Research Papers in Education
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713707783

Assessing the quality of action research

To cite this Article: Elliott, John , 'Assessing the quality of action research', Research Papers in Education, 22:2, 229 - 246
To link to this article: DOI: 10.1080/02671520701296205
URL: http://dx.doi.org/10.1080/02671520701296205

Research Papers in Education
Vol. 22, No. 2, June 2007, pp. 229–246

Assessing the quality of action research

John Elliott*
University of East Anglia, UK

This contribution constitutes an attempt to distil quality criteria for assessing educational action research from examples of action research that the author has directly experienced in the course of his career. As such it effects a reconciliation between the idea of quality as an object of direct and immediate experience and as an object of more distanced criterial thinking and measurement. The author grounds the criteria in a number of narratives of experience or vignettes. In doing so he demonstrates how quality criteria are bound to contexts and multifaceted. The author clarifies the tradition of educational action research that has shaped his work with teachers and to which that work itself has contributed; namely one that stemmed from Stenhouse's notion of 'teachers as researchers'. He then goes on to illustrate three different kinds of teacher research he has engaged with—externally mediated, experimental teaching, and networked learning communities—and how quality shapes up rather differently in each. Each kind of teacher research will need to be judged in its own terms even when it shares certain common features. In the final section, Furlong and Oancea's domains of quality for applied and practice-based research are addressed. The author argues that each set of criteria distilled from his experience can be linked to Furlong and Oancea's dimensions of quality (with the exception of the economic dimension). Each expresses a concern for *theoretical and methodological robustness, value-for-use* (by teachers), and *building capacity* amongst teachers as potential agents of educationally worthwhile change. These universal dimensions of quality may be discerned in the formulation of the criteria illustrated in the vignettes. However, he contends that Furlong and Oancea's dimensions are too abstract and distanced from concrete experiences of action research to serve as meaningful criteria for judgement. Nevertheless, he claims that they can offer a broad orientation to reflecting about quality-as-experienced. This is illustrated by the author discussing the *value-for-use* dimension in the light of his experience of a particular action research project aimed at overcoming disaffection from learning.

Keywords: *Assessing action research; Experimental teaching; Networked learning communities; Quality criteria; Rigour; Value-for-use*

*Centre for Applied Research in Education, School of Education and Lifelong Learning, University of East Anglia, Norwich NR4 7TJ, UK. Email: John.Elliott@uea.ac.uk

ISSN 0267-1522 (print)/ISSN 1470-1146 (online)/07/020229–18

DOI: 10.1080/02671520701296205

Introduction

Stake and Schwandt (2006, pp. 404–418) point out two views of *quality* that are often seen as standing in tension with each other. On the one hand there is the view of *quality-as-experienced* and on the other the view of *quality-as-measured*. The former implies that the discernment of quality is a form of practically embodied knowledge—'at once both cognitive and emotional'—that is acquired in the course of immediate and direct experience of practical situations and events and manifest in the actions and language of participants. On this view, the evaluation of quality takes the form of 'experience-near understandings' that involve grasping 'the subjective and intersubjective meanings' that the evaluand attaches to 'events, personal encounters and places' and their 'sensitivities to virtue and trauma …'. Under these conditions *quality* is represented through narratives of personal experience.

The view of *quality-as-measured* involves an approach to evaluation that involves a *distancing-from-experience*. From this standpoint discernments of quality involve 'explicit comparison of the object in question with a set of standards for it'. Stake and Schwandt point out that from a *quality-as-measured* perspective the meaning of quality 'is structured … by a set of constructs' that tend to be derived not so much from the actions and language of the evaluand, as from the communities of discourse to which the evaluator belongs. The need to develop such standards stems, they suggest, from confrontational situations where 'few people are willing to accept personal perceptions of quality from opponents'.

