies focused on recognising the diversity of all learners, the contribution they make to the cultural and educational enrichment of schools and communities, and the different contexts in which learning occurs. (Ministry of Education 2005a, p. 3)

The development program was conducted over 14 months with 49 diverse schools throughout NZ using either an AR (25 schools) or AL (24 schools) approach. The AR schools were led by a university researcher; the AL schools by nominated Ministry of Education Group Special Education (GSE) district facilitators.

The development program was expected to impact on student learning, social outcomes and cultural identity/connectedness. Sustainable change was the goal. The AR and AL approaches were seen as vehicles for teachers to examine and critique their own practice in a systematic, intentional way. Ultimately the approach focused on teacher inquiry—teachers taking a close critical look at their practice with the assistance of an external facilitator (for AL) or researcher (for AR). The role of the facilitator/researcher was not to solve problems or direct change, or do this for teachers, but to teach them an approach to do this for themselves.

The Contract for Services for the evaluation of this development designated the following key research questions:

- What are the current structures in these settings that support effective pedagogy for maximising participation of the students in question?
- What specialist supports have contributed to the improved learning, social and cultural outcomes for the students in question?
- What are the most effective models of professional learning identified for teachers to optimise the learning and participation of the students in question?
- What ongoing supports can effectively maintain and enhance teacher capability to meet the needs of the students in question?

(Ministry of Education, 2005b, p. 12)

These questions were used to guide the recording of results for each phase of the evaluation.

**Evaluation design**

Three phases of a mixed method design (questionnaires, focus groups and success case studies) were planned for the overall evaluation in order to triangulate data and enhance the validity of the research (Burns 2000). Johnson and Onwuegbuzie (2004) note that this varied approach involves methodological pluralism, which results in superior research. The Johnson and Onwuegbuzie (2004) model positions the researcher to:

1. collect data using quantitative methods (for example, as in the questionnaires used in Phase One of this evaluation study)
2. collect data in naturalistic settings (for example, in the focus groups in Phase Two and the success case studies carried out in Phase Three of the evaluation)
3. be responsive to local situations, conditions and stakeholders’ needs (the focus groups and success case studies were designed to allow for this)
4. collect data in words and categories of participants who lend themselves to exploring how and why phenomena occur (designed for each phase of the evaluation).

The Unitec Research Ethics Committee (UREC) approved the design, all data collection instruments, the sampling process, and information and consent forms that were employed throughout the evaluation study. Formal consent was obtained from all participants, including caregiver consent from student guardians.

**Method/data collection tools**

**Phase One**

A broad questionnaire using both closed (continuum rating-scale) and open-ended questions based on the key research questions provided both quantitative and qualitative responses. The non-discrete continuum rating scales provided a ‘measurement
of peoples’ subjective states: their knowledge and perceptions, their feelings, and their judgments’ (Fowler 1995, p. 46). Open-ended questions were designed to add value to other answers from respondents (Cohen, Manion & Morrison 2000). Piloting of the draft questionnaire conducted with special school staff who were not involved in the program, resulted in only minor amendments for the final format.

The questionnaire, information and consent forms, and a covering letter, were mailed to all facilitators (GSE staff working with AL groups) and researchers (working with AR groups), and those principals, board members and participating teachers who were nominated by the facilitators/researchers (120 in total). It was intended that caregivers and students would be included; however, despite continued attempts to gain access information on these groups, inadequate or inaccurate contact detail provided by facilitators/researchers resulted in no questionnaires being issued to either of these stakeholder groups.

**Phase Two**

Eight focus groups were designed to further triangulate and strengthen the data gained in Phase One by using open-ended questions based on the same key research questions. Specifically, the focus groups aimed to: include student and caregiver perspectives; probe responses more deeply; and hopefully gain further evidence to support perceptions/participant self-report.

