



Collaborative Action Research within Developmental Evaluation

Learning to See or the Road to Myopia?

TINA COOK

Northumbria University, UK

This article investigates the use of collaborative action research in both the development and evaluation of a project designed to develop inclusive practice in Early Years and childcare settings. The purpose of the evaluation was to ascertain how practitioners understand the term 'inclusion', how those understandings were translated into practice and how changes in that practice might be conceptualized and carried out. The article explores the use of facilitated collaborative action research as a core element in developmental evaluation. Reflecting on other approaches such as theory of change, it considers whether the use of action research supported the critical examination and development of a project, or whether, by being so similar in design, it constricted the evaluation in terms of reliability of data, accountability and providing a framework for planning and development.

KEYWORDS: change; constructing realities; inclusion; learning; participatory action research

Introduction

The Inclusive Practice Pilot (IPP) project was commissioned in the UK by Newcastle upon Tyne Early Years Development and Childcare Partnership (EYDCP) to look at how inclusive practice might be developed across Early Years and childcare settings within the city. An evaluation was proposed to provide the funders (EYDCP) with information about the necessary conditions for developing inclusive practice to help bridge what was seen as a policy/practice divide. The policy was to be inclusive in the sense that inclusion is legislated for as a human rights issue. It is underpinned by a number of initiatives and treaties such as the United Nations Convention on the Rights of the Child (Office of the United Nations High Commissioner for Human Rights, 1989), the UNESCO Salamanca Statement (1994) and the Special Educational Needs Disability Discrimination Act (2001), and it is a requirement for all Early Years and childcare providers

(DfEE, 1997, 1998). However, in practice, the EYDCP recognized that exclusion still occurred. The purpose of this evaluation was to document how inclusion was conceptualized and practised across Newcastle and establish what might allow settings to develop and/or change their thinking and behaviour to enable more inclusive practice. The evaluation aimed to support knowledge building about inclusive practice through providing an account of the specific situation that would, in Winter's words, get 'sufficiently close to its underlying structure to enable others to see potential similarities with other situations' (2000: 1).

A form of evaluation was needed, therefore, that would explore below surface representations of inclusive practice and tease out principles to be further developed in Early Years settings across the city. It had to investigate how inclusive practice could be characterized, to identify what enabled participants to be aware of and even change their own thinking, the thinking within their setting, and ultimately translate changes in thinking into changes in practice.

Action research is a form of inquiry that uses the experience of trying to improve some practical aspect of a real situation as a means for developing our understanding of it (Winter, 2002: 27). It seemed, therefore, to be an appropriate approach to use in an evaluation that included facilitating the development of philosophies of practice, being a practical conduit for change in practice and recording that which enabled progress and change in practice to take place. It seemed that the closeness between the agreed principles and methods of action research and the needs of this evaluation would, when working together, provide the 'right tool for the job'. However, the question remained as to whether, by using participatory inquiry methods to both develop and evaluate, there was a danger of losing critical perspective. This article articulates that debate. It begins by giving a brief overview of the IPP project and the fundamental principles of evaluation and action research as defined and agreed with the funders of the project. It considers how inclusion and progress towards inclusion were conceptualized and articulated, how a range of perspectives and interpretations were gathered across the multiple realities and meanings that existed across a variety of Early Years settings, and how this way of gathering and using information contributed to the evaluation.

The Inclusive Practice Pilot Project

The Inclusive Practice Pilot (IPP) project was part of a series of initiatives conducted by the EYDCP to develop inclusive practice in Early Years services in Newcastle upon Tyne in the UK. The project was coordinated and evaluated by a small team from Northumbria University led by the author, a senior lecturer in Early Years and special educational needs. My own involvement in action research spans the use of action research for individual learning and development (my own), action research for service development (as head of a preschool service for children with special educational needs), and action research as part of externally funded evaluation projects.

The project ran for nine months. All Early Years and childcare settings in the city were invited to participate. The only criterion for acceptance on the project was that settings had to be committed to actively developing more inclusive practice. Broadly speaking, therefore, the IPP project worked with a group of participants committed to trying to improve some practical aspect of their situation as inclusive practitioners. Nine settings, of the following types, participated in the study: toy libraries, childminders, out of school clubs, private and voluntary nurseries, playgroups, parent and toddler groups and LEA nursery provision. Two staff members from each setting acted as representatives at project meetings.

The university team met fortnightly with a group of professionals and practitioners who had volunteered to be mentors for the pilot settings. Each mentor worked with an Early Years setting with which they were already familiar and where they had regular working links. Their role was to support the process of setting-based thinking, putting research into practice and providing data for the evaluation. They visited their setting between each fortnightly mentors' meeting and worked with that setting using the content of the meeting as a starting point. All participants, the coordinating team, mentors and practitioners from the settings met together on three occasions, once at the beginning of the project, once midway and again at the end. The university team acted as 'outside facilitators' for the mentors, and the mentors became facilitators for the settings by acting as a conduit with the university team.

The project proceeded in phases, each phase building on knowledge gained from the previous phase in a manner adapted from Kemmis and McTaggart's (1988) notion of an action research spiral and Elliott's (1991) revised notion of Kurt Lewin's action research model. The first phase concerned building understandings of inclusion and identifying contextual indicators of inclusion pertinent to the setting. The second phase involved participants conducting research with their setting both to identify notions of inclusion held by other people (e.g. staff and parents) and to check whether the indicators were helpful pointers for development. The third phase involved identifying action to be taken and reflected upon in the light of understandings developed during the project.