However, Stake and Schwandt point at the tendency in practice, for criterial thinking to 'reduce the number of views of what quality is' in pursuit of a composite score that all evaluators might agree with. The more judgements of quality are reduced to a single measure, the greater the distancing from quality-as-experienced. This may secure agreement in the judgements of evaluators but does so at the expense of quality-as-experienced.

Such agreement, they argue, does not necessarily represent 'the most valid meanings' of quality as it is commonly and universally experienced. Criterial thinking needs to be rooted in narratives of experience. When it is so rooted, the number of quality criteria will tend to increase since quality-as-experienced is always 'multifaceted, contested, and never fully representable'.

Stake and Schwandt make a case for effecting a better balance between *intimacy* and *distance* in evaluating quality. This in my view lies at the heart of the challenge that the editors of this volume have presented to its contributors. In the light of Stake and Schwandt's distinction I will recast the challenge of the editors in the following terms:

- *How should our practical experience as applied researchers inform criterial thinking about the quality of our research when it is called to account?*

I have been asked to specifically focus on 'action research' as a form of applied research in education, and will do so in relation to the study of teaching and learning in classrooms where most of my experience lies.

The action research I have been engaged with stems from Stenhouse's notion of 'teachers-as-researchers' (1975, chapter 10), and focuses on ways of making teaching and learning more intrinsically *educational.* It has been part of a specific curriculum discourse that views the aims of education to refer to 'goods' that are internal to the practice of education, rather than to certain extrinsic purposes—social and economic—that education may serve (see Stenhouse, 1975, chapter 7). These may also be regarded as 'goods', but from the standpoint of a specifically educational discourse they should not be maximized at the expense of engaging students in activities that are intrinsically *educational*. From this standpoint educational action research is an ethical inquiry into the ways educational aims and values can find practical expression in the activities of teaching and learning. It engages teachers and their collaborators in a form of practical reasoning that Aristotle called *phronesis,* where the ends that constitute the internal goods of a practice and the means of realizing them in action are objects of joint reflection and inquiry (see Carr, 2006). In this form of action research, teachers develop their understanding of what constitutes *educational action* by reflecting about their actions in the light of their aims and their aims in the light of their actions.

In this contribution I will not be addressing quality issues that arise in relation to a mode of action research that is governed by another logic of practical reasoning, what Aristotle called *techne* (see Carr, 2006).[1] This kind of action research is concerned to shape teaching and learning as a means of maximizing the production of human capital to serve purposes that are external to education as such. I don't wish to deny a space for this kind of action research in classrooms and schools. However, given the increasing dominance of curriculum and educational reforms during the last 40 years that have been driven by the logic of technical rationality, I have found myself engaged in a mode of action research that emerged as an attempt to reclaim a space for *education* in classrooms.

In what follows I shall offer a number of vignettes that are illustrative of quality issues, as I have experienced them in different contexts of *educational* action research. Following each vignette I will reflect about the criteria that might validly inform assessments of the quality of the action research depicted. These reflections have been largely influenced by philosophical ideas that I will only briefly indicate in passing, since I have recently published fairly elaborate accounts of how these ideas have shaped my thinking about educational research (see Elliott, 2006a, b). As I proceed I will explore the extent to which quality in action research is context dependent, and argue that although there are criteria that can be validly applied across a range of different action research contexts there are some that are more context specific.

Externally mediated action research

The notion of 'teachers as researchers' was introduced into the UK during the late 1960s by Lawrence Stenhouse, who directed the Nuffield/Schools Council Humanities Curriculum Project (1967–1972). HCP was part of a widespread curriculum reform movement that involved a mixture of school-based initiatives and national projects.

Stenhouse (1970) recast the contribution of the humanities subjects to the education of 'young school leavers' (14 to 16 years old) in terms of the study of value issues in society. He argued that this required a pedagogical transformation at the level of the classroom from a pedagogy that consisted of the direct transmission of factual information by teachers to a process of discussion chaired by a procedurally neutral teacher. The teacher was responsible for 'discipline in discussion' by introducing relevant content as discussion unfolded, and ensuring that different interpretations of its meaning and significance were respected and treated in an open minded and reflective manner. Content was to be treated as 'evidence' for an inquiry aimed at 'developing an understanding of human acts, social situations and the controversial issues they raised in society'. Such evidence was drawn from a variety of genres, such as poetry and literature, film, documentary evidence, journalism, historical accounts, and research evidence from the behavioural sciences.