A focus group is described as ‘a small gathering of individuals who have a common interest or characteristic, assembled by a moderator, who uses the group and its interactions as a way to gain information about a particular issue’ (Williams & Katz 2001, p. 1). Hakim (2000) believes that eight is the optimum number for a focus group and in this evaluation, although eight participants were invited to each group, in reality frequently only four to six attended. Hakim (2000, p. 35) considers that focus groups:

produce less information on individual motivations and views than in-depth interviews can achieve, but they can yield additional information as people react to views they disagree with, or the group as a whole develops a perspective on the subject.

Both dissenting ideas and group perspectives were sought in the focus groups in this evaluation.

Focus group participants received the key research questions, as well as a summary of the Phase One broad questionnaire results, prior to meeting. Facilitator/researcher advice and support on protocols was sought for the Kura Kaupapa Māori (full immersion) schools. All focus groups were conducted at the school site and were tape-recorded and transcribed independently.

**Phase Three**

Success case studies were conducted with four schools. Brinkerhoff (2003) describes the ‘Success Case’ methodology as involving the following two-part structure: locating potential success cases; and determining and documenting the nature of the success. The latter was established via specific criteria for effectiveness that originated from the Ministry of Education and the criteria were utilised to help locate successful schools to enable selection.

Essentially observation, interviews and documentary analysis were employed as methods of data collection for the case studies. Once again, the key research questions guided development of these methods.

**Sample selection and response rate**

**Phase One**

No sampling process was engaged in for the questionnaire; it was issued to all stakeholders involved in the development program from each of the 49 schools. Of the 120 questionnaires posted, 40 were returned, that is, 30 per cent of the total issued. A breakdown of the number of responses categorised by development type, school type and respondent role is provided in Table 1.

The number of AL responses was approximately three times that of AR. An almost equal number of respondents (31 per cent primary; 36 per cent secondary) were located in mainstream primary and secondary schools and this was in keeping with the proportion of school types that were issued questionnaires. Just over four times (57 per cent) as many facilitators/researchers as teachers (14 per cent) responded, despite the fact that the total number of teachers issued the questionnaire was almost twice that of the number of facilitators/researchers. At 30 per cent, the overall response rate was low, even for a posted questionnaire which traditionally has a lower response rate than questionnaires issued face to face. Possible reasons (actual and hypothetical) for the low response rate include that: incomplete contact details were provided by some facilitators/researchers; there were concerns from facilitators/researchers that they would be expected to collect the data (a misinterpretation); the timing of the questionnaire distribution coincided with end of the term overload; and some facilitators/researchers believed that the evaluation was being conducted too early in the implementation process (a factor that the contracted evaluators had no control over).

**Phase Two**

Eight schools (four for AL; four for the AR initiatives) were sampled for the focus groups by choosing every sixth school on the list of each of the AR and AL schools. Each of the school types (see Table 1) was drawn in this sample and they were designated as A to G. A nominated coordinator in each school recruited focus group participants: the principal, at least one student, any teachers who
had participated, the board chair, a caregiver, and the facilitator/researcher. Only two students in total attended, one in each of schools B and G.

Phase Three
Four schools were nominated by Ministry of Education contract directors and facilitators/researchers as having met the success case criteria: one school had self-nominated that they were successful in the Phase One questionnaire and one focus group school demonstrated the success case criteria. The six schools that met the criteria were reduced to a sample of four by selecting a range of AL and AR schools and school type.

Analysis
Phase One quantitative data were manually collated and analysed using Microsoft Excel. Graphic representation of data was derived from this information where appropriate. The questionnaire was designed so that bar graphs of means, t-test (to determine whether the means of groups were statistically different) and correlation analysis could be generated from the non-discrete data provided. It was intended that data would be analysed comparatively based on development, school and stakeholder role type. Unfortunately, the ability to conduct the t-test and comparative data was deterred by the low response rate. In all three phases qualitative comments were coded under themes associated with the key research questions for the evaluation.