The Basis of the Evaluation

While evaluation can be characterized as a relatively new discipline, it has experienced tremendous interest, growth and development in recent years. However, many programme managers, when they call for evaluations to be undertaken, have limited experience, knowledge and understanding of evaluation. They tend to regard it as an audit-like accountability measure. Evaluation is expected to be a process of counting and checking up, of measuring value against an externally imposed accountability structure. Cook (1997) suggests that historically evaluation has been seen as a quantitative procedure consisting of the 'task of generating unbiased, precise estimates of the causal consequences of programmes of their major constituent parts' (Cook, 1997: 32).

Evaluation, however, has many variables that encompass a broad set of understandings and purposes identifiable in the ongoing differentiation of

evaluative procedures, for example, 'bureaucratic' evaluation (MacDonald, 1977), 'constructivist evaluation' (Guba, 1990), 'reflective evaluation' (Eisner, 1991), 'evaluation for social justice' (House, 1990), 'responsive evaluation' (Stake, 1997) and 'realist evaluation' (Kazi, 2003; Pawson and Tilley, 1997) to name but a few. It is not surprising then that when the evaluators and project/programme managers met together to begin to plan the IPP project, they had a variety of expectations of evaluation and its purpose. The first step was to identify the purpose of *this* evaluation.

Through discussions with the managers and funders at the outset of the project it was agreed that the project would be formative and would include:

- helping to establish and maintain communication amongst project participants by supporting 'enabling conversations' about programme quality and direction;
- helping to clarify appropriate indicators and criteria to account for success;
- assisting project managers and participants to adopt a (self-)critical perspective on the project and its activities; and
- providing opportunities for project members to develop evaluation skills incorporating a strong stakeholder focus.

As evaluators of the IPP project, the university team was mindful of the guiding principles of the evaluation as agreed between the evaluators and funders. In particular they paid attention to the requirement to support 'enabling conversations' about programme quality and direction amongst participants and to assist participants to take a reflective stance towards the project and its activities. This had to be weighed against a less well-articulated need of officers and managers (and to a certain degree project participants) for a set of recognizable specific understandings and actions that would become identifiers for/indicators of 'inclusive practice'. A balance needed to be struck between the search for concrete outcomes and a recognizable order, and the opportunity to raise the profile of the process as a meaningful happening in itself, without fixing it as a particular product. There was some tension between identifying certain practices as being inclusive and retaining an open and ongoing dialogue with participants about what currently characterized inclusive practice that would leave the door open to the possibility of new characterizations in the future. It was important to devise methods that would document and capture the insider experience and interpretations of that experience but at the same time be responsive to external evaluation requirements and understandings. Choosing a collaborative action research approach as the basis of the evaluation was a way of providing a framework for a process of thinking and action that could continue beyond the life of the project.

The evaluation of the IPP project was to be more than the use of applied social science methods to inform the project about whether it was meeting a given set of targets. It sought to get below externally predetermined representations of quality, efficiency and conceptualizations of practice. It was therefore a broader understanding of evaluation, based on Chelimsky and Shadish's (1997) three

perspectives on evaluation – evaluation for development, evaluation for knowledge building and evaluation for accountability – that was chosen to frame the approach used by the IPP project. It would support participants in the project in taking a new perspective on their work and in uncovering new and unrecognized knowing, characterized by Eraut (2000: 256; after Polyani, 1967) as tacit knowledge, that is, knowledge that we have but cannot easily describe or explain. Marra (2004: 269) notes that the bulk of evaluation literature finds participatory design and interactive processes of data collection and analysis to be the most effective ways to socialize tacit knowledge. Socialization, according to Marra (2004), means that the participants ‘not only come to understand each others’ definition of shared situations but also agree on a common identification and “justified true belief” about how to act in that situation’ (Marra, 2004: 269). She goes on to say that

. . . evaluation-based information is more than the specific information required immediately by each individual. The sharing of the extra information between individuals promotes the sharing of individual tacit knowledge and members share overlapping information. (Marra, 2004: 279)

This joint understanding could then be used to develop indicators for evaluating practice and as the basis for transforming practice. It would also lend itself to theory and knowledge building, particularly in the area of conceptualizing practice.

Methods

The evaluation of this project gathered documentation of action research undertaken by participants and discussions about the meaning of that research at fortnightly meetings. As they struggled to deconstruct preconceived notions of inclusion, the attempts of participants to make practice meaningful to terms of new understandings and to build new ways of doing were collected and collated. Documentation such as diary or field notes kept by participants as part of their own reflective practice and research was also brought to the table at the meetings to be discussed and key issues identified. Key themes emanating from all documentations were then returned to the participants to form part of further discussion at the fortnightly meetings. Semi-structured interviews took place before the project started, during the project and at the end of the project and were recorded by the university team. In this way fundamental ‘whys’ of practice from a number of perspectives were gathered together without reducing the complexity of diverse understandings by trying to make them fit into a particular reporting structure.

On the basis of these documentations the group (participants and evaluators) focused on identifying indicators of inclusive practice gathered from the range of Early Years and childcare settings, to help map programme development. The EYDCP already had a set of Five Key Principles for Inclusive Practice which underpinned the work of the project (Box 1).

Box 1. Key Principles for Inclusive Practice

1. We value all people in society.
2. We will provide quality services for all children.
3. We recognize the right of children to be involved in all decisions that affect them.
4. We recognize the role of families and carers to be decision makers with children.
5. We will not limit our expectations.

We recognize the need to identify and remove all barriers to inclusion which relate to:

- the environment;
- people; and
- our organization.