Stenhouse believed that the traditional culture of teaching and learning in schools could only be transformed in ways that were consistent with the aim and pedagogical principles of HCP, if teachers adopted a research stance towards their teaching (see Elliott, 1983). One of my roles on the project was to facilitate such a stance amongst teachers in our 'experimental schools'.

Vignette 1. The problem of information handling in classrooms

I was called into a school where the students were apparently unable to discuss the 'evidence' from the 'war and society' materials (which I had edited). The teachers had come to the conclusion that this was because it was above their reading level but wanted me to see for myself. I observed a lesson in which students were asked to discuss some information. They remained silent and the teacher, presuming that they had difficulties understanding it, went into instructional mode. He explicated the meaning of terms he believed students might have difficulty comprehending, administered a question and answer test, and then renewed his request that they discuss the information before them. They remained silent.

At the end of the lesson I interviewed a sample of students selected by the teacher, but in his absence. I asked them why they remained silent when invited to discuss, and they all gave the same explanation. They didn't like the material, not because they didn't understand it, but because they disagreed with the information it contained. I replied 'That's great so you can discuss it then?' They looked at me as if I were a novice to life in classrooms. One boy explained, 'If teachers give you things to read in lessons they expect you to agree with what they say'. He was explicating a cultural rule governing the use of information in classrooms that he and his peers had learned over their school careers.

I gained the permission of the students to play a voice recording of the interview back to the teacher. He became aware that he and his colleagues had misdiagnosed the problem of 'silence'. The teachers came to recognize that changes in classroom practice have a normative basis and require them to discuss and negotiate the new norms with their students.

This vignette describes a process of action research in which a member of a project team helps a group of teachers to gather evidence around a practical problem that had arisen in the context of their attempt to effect change in line with the aim and principles of the project. This evidence consisted of data that provided a view of the problem from three different angles: those of the teacher, students, and an observer (triangulation).

The vignette illustrates what the gathering of 'triangulation' data did in the context of the action research. It challenged teachers' stock diagnoses of classroom situations and in doing so led them to extend their *situational understanding* in ways that opened up new possibilities for action.

The student data particularly challenged what Becker (1998, pp. 90–91) calls 'the hierarchy of credibility'. He is referring to the way 'knowledge' in organizations like schools is hierarchically structured. Certainly teachers tend to be regarded as having more credible knowledge about what goes on in classrooms than their students. Therefore they tend not to be asked for their views on problems that teachers experience in teaching them. In this context my interview with the students might be regarded as potentially subversive to the maintenance of the hierarchy of credibility. Yet the data it provided were critical for the extension of their teacher's understanding of the problem he experienced, and for opening up some resolution to it.

How would I assess the quality of the action research depicted in this vignette? I certainly experienced it as a good piece of action research in the context depicted. What made it so?

1. It focuses on a problem that is of practical concern to the teachers involved.
2. It involves a gathering of data from the different points of view of the teacher, an observer, and students (triangulation).
3. It enables teachers to call their existing stock of professional knowledge (tacit theories) into question, and to test it against evidence gathered in their practical situation.
4. It extends teachers' understanding of their situation in a way that opens up new possibilities for action.

If these four criteria are taken out of the context in which the action research process was forged, and elevated to the status of universal quality criteria for all action research, they would only partially capture quality as I have experienced it in other action research contexts. The quality of the action research depicted in the vignette is dependent on an external researcher, who takes control of the process of data gathering and interpretation. Although such external control may enhance the quality of action research in circumstances where 'the hierarchy of credibility' makes it difficult for teachers to call their existing stock of professional knowledge into question, it will not be necessary in a context where teachers have become capable of sustaining an *experimental stance* towards their teaching. Here the locus of control over the action research process shifts to the teacher. In this context other aspects of quality in action research become apparent. I have come to regard the kind of externally mediated

action research that is depicted above as a stage many teachers will have to go through in developing a capability for experimental teaching.