Ethical considerations
In all phases of this evaluation study formal consent was obtained from participants, with caregiver consent provided for any students involved. Detailed information outlining each phase of the research was provided to participants alongside relevant research questions. Confidentiality and anonymity of individuals and the school were also assured.

Summary of Phase One, Two and Three results
Despite a disappointingly low response rate (30 per cent) for the broad questionnaire, participants reported multiple strong positive outcomes for the development program. This optimism was largely reiterated also in Phase Two where results almost always confirmed and expanded the responses gained in Phase One. In the following section, overall responses for both phases linked to the key research questions are reported.

Focus
The results revealed that school projects were most often at a small scale, beginning point, with the aim
Evidence of impact

Several respondents in both phases noted that it was too early to evaluate the effectiveness of changes or to show causal effect between the development engaged in and impact. The majority of respondents in Phase One self-reported that they had data collection systems in their schools to demonstrate improvement, with systems for tracking academic achievement reported most frequently. Focus group respondents reported similarly but could show little evidence to verify such systems. Despite this, a considerable number of improvements and impacts from projects were noted with the greatest number falling under the heading of ‘social’ impacts for students. These included: enhanced cooperation and awareness of others; greater confidence, self-efficacy, self-management, and happiness; and calmer students who had improved behaviour. One student reported this impact in the following way:

I have learned more. I’m not rude to the teachers this year as much as I was last year. I am much better in behaviour and attitude.

Changes in values and attitudes of teachers were the next most frequently recorded the impact with teachers: working more collaboratively and cooperatively; having a greater awareness of inclusion; enhancing communication; changing their mindset; adapting teaching approaches for the student’s individual level and learning (including re-adapting the mainstream curriculum); showing greater flexibility and a less teacher-directed teaching approach; ensuring consistency/continuity and an integrated approach between classes; and creating new resources. The following primary teacher comment encapsulates their thinking about how they changed:

It made me really look within myself and to assess myself and to say ‘maybe you are not as good as you think you are?’ It showed up some real grey areas in my teaching. That course for me personally made me sit up and think, ‘I need to get off my backside and really think how I should teach that child writing.’

Multiple anecdotal academic achievements were noted for students, including: better retention of information and concentration; general improvement and work habits; better understanding; improved reading levels; and enhanced listening and reading skills. The following student and parent involved in Phase Two confirmed such achievement.

A student in a maths project said:

I got brainy.

The student’s parent stated:

My child is a very shy child, and I noticed that she is much better in problem solving, is much more confident in approaching any new subject. Her teacher advised us to provide a Year 9 maths book for her [student is in Year 8] and we did and she really enjoyed it. She would go and do any activity in the book. I have noticed a lot of difference.

It is important to reiterate that despite a multitude of anecdotal reports of impact, little evidence supporting self-report was provided.

Support structures—policies, systems, personnel

The majority of schools in the development program already had policies and/or guidelines on special education in place and either the curriculum was already adapted for special education or was at the development stage.

In terms of existing structures and support systems for effective pedagogy in special education, respondents cited the importance of board financial assistance for specialist staffing and resources, support from senior management, and having someone in the school driving initiatives. Existing school cultural norms that were seen as necessary included: commitment, good relationships, communication, cooperation, and a will to improve students and wanting them to be responsible, independent and self-managing.

Specialist support

Strong positive feedback was provided about the support from researchers/facilitators guiding the development program who were noted to have offered feedback, guidance, resources and tenacity. One teacher expressed the impact of the support in the following way:

The other reason for the success of the project at our school was that the GSE staff member herself was sufficiently creative and allowed flexibility for our staff. She wasn’t tunnel visioned about expectations.