Source: EYDCP September 2001.

The engagement of those who struggled daily to balance actual working lives with perceptions of 'perfect' practice occasioned by external requirements produced a set of active, needs-led indicators (Box 2) that also seemed to fit within the Five Key Principles for Inclusive Practice outlined in Box 1.

These indicators were a synthesis of the combined 'knowing' of what was considered worthwhile by participants, critiqued by themselves in collaboration with others. Collected from within practice, they were expected to provide direction for future programme planning.

Box 2. Examples of Indicators of Inclusive Practice, IPP Project, 2002

- Happy children, parents and practitioners
- Families are involved in dialogue, planning and decision making
- Children are involved in dialogue, planning and decision making
- 'Being made welcome' is paramount when engaging with children and families
- Inclusion is actively discussed and all perspectives are carefully thought about
- There is a positive attitude towards problem solving
- Confident staff who make parents feel confident about leaving their children
- Practitioners/professionals and managers ask questions about *what* they are doing and *why* they are doing it
- Access for all really means for all
- The environment is included in planning
- There is visual evidence of respect for other cultures and difference
- Staff organization is carefully considered on a regular basis
- Good work is being recognized and celebrated
- Active thinking is leading to changes and evolving practice
- Policy and practice are interlinked
- New knowledge is always being sought and utilized

Analysing the data to uncover key themes was achieved through a modified Delphi technique. Ziglio (1996) describes the Delphi method as a structured process for collecting and distilling knowledge from a group of experts by means of a series of questionnaires interspersed with controlled opinion feedback. In this case the experts were the participants from Early Years settings and questionnaires were replaced with meetings, groups and individual interviews. The feedback from each session was either given informally as a group topic for discussion or, halfway through project, 'thoughts so far' were written into a 'snapshot' report and delivered to participants with a questionnaire feedback option or the option to meet and discuss with the researcher. Both methods were used by participants.

Ultimately the project participants identified that it was not the use of indicators, but the interpretation of the indicators, that was the key to the development of inclusive practice and the aspect of the project that really needed to be evaluated was the process, not the outcome. Some participants suggested that target or checklist methods for measuring success in inclusive practice could, in essence, be seen as part of the problem rather than the solution. For instance:

The Government think that more is better [more children in mainstream schools] but more children in mainstream settings is not necessarily a good thing at this point if they are just put in and nothing is done to change practice to accommodate them. But at this stage of the game we might need to really think about whether practices we have in our settings actually exclude children from taking part, so more is not better, more is worse for those children, but they get seen as success stories. (Interview c)

Methodological Dilemmas and Choices Made

Setting an evaluation so strongly within an 'insider participatory' paradigm raised specific tensions. Externally imposed measures of quality have traditionally been held up as more objective and hence reliable indicators and measures. Scriven (1997) argues that, while there is a strong attack on distancing, distancing and 'objectivity' remain the correct ideals for the external evaluator. He states that: 'the closer we come to them [distance and objectivity], the more accurate our conclusions are likely to be, other things being equal'.

Validity is considered to be higher when the evaluator is distanced from project participants and not drawn into the complexity of their discussions, perceptions and formulations about what constitutes programme quality. If the evaluator remains aloof and maintains a given criterion for evaluative critique, their judgements are characterized as being unbiased and more valid.

The strength of performance management type approaches to evaluation, which emphasize predetermined programme outcomes as measures of quality, tend to lie in providing 'important short-term, quick turn-around information for tracking progress against stated goals' (Blalock, 1999: 142). The absence, however, of what Blalock (1999: 117) calls 'a commitment to collect information about why and how those results occurred' renders the performance management model less useful as a basis for programme development. Interpreting 'worthwhileness'–

what fundamentally enables a programme to function, what gives it meaning and ultimately why it is initiated – is a complex and often contentious issue within the evaluation task that is often left unaddressed and ill-defined. Despite continued debate, measurable outcomes that search for stable, objective truths continue to remain the ‘gold standard’ for many.

In the case of the IPP evaluation, while the questions ‘what’ and ‘how many’ were part of the evaluators’ armoury, the question ‘why’ took on particular significance within the enquiry. As one participant in the project stated, the key element of developing inclusive practice lay not in identifying what practitioners do, or could do, but in finding out: ‘if this is the way we do it, *why* do we do it this way and what is the effect of that?’ (Fieldnotes).

The ‘why’ became a crucial aspect of the process of developing inclusive practice and the answers to the ‘why’ questions in terms of what made practice worthwhile were seen as being a starting point for action planning for change.

Suggesting that evaluators and participants might work together in making decisions about ‘worthwhileness’, good practice and quality, using an insider-researcher/evaluator approach, requires a paradigmatic shift from an approach aimed at proving something against a given standard. It means moving towards accepting that there is no homogeneous standard of value to unite all parties. The identification of meaningful processes that ‘make sense of what is going on’ (Dahlberg et al., 1999: 107) are seen, at best, as the icing on the cake which can be disregarded or abandoned according to taste or in the interest of efficient use of evaluative resources. As Somekh (2002: 89) points out, however, there are multiple realities, not just one: ‘There is more than one construction of the world, which means that there is more than one way of deciding upon what “counts” as knowledge.’

Blumer also warned that trying to catch the interpretative process by remaining aloof as a so-called ‘objective’ observer and refusing to take the role of the acting unit is ‘to risk the worst kind of subjectivism – the objective observer is likely to fill in the process of interpretation with his own surmises in place of catching the process as it occurs in the experience of the acting unit which uses it’ (Blumer, 1969: 86).