Experimental teaching

In *experimental teaching* the teacher controls a data gathering process designed to provide information about the extent to which her actions and their effects on students are *ethically* consistent with her pedagogical aims and principles. She seeks to explain inconsistencies in terms of constraints operating on her practice that stem from both the system and a wider social context. New insights into her situation lead the teacher to experiment with new action strategies that in turn change the situation sufficiently to reveal previously unrecognized dimensions of the problem that need to be understood. In experimental teaching *situational understanding* develops interactively with action to effect improvements in the situation.

The boundary between research and teaching is dissolved in experimental teaching. Data gathering becomes an integral part of teaching, and the teacher is the locus of control with respect to the links between research and action. The latter need not imply that there is no role in the process for external researchers. They can provide intellectual resources and skills at gathering and interpreting data, which a teacher engaged in experimental teaching might be wise to call upon. In doing so the teacher does not relinquish control over the research process in her classroom, but rather shapes it around her information needs as they emerge over time.

The kind of action research depicted above will tend to involve less formal methods and procedures for gathering and interpreting data than those depicted in the earlier vignette. For example, the teacher-as-researcher is more likely to directly engage their students in ongoing conversations about how to develop their teaching as and when the need arises, rather than depending on formal interviews or surveys carried out at pre-planned intervals. Students then become active partners in the action research process as opposed to being cast in the role of mere research subjects.

Vignette 2. Gendered constraints on freedom of discussion (see Elliott, 1974)

I spent a period during the course of HCP in a school teaching with the 'war and society' materials alongside a number of permanent staff. My group consisted of seven boys and nine girls. I voice-recorded 'the discussions' with them and gave the other teachers opportunities to discuss the recordings with me. In this way I hoped to build their capacity for undertaking action research.

In the first session I passed around about 50 photographs of situations that were relevant to the topic, and asked each student to select two that were of interest to them. I then invited students to volunteer the reasons for their choice of photographs and the rest of the group to comment on these reasons. In this way I hoped that certain issues would emerge to form an agenda for discussions in the subsequent sessions.

In this first session there was a high degree of participation amongst the boys, but less evidence of participation amongst the girls. At the end of the session I asked the students what they thought of the discussion. One girl remarked: 'They [other girls] just sat and listened; they didn't feel so strongly that they would say anything. No they are too shy—'.

After the session I made a note: 'The problem of non-participation amongst girls was explored. There was some hostility from Roy [a student] towards these girls [those that contributed] and I sensed some hostility from them towards him'.

It was beginning to dawn on me that in trying to wean the students away from dependence on the traditional authority position of the teacher I had also exposed a highly gendered structure of power relations that shaped students' freedom of opportunity to participate in discussions. This situation was normally masked by the dominance of the teacher in the instruction-based classroom.

The HCP teachers in the school did not see the girls' non-participation as such a problem. From their point of view it was a normal situation. Girls are 'naturally' shy about discussing their opinions in front of boys, particularly on a topic that might be seen as the preserve of boys. The dominance of the boys was simply a natural state of affairs that one had to accept. During the next three sessions I tried a variety of different strategies to encourage the girls to participate in discussion, including giving them space to talk about the reasons why they selected their photographs, and to identify themes and issues that were of concern to them as girls. At times the discussion sparked to life as boys and girls began to share their different points of view, but the process was not sustained.