Experts from outside agencies also provided specialist support, as did internal staff and student mentors. Although rare, barriers to specialist support also existed where changes in the assigned researcher/facilitator had occurred, where little assistance was provided by this key person, or where they had a limited background of AR or AL.
Effective forms of development
In Phase One, respondents considered the most positive component of development to be systematic, evidence-based, recording of student learning. Alongside this they noted the importance of clarity of planning and expectations, shared outcomes, and a realistic timeframe. In Phase Two the elements of effectiveness were expanded to include having: collaboration; buy-in or ownership; enthusiastic and committed teachers; the incorporation of reflection, dialogue and regular meetings; a clear vision and achievable goals for projects; small projects; support from, and access to, a facilitator or school management; and training components. Very few respondents in Phase Two indicated that they had any awareness of the principles of AR or AL or that they were conducting projects based on stages of evidence-based inquiry.

In both phases, barriers to effective models of development were noted to be linked to contract administration (initial misinformation, lack of clear information, hurried introduction, milestones due before schools started), a confused process, school coordinators not knowing what they were doing or not knowing enough about AR or AL, an unexpected requirement to present at a symposium, doing or not knowing enough about AR or AL or that they were conducting projects based on stages of evidence-based inquiry.

Maintaining and sustaining impacts
A raft of ideas was offered for maintaining and sustaining the effective impact of projects. These included: ongoing checking in; having good planning; reminding people of what works; ensuring ongoing communication; continuing meetings about the project; ongoing funding for learning support and teacher release; provision of more time; beginning with involving people who are willing, then extending from there; ongoing external and internal specialist support; continuing support from the board, principal, and senior management team; bringing other staff (wider than the project team) on board with the project focus to ensure wider collaboration, understanding and acceptance by mainstream teachers; and project team members taking on a training role with other staff where sharing of ideas could occur. The following comment from a teacher illustrates the latter:

*If I had to do this over again, I would pick these people here who were willing and say right let’s work together as a group. Then this group disseminates the information to the faculties.*

Results from success case projects
The success case studies (Brinkerhoff, 2003) were an important element of the design of the evaluation. The overriding feature of the Phase Three success cases that distinguished this small proportion of projects was that the participant action researchers/learners utilised ‘informed’ or evidence-based decision-making. Although many participants in Phases One and Two rationalised their limited use of data by noting that it was either too early to validate project outcome changes, or that it was difficult to show causal effect (that changes could be attributed to the development program), a hallmark of the four success cases was the use of strong evidence/data. Data was collected and reflected upon in the reconnaissance phases of the AR or AL and improvement initiatives were informed by both this data and relevant previous literature. Further, once the improvements had been implemented, additional data was collected to demonstrate change.

It is worth noting that the success case schools could demonstrate improvement outcomes for students, including increased confidence and self-efficacy, improved behaviour, enhanced achievement, better concentration and work habits, and increased inclusion by mainstream teachers and students. These schools were also clear about how to maintain and sustain momentum from the projects. Ideas included using committed staff to further drive the development, using ongoing external and internal specialist support, continuing financial and resource support from senior management, further inclusion of special needs in planning, celebrating success of students more widely, expanding the project team to bring other staff on board in order to strengthen understanding and acceptance, project team members taking on a training role with other staff, and allowing for time for transfer of knowledge about the project.

Themes that emerged
There are many similar themes that emerged from the comparative examination of the success case schools and this is perhaps the most interesting component of the evaluation because the themes both confirm what we already know about effective development and signpost further considerations. The themes are grouped under the key research question subheadings and include the success case project teams.