Externally imposed systems and measures are not necessarily sufficiently finely tuned to enable an evaluator to get to the heart of what gives a project or programme meaning. Claims for accountability that use predetermined preconceptions and standardized external measures are not always helpful in making judgements or in capturing the fine threads that weave together the relative merits of programmes and practices. They can be a blunt tool that reduces the ‘knowing’ to the measure of particular observable or reproducible variables. They may involve identifying certain predetermined features on a basis that can be quantified and pass over some of the more esoteric aspects of a project that may be difficult to quantify but have a major impact on programme development. Deciding on representations of quality without inquiring into the complex features of everyday practice can be seen as tantamount, in traditional scientific terms, to making decisions without knowing all the ‘facts’. The ‘facts’ in qualitative

evaluations may be far from tangible and observable but have equally important effects on the way organizations develop practice programmes.

A tight framework of externally imposed, measurable imperatives may not allow for the construction of concepts of quality that capture the changing and multiple perspectives of programme development over time. They may bias the evaluation towards gathering information about and reinforcing the 'known', rather than the 'yet to be understood'. This was particularly pertinent in the IPP project, which invested energy engaging participants in collaborative reflection on topics to support emerging knowledge and to find appropriate indicators of development. The long arm of history does, however, still reach deep into the mindsets of both organizations and practitioners. The question 'how can this be a good evaluation if we haven't counted anything?' is, in my experience, one that is still asked, or at least is implicit. The next question tends to be 'how can this be reliable evidence if it is given by the participants?' My question would be, 'how could evidence be reliable if participants are not involved?' 'Knowledge constructed without the active participation of practitioners can only be partial knowledge' (Somekh, 2002: 90). Practitioners do know, but they sometimes need some help to develop that knowing into something they can then recognize and act upon.

If externally derived indicators are used for describing what makes effective programme development, an error in judgement may already have been made in terms of what constitutes quality, good practice and worthwhile development. Evaluators and programme policy makers then invest in that error, with evaluators meticulously monitoring the process of practitioners learning how to pursue a course with a great deal of exactitude but with little worth. Eisner (1998) suggests that the features of the work itself should guide the criteria applied to judge it. In addition, if standards/targets and/or indicators are not contextually appropriate, evaluators may find themselves in the position of identifying what programmes have not done, even though the programme has taken an appropriate course of action under prevailing circumstances, rather than identifying positive characteristics in a programme's development and practice. The subject of the evaluation may then be incorrectly judged as failing, with all the damaging consequences of such a diagnosis.

If quality is characterized as a socially constructed notion of what merits being termed worthwhile, it is affected by context, history and perspective. How then, could a construction of worthwhileness and quality that includes being meaningful to programme development be recognized through checking against a set of externally imposed criteria? A more informed recognition of worthwhileness is found through a synthesis of both internal and external understandings that are made meaningful in context. If the notion of 'worthwhileness' was to underpin the basis of judgements about quality there was a need to acknowledge and address the multiple perspectives and realities present amongst the range of practitioners and their settings. Removing the multiple perspectives that come together in complex social organizations, you remove a large part of the contextual element that affects behaviour and decision making in policy and practice. As Somekh

(2002) suggests, researchers who are not part of the action context have a tendency to oversimplify their analysis and assume a simplistic cause–effect relationship between phenomena and events.

Theory-Based Evaluation and the IPP Project

The approach taken to the evaluation of the IPP project drew on aspects of two theory-based evaluation approaches that have come to the fore in recent years: theories of change (ToC) and realistic evaluation. ToC is a term developed through the work of the Aspen Institute Roundtable initiative as a way of describing ‘the set of assumptions that explain both the mini-steps that lead to the long term goal of interest and the connections between program activities and outcomes that occur at each step of the way’ (Weiss, 1995).

ToC employs the integration of process and outcomes in evaluation, asking what happened, how and why that happened, with a focus on developing new theories for action. It ‘delineates the pathway of an initiative by making explicit both the outcomes of an initiative (early, intermediate, and longer term) and the action strategies that will lead to the achievement of these outcomes’ (Connell and Klem, 2000: 94). Connell and Kubisch (1998) suggest that a good ToC design begins with programme staff identifying the outcomes they hope the programme might achieve. In this way it can

1. sharpen the planning and implementation of an initiative;
2. facilitate the measurements and data collection as part of the process so avoiding the risk that evaluation will be driven by the tools rather than vice versa;
3. through articulating the theory of change at the outset of the programme, strengthen the scientific case for attributing subsequent change to the activities included in the initiative.

This articulation of the ToC framework led to a number of questions in relation to the IPP evaluation about when and how a theory of change approach might be developed. If programme staff are asked to identify outcomes at the outset, whilst this does support the development of contextually appropriate, understandable and achievable measures/indicators for change, how do participants see beyond what is already there to what might be possible – and whose views about the future prospects should they build on? A strength of the ToC approach appeared to be in the detailed analysis of the programme in order to identify what it is about the current programme that could lead to development. This was an aspect heavily drawn on by the IPP project. However, where this occurred in the evaluation framework was problematic. The question for the IPP evaluation was ‘how could participants plan for change and understand what might enable change to happen before the process of developing understandings about what things might look like had taken place?’ While it is suggested that the ToC approach ‘forces program staff to examine their own beliefs about what works, for whom

and under what circumstances' (Kagan, 1998: 115), they give little indication as to how this might happen.