At one particularly low point during the fourth session I decided to abandon the discussion and asked the class, especially the girls, to analyse the problems they were experiencing. They were prepared to talk about these and there came a point when I asked them which of the characters in the extract from a novel about the war in Vietnam they had most identified with. This led to a most interesting discussion amongst the girls about the relationship between a professional soldier and a raw recruit, in which they expressed very different views. It now appeared that far more students, including the girls, had sympathized with the professional soldier than was apparent in the earlier discussion.

Why had the expression of individual differences of view emerged in the context of 'talking about the discussion' rather than during it? One possible explanation is that the norms governing 'discussions' in mixed-sex classrooms do not apply at the level of 'talking about discussion'. They only apply to work, and talking about the problems of discussion is not considered to be 'work'.

I eventually refocused the discussion back on the problem students were having with discussing the topic. One girl, Eileen, remarked that she had talked too much and I asked the rest of the group whether they thought she had. One girl said she had, and I remarked 'Took the responsibility off you, off everyone else for speaking—'. A chorus of 'Yeah, yeah—' broke out in response.

From the meta-level discussion it became clear that the group was beginning to split along completely new lines. There was a subgroup of boys and girls, consisting

of just over half of the class, who continued to cling to the security of expressing their views in very gendered terms, and a subgroup that were tentatively beginning to value the expression of divergent points of view by individuals within the group, irrespective of their gender.

The HCP teachers in the school offered information that extended my understanding of the situation further. The quiet girls were 'country girls', while the dominant girls came from military families and were only temporarily resident in the area. They were widely travelled and as a result had developed a broader and more liberal outlook on life than 'the country girls', who tended to have a more traditional view of gender relations.

At the start of the fifth session I introduced a poem that was greeted with a wall of silence. Thinking that the material might be part of the problem I then introduced some photographs, but to little effect. I abandoned the discussion of materials and suggested once again that we moved into a discussion of the problems we were having with discussion. The discussion once again began to flow freely.

The normally dominant girls explained their silence as an attempt to force the passive girls to participate. I questioned the wisdom of this strategy on the grounds that it would be counter-productive by further increasing feelings of resentment towards them. The dominant girls and some of the boys began to analyse the problem of non-participation in some depth rather than seek easy answers. They talked about the fear many students have of 'making mistakes' in front of the teacher and their peers, and their lack of confidence in their own opinions and judgements.

For the sixth and final session I introduced an extract from Jungk's *Children of the ashes*, depicting the human misery and suffering that followed the dropping of the A-bomb on Hiroshima. Not one of the girls participated in the discussion. I asked the girls who normally participated why they had opted out. They accused me of allowing the normally passive girls to opt out. One of them then spoke on behalf of the others and argued that 'we didn't want to talk, and he [me] asked us "Why?" but didn't ask them [the passive girls]'. To her I was being unfair. I suggested that she and her friends normally enjoyed talking, while the other girls did not. She retorted 'But why should we talk this morning?' I was being accused of bullying and charged with inconsistency of treatment.

The silence of the normally dominant group of girls at the start of the session was testing the extent to which I was prepared to allow them, as well as the normally passive girls, the freedom to remain silent. In their eyes I failed the test. I had seen myself as 'encouraging' rather than 'forcing' their participation and protecting their 'rights' when threatened by the boys. What I had underestimated was the extent to which my interventions would arouse considerable guilt amongst these girls, by reinforcing a degree of participation that came perilously close to being 'unfeminine'.

I had not realized that the willingness of the less passive girls to participate in discussion was itself governed by norms of participation that stemmed from their gender identities. These girls wished to contribute but not at the expense of appearing to the other girls to be behaving like dominant males. They wanted to share the work

of discussion more equitably with the other girls. Their construct of 'freedom to express their views' was of *freedom within limits.* From their point of view I failed to acknowledge those limits.

Although I assumed that I was construing my pedagogical aims and values in gender free terms I now realize in retrospect that I was operating with a masculine construct of freedom. I was construing 'freedom of thought' as the expression of a purely personal and individualized 'self' freed from social limitation. The data that emerged from the students' feedback on my interventions challenged this construct but I failed to recognize the challenge until it was too late. Had I done so earlier I may have discerned possibilities for action that would have resolved the situation more satisfactorily for both my students and myself.