Focus
- defined their issue with considerable clarity
- planned well for the development and broadly aligned it to the school’s strategic planning goals
- started small with their projects, usually in the form of a trial, or small number of students, classes or teachers (only two in school 4)
- projects were manageable and adhered to the philosophy of ‘do a few things well’ (Piggot-Irvine 2006)
had a cautious and well-planned approach to development that was in keeping with the notion of deep learning (Biggs 1992; McKay & Kember 1997)

Evidence of impact

adapted the environment for students with special education needs (especially evident in school 2 where class size, room arrangement, curriculum and assessment opportunities were all adapted)

reported comprehensively on student outcomes (particularly in schools 2, 3 and 4)

reported that teaching practices had improved

Support structures—policies, systems, personnel

had overt board, principal and senior management team support (in three of the four success cases)—another characteristic associated with effective professional development (Baldwin 2005; Fletcher 2003; OECD 1998)

centred their projects on enthusiastic and committed teachers—a factor that is often associated with effective professional development (Fullan & Mascal 2000; Lewis 2003). Although this was probably most evident with schools 3 and 4, in all four cases there was an element of willingness of these teachers to give it a go

had high levels of collaboration and teamwork—yet another feature associated with effective professional development (Darling-Hammond 2000; Hill, Hawk & Taylor 2002; Lambert 2003)

used reflection and dialogue in collaboration—cornerstones of AR and AL (Zuber-Skerritt 2002)

collaboration/participation with parents and students was also reported, particularly in the case of schools 1 and 2 for parents. The full induction of students in the development was most evident in school 4

Specialist support

had strong external and internal specialist support, guidance and resources

Effective forms of development

followed the classic stages (issue definition, reconnaissance, implementation and evaluation) of AR and AL (even if not articulated as such)

used data/evidence to examine both the current situation and outcomes (pre- and post-implementation evaluations were conducted)—such use of evidence is considered to be an important feature of effective development (Allen 2005; Lewis 2003; Timperley 2004; Timperley et al. 2006)

referred to the examination of ‘best evidence’ or a relevant literature base in their project—they were ‘avid seekers of research and best practices that will help themselves and others’ (Lewis 2003, p. 2)

incorporated focused professional development (training) as a feature of the implementation phase of the projects.

Maintaining and sustaining impacts

celebrated success of students (especially schools 2, 3 and 4)

noted ways that they would sustain the development

Conclusion

Overall, all schools that demonstrated effectiveness in the development program, and the success case studies in particular illustrated commitment to long-term, ‘deep’ development, that is, development that went beyond a surface, or quick-fix, short-term training approach (Piggot-Irvine 2006). The development was part of a context-specific, situated, work embedded approach (Dempster 2001; Guskey 2002). However, despite being school-focused, the development did not show signs of being insular or introspective because there was contact with, and support from academics, outsiders and other organisations in order to broaden reflection, thinking and behaviour. Fletcher (2003) describes this as co-construction, where teachers and academics collaborate, offering a mutually informing, reciprocal model of support. Further, and in keeping with Hargreaves’ (1998) suggestion, the teachers were involved in the construction of the agenda and execution of projects in their schools.

In summary, the triangulated data adopted in this evaluation study has shown that schools engaged in effective development demonstrated ‘inquiry’, defined by Robertson (2005, p. 4) as ‘researching practice and seeking information’. In this inquiry, the teachers were self-directing yet collaborative, strongly focused, committed to a well-planned, evidence-based AR or AL and, most importantly, dedicated to student improvement.

Limitations

The most significant limitations associated with this evaluation involved the low response rate in Phase One, which was possibly linked to resistance to involvement associated with perceptions that the evaluation would fall on the shoulders of facilitators/researchers and that it was being carried out too early. This low response rate restricted the use of any comparative analysis between development, school or role type.

A second limitation was the lack of rigorous data to support respondent anecdotal self-reporting of outcomes in Phase One. It was intended that the focus groups would offer an opportunity to request...
evidence of outcomes but, in reality, little data was provided. Phase Three of the evaluation (the success case studies) did somewhat address the limitation of anecdotal self-report via in-depth success case studies in four schools.

These limitations do not render the findings reported on in this article invalid, but a degree of caution is centred on the generalisability of the outcomes.

References
Biggs, J 1992, Why and how do Hong Kong students learn? Using the learning and study process questionnaires, Education Paper 14, Hong Kong University.