Blamey and Mackenzie (2002), when using a ToC approach to evaluation of two Scottish National Health Demonstration Projects, also raised this issue. They noted that ToC tended to lead to a very linear approach to planning and evaluation which may miss or mask some of the very complex interactions within and between projects. It may be unable to uncover unexpected outcomes or synergies and the skills and procedures for monitoring are not sufficiently sensitive or responsive to the complexity that exists (Blamey and Mackenzie, 2002: 14). What a 'final, usable version' of ToC might look like was lacking. While ToC led directly to improved planning they questioned whether the approach could really get to the heart of 'which aspects of a complex programme of activities work with which sub-groups of the population and in what circumstances' [italics in the original] (Blamey and Mackenzie, 2002: 15).

The second approach drawn on by the IPP project is known as realistic evaluation. This places a particular focus on generating theories underlying programme design through detailed analysis, in order to identify what the programme is about and what might produce change. This detailed analysis is then used to identify activities/measures that might produce change and 'which individuals, subgroups and locations might benefit most readily from the program, and which social-cultural resources are necessary to sustain the change' (Pawson and Tilley, 1997: 85).

Pawson and Tilley (1997) claim that those who follow a realistic evaluation approach are 'whole-heartedly *pluralist* when it comes to the choice of method' and attempt to carefully tailor the method to the form of hypothesis being used. The notion of 'what works for whom' is strong in this type of evaluative approach and therefore attractive to an evaluation process such as the IPP that sought to match method to form and build on diversity.

The principles behind both these approaches informed the design of the evaluation of the IPP project.

The Role of an Action Research Approach in Evaluation

Discussion abounds about key principles of action research, how it is distinguishable from other forms of enquiry and what its role and function might be. For the purposes of the IPP project, action research was conceptualized as being aligned with processes of research that strive to represent all voices. The value of action research lies in its intent to develop collaborative thinking on and through action. It engages participants in creating better understandings of what they are doing now, why they are doing it now, whether it matches their understandings of what should be done, whether there are other ways of doing this that meet their understandings and what is needed to create change.

Elliott, a member of the team involved in the Humanities Curriculum Project (Stenhouse, 1975, 1980), suggested that action research consists of: 'review, diagnosis, planning, implementation, monitoring effects [and] provides the necessary

link between self-evaluation and professional development' (Elliott, 1982: ii, 1). For Elliott, then, action research, grounded firmly in the personal, was part of self-evaluation, the fundamental aim being to improve practice rather than to produce knowledge. 'Within this form of educational inquiry theoretical abstraction plays a subordinate role in the development of a practical wisdom grounded in reflective experiences of concrete cases' (Elliott, 1991: 53).

As already outlined, the IPP evaluation had, as part of its remit, the intent to engage in theoretical construction, both about the nature of inclusive practice and about the use of action research as an evaluation and development tool. The conceptualization of the action research approach in this case went beyond Elliott's (1991) view just quoted. It included the need to create what Whitehead (1989) termed a living educational theory about what gives action 'validity' in context and how certain principles might inform the practice of others. Questions about what is worth doing, what is 'good practice' and what is worth recording are inextricably linked.

Central to the action research process is that it separates rhetoric from well-rehearsed notions of practice. Thus action research can be used to puncture and critique the general worldview of practitioners. It can disturb their current satisfaction with what they have, uncover tacit knowledge and understandings and support participants in moving beyond the familiar to learn something new from their own work: 'action research . . . involves questioning the meaning of data so that participants can go beyond the already "expert" understandings which defined their starting points' (Winter, 2002: 36–8).

Participants in the IPP project had volunteered to participate in the project because they were committed to 'inclusion'. For many, however, whilst inclusion had been a constant companion it was a fairly vague notion. Being 'committed to inclusion' had not necessarily engendered in-depth thinking with a resultant reshaping of both general and specific practices involving children with different needs. Inclusion was fraternized with: understood in terms of its intentions without embedded ways of working being confronted and critiqued in terms of philosophies, beliefs and understandings. It had invoked what Roland Barthes, the French philosopher, linguist and literary critic, termed 'docile interest' (1982). The degree of interest invested in inclusive practice was never, to borrow from Barthes's terminology, a 'delight or . . . pain' for practitioners (1982: 28). Oliver (1996) argues that for inclusion to take place it must be struggled for.

When talking about photography and its effect on thinking Barthes used the term 'studium' to describe the average effect or 'docile interest' that is engendered when looking at certain photographs. These photographs have meaning for him as he is familiar with what they portray and with their message but he only feels 'a kind of general, enthusiastic commitment . . . but without special acuity' (p. 26). They do not inspire him to question or think in depth about the photographs. He accepts their generality as it is. Barthes continues by reflecting on other photographs that affect his world in a more active sense. When he looks at these photographs, there is something about a particular aspect of the photograph that

has the ability to 'prick, disturb and wound' the comfortable studium. This aspect he termed the 'punctum point'. This seems to reflect a key role of the action research process in evaluation. It provided a prism rather than a mere window through which participants looked at their own work to find meaning, develop thinking and articulate new knowledge. Action research shifted the discussions held amongst practitioners from description to reflection, then from reflection to critical self-reflection and to begin to yield the details that constituted the very raw elements of their work. It enabled a number of participants in the IPP project to find their 'punctum point' in the picture of inclusive practice: the point that had meaning for them and that would both inform and effect change. The punctum point may be different for each participant, but serves the same purpose. It moves thinking beyond the chimera or gloss of accepted ways of being and behaving to identifying contradictions and areas of rub within that practice. Action research had a facility to support participants to have that 'bolt from the blue' realization. It enabled them to get below their own rhetoric, and that of others, to identify actual practice and the meaning behind practice rather than theories of idealized practice.