What criteria for assessing the quality of *experimental teaching* might be distilled from the experience depicted in the vignette above? The vignette highlights the importance of certain aspects of quality in educational action research in addition to those specified earlier. These are linked to the fact that in the context of experimental teaching it is the teacher who is the locus of control. I would therefore suggest that the following criteria are also relevant to the assessment of action research when it takes the form of experimental teaching:

5. It is a deliberative and self-reflexive process in which the teacher calls into question both her teaching strategies (means) and the aims (ends) to which they are directed, and then modifies each by reflecting on the other.
6. It is a rigorous conversational process in which the teacher opens up her practice to the rational scrutiny of students and peers, 'in-voices' their views of the action situation, and in the process demonstrates a disposition to subordinate her own prejudices to the search for an overlapping and un-coerced consensus.
7. It is a process in which the teacher displays:
   Integrity in the pursuit of her educational aims and values.
   Curiosity about other people's interpretations of the action situation.
   Objectivity and honesty about her own motives and reasons for action.
   Open-mindedness towards the views of others and respect for their freedom of thought and action.
8. It enlarges the teacher's sphere of personal agency in the practical situation through the realization of her educational aims in a sustainable form.

Criteria 5–8 refer specifically to the personal qualities manifested in good experimental teaching. Criterion 5 captures the quality of reflection that Aristotle called *phronesis* (see Nussbaum, 1990, p. 62), and which is reflected in Stenhouse's notion of 'teachers-as-researchers'.

The democratic conception of rigour cited in Criterion 6 is consistent with the view of the pragmatist philosopher Richard Rorty. He argues that there are no methodological constraints on inquiry 'derived from the nature of objects, or of the mind, or of language'. The only constraints are conversational ones, 'those retail constraints provided by the remarks of our fellow inquirers'. Those of us engaged in inquiry 'have a duty to talk to each other, to converse about our views of the world,

to use persuasion rather than force, to be tolerant of diversity, to be contritely fallibist' (1991, p. 67). Rigour, Rorty argues, is something 'you can have only after entering into an agreement with some other people to subordinate your imagination to their consensus' (1998, p. 339). Rorty's views on the nature of inquiry are continuous in many respects with Dewey's (1916, pp. 173–179) account of scientific method. Dewey depicted the kind of dispositional qualities cited in Criterion 7 as virtues inherent in any democratic process of inquiry.

Criterion 8 highlights a quality that is central to the notion of experimental teaching. Such teaching reclaims the teaching situation as the sphere of ethically committed action or *praxis*. It is in this sphere that the teacher experiences her 'self' as an agent of change. The Scottish philosopher John Macmurray (1957, p. 90) argues that a person experiences 'the self' most fully in action directed towards the fulfilment of practical goals.

How should the piece of experimental teaching depicted in the vignette be assessed in the light of the above criteria? I presented the quality of my deliberation and reflection as at issue. It was only after the final session with the students that I began to reflect about the constructs that shaped my ends-in-view in the light of the students' interpretations of my actions and their explanations for their reactions. This might be explained in terms of my lack of rigour in using conversational feedback from the students much earlier to question my own masculine construct of 'freedom in discussion'. A more rigorous engagement with that feedback would have called for a greater degree of open-mindedness and curiosity on my part towards the views expressed by students. Finally, it might be argued that this piece of experimental teaching did not significantly enlarge my sense of agency in the action situation. Although I did not entirely resolve the situation to my satisfaction, the evidence presented in this vignette suggests that an increasing number of students began to transcend the negative aspects of their gendered identities to embrace the values inherent in discussion as a learning process.