We realized that although we had accepted him into our playgroup, [the basis on which they had previously defined themselves as being inclusive] he was only in the building and now we had to work on helping him be part of what goes on here. That is the difference really, the difference I now see between integration and inclusion. (Interview b)

I think we had been fairly at an integration level [as opposed to inclusion]. Now it's not just the case of people coming into the nursery and fitting in with our routine . . . it's about us changing too. (Interview d)

Both these participants had a realization that their current practice, whilst meeting certain administrative conditions for inclusion, was not really fulfilling a central tenet of inclusion, that is, that a child should be an active participant. This led them to think not just about current practice, but how they could develop future practice and, more importantly, continue to critique any future practice they may develop.

Deconstructing both embedded and consigned notions of inclusion was not easy and involved some participants in wrestling with standards (their own as well as externally imposed standards), targets and organizational dictates that had been accepted as, and broadly translated into, guidelines for practice. Participants suggested that in the past inclusion had been a set of 'things to check', such as access to buildings, and 'ensuring that children with special educational needs are given a place, an opportunity to be in the same building as the others' (Interview b).

For many within the project, the process of learning to reflect was a key element in the change process. 'Staff feel the most important aspect of their involvement in the project has been their reflection on the principles of inclusion' (Interview a). Action research goes beyond critical reflection; it has a commitment to instigating action based on that reflection. In the IPP project, action plans for addressing the 'rub' within practice varied across settings but insider imperative to create steps for change was an important aspect of the project:

Everyone should have the opportunity to attend something like this – it's made us all think and then feel confident about trying to do something about it – knowing we can do something here. (Researcher fieldnote)

Outcomes of the Evaluation

Whilst one outcome of the IPP project was a set of indicators for inclusive practice, a key outcome was the recognition that it was the way of working towards those indicators that was crucial for development. Participants suggested that the activity of engaging in critical reflection on and in practice – not the identification of a given set of standards – provided a basis for active change (see Box 3).

Box 3. Key Elements to Developing Practice

The following were identified by participants as key elements that enabled them to develop new understandings about inclusion in theory and practice.

- Access to frequent/regular support for both settings and mentors
- Starting from where you are – being realistic
- Theory and practice linked to your own setting
- An expectation of action and reporting on that action
- Pertinent and helpful activities to develop thinking
- A framework for thinking as well as doing
- Providing new perspectives of the familiar
- Opportunities to hear and learn about the work of others
- The process of having to identify *why* your good practice is good
- Opportunities for focused discussion that made you investigate thinking alongside practice
- Opportunities for focused collaboration both across settings and within settings

A second outcome of the evaluation was the intention to develop a training programme for Early Years practitioners in relation to working with children with special educational needs. This would not be a didactic programme on the features of inclusive practice as these would differ across organizations and change over time. This training would be about developing the persona of an action researcher with inclusion as the focus. Participants in the IPP project, as part of defining what has supported their thinking and development, suggested that the future learning would need to include the elements listed in Box 4. It would aim to provide the opportunity for participants to address the 'studium' and hopefully to be hit with a 'punctum point' or two! The punctum point, where tacit knowledge, known knowledge, seeing and learning come together was seen by participants as a key facilitator of change.

Box 4. Key Elements for Supporting Development and Change

Participants suggested that the following would be necessary to continue the development of inclusive practice across the Early Years sector.

- Current participants should have continued opportunities to get together and discuss practice, but not on such a regular or intense basis.
- New participants should have the opportunity to investigate their own practice in a similar manner to the current project. They suggested the development of a course that would encapsulate the key elements of the IPP project (see Box 3).
- Participants on the course should represent and collaborate with all staff in their setting and a representative from all settings should eventually be *required* to attend the course.
- Mentors were a necessary element in successful change.
- Policy makers, managers and practitioners should attend training.
- Training should be made accessible to all, including parents.

Discussion

How is practice characterized as worthwhile? What should count as evidence of worthwhileness? Who decides? How can knowing facilitate action? These questions lie at the heart of the debate about what evaluation is for and what it does, and hence they lie at the heart of this article. If the overall aim of evaluation is, as Weiss (1995) suggests, to assist people and organizations to improve their plans, policies and practices on behalf of citizens, then it is important that real understandings of practice and philosophies of practice are the foundations for planning. Differences in perspective and emphasis across stakeholders in projects need to be teased out and engaged with. Not to do so would result in building on a chimera of understanding that could not offer firm foundations for development.

The use of collaborative action research offered a means of getting close to finding out what might produce new understandings and how that might link to a change in practice. The strength of action research lay in making meaning of current activities to inform future change practice, an aspect of evaluation that has been cited by Connell and Kubisch (1998) as the 'hardest part of the theory articulation process'. When Weiss (1995) hypothesized that a key reason complex programmes are difficult to evaluate is that the assumptions that inspire them are poorly articulated and that stakeholders of complex community initiatives typically are unclear about how the process will unfold, she stated that one key reason for this was that evaluations paid insufficient attention to early and mid-term needs in order for a long-term goal to be reached. A strength of the IPP evaluation was its emphasis on early understandings in the first stages of the evaluation design. Using collaborative reflections to begin to break down generally held beliefs and assumptions and build on tacit knowledge through articulation of issues drew in the multiple perspectives endemic in such a project.

It brought together understandings, development and implementation in practice. As Somekh points out, action research takes account of the need to integrate the construction of knowledge with its enactment in practice: 'The epistemology which underpins action research methodology is distinctive in that it rejects the notion that knowledge can be de-contextualised from its context of practice' (Somekh, 2002: 90).

In terms of strategic development based on learning from the evaluation, action research may have much to offer in terms of improving understanding, development and change in practice and, perhaps more importantly, finding out how change might occur in the future. However, it was weaker in the area of organizational planning. As used in this evaluation, action research was unlikely to produce a blueprint set of change pathways, but as suggested by the discussion throughout the article, that was not deemed appropriate for this type of evaluation. Change that came from this evaluation tended to be a pragmatic response to reflection and sudden realization (the 'punctum point' effect) engendered through the practical processes of the evaluation. An emphasis on strategic planning for change embedded at the beginning of a project, informed by ToC approaches, could be helpful in strengthening the strategic planning process within evaluations using an action research approach, and could perhaps go some way to engaging with policy makers. One notable omission from the participant list in this project was the managers and policy makers. Although invited and encouraged to participate, they felt unable to do so. This may have long-term implications for strategic planning from the learning acquired in this project. This is something to learn from and take forward to future evaluations using this approach.

Concluding Thoughts

It would appear then, given the purpose of this evaluation as articulated at the outset, that the use of collaborative action research was a reasonable response to the needs of the participants and their managers/local policy makers. It enabled the evaluator to work with participants to delve deep into their understandings of inclusion and begin to tease out the complex and temporal meanings that form the basis of current practice. It evoked an essence of 'knowing' where multiple perspectives told different stories and supported participant enquiry into what they personally meant by inclusion and what others might mean by the use of the same word. The use of self-evaluation and self-reflection as critique to put common understandings to the test in a collaborative forum supported the unearthing and synthesis of complex and varied meanings from a range of perspectives. The multiple perspectives gathered through the discussions and research, plus the varied opportunities for both data collection and analysis, gave strength, meaning and – to borrow a word from a more positivist paradigm – validity to the project. The work has resonance with similar work undertaken by Marra (2004) where the process of building evaluative knowledge was seen to take place only when organizational members reflected on their actions.

The use of action research also strengthened two other key dimensions of this evaluation. First, it provided a workable base for development that delved

beneath the general representations of practice (studium) that can be mistaken for 'knowing'. Second, the 'hows' of practice change have been addressed through the direct linking of critique and change embedded in the evaluation design at the start of the project. Whilst it is not always comfortable for participants to have their understandings and beliefs questioned in this way, not to do so would have left the project with an unstable basis for development. This process of discussion and supportive critique offered a dynamic learning process that worked towards renewed understandings and continued change.

Searching for one simple truth, capturing one objective measure of worthwhileness, one way forward, could have left the picture whole, its fabric undisturbed and the basis of its development unknown. Perhaps then, drawing on the essence of a number of theories of evaluation and through using a number of methods to develop, research and evaluate the project, action research was a reasonable tool for the job; it was a reasonable evaluation; it evoked some 'truths' and offered ways to further development.

We often photograph events that are called 'news' but some tell the news step by step in detail as if making an accountant's statement. Such news and magazine photographers, unfortunately, approach an event in a most pedestrian way. It's like reading the details of the Battle of Waterloo by some historian: so many guns were there, so many men were wounded – you read the account as if it were an itemisation . . . Life isn't made of stories that you cut into slices like an apple pie. There's no standard way of approaching a story. We have to evoke a situation, a truth. This is the poetry of life's reality. (Cartier-Bresson, 1989: 425)

References

- Barthes, R. (1982) *Camera Lucida*. London: Jonathan Cape.
- Blalock, A. B. (1999) 'Evaluation Research and the Performance Management Movement: From Estrangement to Useful Integration', *Evaluation* 5(2): 117–49.
- Blamey, A. and M. Mackenzie (2002) 'Using a Theories of Change Approach: Examples from the Scottish Health Demonstration Projects', conference presentation, Belfast Healthy Cities, Belfast.
- Blumer, H. (1969) *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, NJ: Prentice Hall.
- Cartier-Bresson, H. (1989) 'An Interview with Byron Dobell 1957', in W. Manchester (ed.) *In our Time: The World as seen by Magnum Photographers*, pp. 424–5. London: W. W. Norton & Co. Inc.
- Chelimsky, E. and W. R. Shadish, eds (1997) *Evaluation for the 21st Century: A Resource Book*. London: SAGE.
- Connell, J. P. and A. D. Klem (2000) 'You Can Get There from Here: Using a Theory of Change Approach to Plan Urban Education Reform', *Journal of Educational and Psychological Consultation* 11(1): 93–120.
- Connell, J. P. and A. C. Kubisch (1998) *Applying a Theories of Change Approach to the Design and Evaluation of Comprehensive Community Initiatives: Progress, Prospects, and Problems*. Washington, DC: Aspen Institute.
- Cook, T. D. (1997) 'Lessons Learned in Evaluation over the Past 25 Years', in E. Chelimsky and W. R. Shadish (eds) *Evaluation for the 21st Century: A Resource Book*, pp. 30–52. London: SAGE.