The vignette is a shortened version of an even more detailed case study that was published in both an educational journal (see Elliott, 1974, pp. 147–155) and an early number of the feminist journal *Spare Rib*, where it attracted the interest of some feminist researchers. The latter's critique of the case study made me more self-reflexive about my own actions in the situation and influenced the construction of the vignette. This raises the question of how a single context-bound case study can provide a focus for general conversation and debate. My experience suggests that many single case studies written by teachers have great power to illuminate the problems and dilemmas experienced by other teachers in their particular contexts of practice. It is an experience that resonates with Stake's notion of *naturalistic generalization* (1978, pp. 5–7), in which the reader is able to generalize from the case to her own experience in a way that illuminates her own experience. I would argue that the greater the particularizations of descriptions of action situations, the greater their potential to throw light on possibilities for action in other situations.

In the light of this claim I will suggest another criterion for assessing the quality of experimental teaching; namely, that:

9. It enables a teacher to generate a description of the complexities of the case in sufficient detail to be of universal significance to other teachers.

I now wish to set my experience of action research in the wider context of professional learning communities.

Collaborative action research across networked professional learning communities

Much of my experience of action research has been in the context of establishing learning communities in the educational system, consisting of teachers as core members who come together on a voluntary basis to focus on problems they experience in common. All involved cross-curricular groups of teachers operating inside their schools under the leadership of a senior member of staff appointed by the head teacher. These groups met regularly with similar groups from other participating schools on a regular basis to share and discuss each other's data. The cross-school meetings were coordinated by an external team of educationalists and researchers, who also provided research support at the classroom level. The following vignette is illustrative of such collaborative action research.

Vignette 3. Inquiry/discovery teaching (1972–1974)

The Ford T Project (see Elliott 1976/77) brought together teachers engaged in curriculum development at different phases of education, and within different subject areas.

The external team of the project began by bringing the participating teachers together to focus on examples of teaching and learning, captured on audio and audio-visual recordings, where the teacher claimed to be using an inquiry/discovery approach. The teachers were asked to discuss which approaches could be described as forms of inquiry/discovery teaching and which could not. These discussions revealed considerable disagreement amongst the teachers on this matter, and from them we were able to distil a range of pedagogical theories that different teachers used to judge whether an approach counted as a form of inquiry/discovery teaching. Subsequently the teachers were invited to gather evidence about their own teaching approach in the light of these practical theories, and how they influenced the learning process. In effect we asked them to test the pedagogical theory embedded in their practice (see Criterion 3, above).

In spite of disagreements the initial conference enabled the teachers to reflect about the kind of learning process they were trying to realize in their classrooms. Reflection about pedagogical means stimulated reflection about their aims. In this respect an overlapping consensus emerged concerning the defining characteristic of inquiry/discovery learning. What most characterizes it they agreed is the scope it allows for 'self-direction' or 'independent thinking' on the part of the learner. In such a process learners experience the freedom to do their own thinking, as opposed to depending

on the teacher as an indubitable source of knowledge. This consensus provided teachers with a criterion against which to evaluate their pedagogical strategies in the classroom and test the practical theories embedded in them.

In order further to assist with the testing of practical theories, I attempted, in line with my HCP experience, to clarify the pedagogical principles (principles of procedure) implicit in the teachers' conception of their end-in-view. The framework of pedagogical values and principles that emerged from this exercise was subjected to further discussion with the teachers, and an acceptable version negotiated (see Elliott, 2007, pp. 41–42). It provided criteria that they could use to test and develop their practical theories of inquiry/discovery learning through action research. At issue here was the appropriate relationship between *structure*, *guidance*, and *independent reasoning* in the teaching and learning process.

The findings that emerged from Ford T reflected an increasing consensus, based on evidence, that *structured teaching* was counter-productive when it involved setting short time-scales for learners to achieve the desired outcomes. When longer-term outcomes were set the appropriate degree of structuring appeared to be in less conflict with the aim of fostering independent thinking. An emerging agreement also surrounded the issue of whether the teaching strategies should be simply *open-ended* or provide *positive guidance* to the learner. The evidence demonstrated that some teachers were able to provide positive guidance in ways that enhanced rather than diminished students' capabilities to engage in independent thought.