- Dahlberg, G., P. Moss and A. Pence (1999) *Beyond Quality in Early Childhood Education and Care: Postmodern Perspectives*. London: Falmer Press.
- DfEE (1997) *Excellence for All Children*. London: DfEE Publications Centre.
- DfEE (1998) *Meeting Special Educational Needs: A Programme of Action*. London: DfEE Publications Centre.
- Eisner, E. W. (1991) *The Enlightened Eye: Qualitative Inquiry and the Enhancement of Educational Practice*. New York: Macmillan.
- Eisner, E. W. (1998) 'The Primacy of Experience and the Politics of Method', *Educational Researcher* 17(5): 15–20.
- Elliott, J. (1982) 'Action-Research: A Framework for Self-Evaluation in Schools', Working Paper No. 1 in R. Winter (ed.) *Learning from Experience: Principles and Practice in Action-Research*. London: Falmer Press.
- Elliott, J. (1991) *Action Research for Educational Change: Developing Teachers and Teaching*. Buckingham: Open University Press.
- Eraut, M. (2000) 'The Intuitive Practitioners: A Critical Overview', in T. Atkinson and G. Claxton (eds) *The Intuitive Practitioner: On the Value of Not Always Knowing What One is Doing*, pp. 253–268. Buckingham: Open University Press.
- Guba, E. G. (1990) 'The Alternative Paradigm Dialogue', in E. G. Guba (ed.) *The Paradigm Dialogue*, pp. 17–27. Thousand Oaks, CA: SAGE.
- House, E. (1990) 'Realism in Research', *Educational Researcher* 20(5): 2–9.
- Kagan, S. L. (1998) 'Using a Theory of Change Approach in a National Evaluation of Family Support Programs', in K. Fulbright-Anderson, A. C. Kubisch and J. P. Connell (eds) *New Approaches to Evaluating Community Initiatives*, vol. 2, *Theory, Measurement, and Analysis*, pp. 113–21. Washington, DC: Aspen Institute Roundtable.
- Kazi, M. A. F. (2003) *Realist Evaluation in Practice: Health and Social Work*. London: SAGE.
- Kemmis, S. and R. McTaggart, eds (1988) *The Action Research Planner*. Victoria, Australia: Deakin University Press.
- MacDonald, B. (1977) 'A Political Classification of Evaluation Studies', in D. Hamilton, D. Jenkins, C. King, B. MacDonald and M. Parlett (eds) *Beyond the Numbers Game*, pp. 224–7. London: Macmillan.
- Marra, M. (2004) 'The Contribution of Evaluation to Socialization and Externalization of Tacit Knowledge: The Case of the World Bank', *Evaluation* 10(3): 263–84.
- Office of the United Nations High Commissioner for Human Rights (1989) *United Nations Convention on the Rights of the Child*. Geneva: UNHCR.
- Oliver, M. (1996) *Understanding Disability: From Theory to Practice*. London: Macmillan.
- Pawson, R. and N. Tilley (1997) *Realistic Evaluation*. London: SAGE.
- Polanyi, Michael (1967) *The Tacit Dimension*. New York: Anchor Books.
- Scriven, M. (1997) 'Truth and Objectivity in Evaluation', in E. Chelimsky and W. R. Shadish (eds) *Evaluation for the 21st Century*, pp. 477–500. London: SAGE.
- Somekh, B. (2002) 'Inhabiting Each Other's Castles: Towards Knowledge and Mutual Growth through Collaboration', in C. Day, J. Elliott, B. Somekh and R. Winter, *Theory and Practice in Action Research: Some International Perspectives*, pp. 79–104. Oxford: Symposium Books.
- Special Educational Needs Disability Discrimination Act* (2001) London: Stationery Office.
- Stake, R. E. (1997) 'Advocacy in Evaluation: A Necessary Evil?', in E. Chelimsky and W. R. Shadish (eds) *Evaluation for the 21st Century*, pp. 470–6. London: SAGE.
- Stenhouse, L., ed. (1975) *An Introduction to Curriculum Research and Development*. London: Heinemann.

Evaluation 12(4)

- Stenhouse, L., ed. (1980) *Curriculum Research and Development in Action*. London: Heinemann Educational.
- UNESCO (1994) *The Salamanca Statement and Framework for Action on Special Needs Education*. Paris: UNESCO.
- Weiss, C. (1995) 'Nothing as Practical as a Good Theory: Exploring Theory-Based Evaluation in Complex Community Initiatives for Children and Families', in J. Connell, A. Kubisch, L. Schorr and C. Weiss (eds) *New Approaches to Evaluating Community Initiatives*. Washington, DC: Aspen Institute.
- Whitehead, J. (1989) 'Creating a Living Educational Theory from Questions of the Kind, "How do I Improve my Practice?"', *Cambridge Journal of Education* 19(1): 41–52.
- Winter, R. (2000) "'Truth or Fiction": Problems of Validity and Authenticity in Narratives of Action Research', paper given at the Philosophy of Education Seminar European Education Research Association Conference, Edinburgh, Sept.
- Winter, R. (2002) 'Managers, Spectators and Citizens: Where does "Theory" Come from in Action Research?', in C. Day, J. Elliott, B. Somekh and R. Winter, *Theory and Practice in Action Research: Some International Perspectives*, pp. 27–44. Oxford: Symposium Books.
- Ziglio, E. (1996) 'The Delphi Method and its Contribution to Decision-Making: Gazing into the Oracle', in M. Adler and E. Ziglio (eds) *The Delphi Method and its Application to Social Policy and Public Health*, pp. 3–33. London: Jessica Kingsley Publishers.

TINA COOK is a reader at Northumbria University. She specializes in participatory enquiry working particularly with community-based projects and people/children with learning difficulties. She also works with an NHS Trust developing qualitative research. Please address correspondence to: School of HCES, H Block, Coach Lane Campus East, Northumbria University, Benton, Newcastle upon Tyne NE7 7XA, UK. [email: tina.cook@unn.ac.uk]