During the first term of the project only twelve of the forty teachers who volunteered had shown any capacity for experimental teaching. Eight teachers dropped out in that first term. However, with the agreement of the teachers we used the data we gathered in their classrooms as a basis for discussion at local and central project meetings. This activated more teachers to accept responsibility for the research. Some developed collaborative projects in their school-based teams while others did their own individual research and discussed their data with team members. By the end of the two years of the project's lifetime we estimated that twenty-five teachers had become capable of self-initiated action-research. This did not mean that we gave no help with data gathering and interpretation. Unlike the first term of the project, teachers increasingly asked us to help them gather data. However, they were able to specify the kind of questions they needed us to address in the process and the kinds of data they wanted. They also wanted us to discuss the data with them but were happy to have the final say in how it was to be interpreted.

What quality criteria, in addition to those listed earlier for experimental teaching, might be distilled from this experience of coordinating action research in the context of a networked professional learning community? The following criteria pick out aspects of quality that are manifest in good examples of action research where teachers and others collaborate to construct knowledge together about educational action. As such they highlight additional dimensions of quality in action research to those already specified in relation to the first two vignettes.

Good collaborative action research should:

10. Exemplify a democratic process in which teachers submit their actions and reasons for actions to the rational scrutiny of their professional peers, and demonstrate a willingness to modify their actions in the light of an emerging consensus about their educational significance.
11. Foster the development of experimental teaching across the learning network in a sustainable form.
12. Enable teachers to discern common features across a range and variety of teaching situations that are practically relevant for educational action.
13. Enable teachers collectively to construct knowledge about how to realize their educational aims and values in particular concrete situations.
14. Enable teachers to develop shared understandings of their educational aims and values by deliberating together on the actions they take to give them practical expression in their teaching.
15. Make a significant contribution to the development of a tradition of understandings—a theory—about how to bring about worthwhile educational change across contexts of teaching and learning, and thereby extend the knowledge-base of the teaching profession.
16. Enable teachers systematically to present a body of shared understandings and insights in a publicly accessible form for other teachers to test in their contexts of practice.

These criteria presume that a 'community of practice' has the capability to discern similar patterns and regularities across particular contexts of action. The view that the findings of action research cannot be generalized because they stem from small-scale case studies of action in particular situations tends to rest on the presumption that *generalizability* is a matter of aggregating statistical data to yield *context free* findings. In the Ford Teaching Project teachers gave each other access to their case studies and/or case data for the purposes of discussion. Through such discussions teachers were able to discern commonalities in their experience, and represent them as action research findings (see Ebbutt & Elliott, 1985, for a further example of this kind of action research).

The diagnostic and action hypotheses developed in the contexts of the Humanities Curriculum and Ford Teaching Projects can be regarded as the kind of useful summaries of insight and judgement that Nussbaum (1990, pp. 67–68) calls 'universal principles' as opposed to 'general principles'. Their function is to guide further experimentation—reflection and action—rather than prescribe (see Elliott, 1976/77, pp. 15–18; 1983, pp. 114–116).

Action research and its 'value-for-use'

The dimensions and sub-dimensions of quality in applied and practice-based educational research cited by Furlong and Oancea (2006), with the exception of the economic dimension, are reflected in the criteria that I have distilled from my experience in each of the contexts depicted in V1, V2, and V3. Each set of criteria expresses

a concern for *theoretical and methodological robustness, value-for-use* (by teachers), and *building capacity* amongst teachers as potential agents of educationally worthwhile change.

Furlong and Oancea's dimensions and sub-dimensions of quality cover a variety of genres that might fall under the general description of 'applied and/or practice-based educational research'. Therefore, as descriptions of quality they tend to operate at a more abstract level than the sets of criteria that I have distilled from my experience of action research. They do however offer a broad orientation to reflecting about quality-as-experienced within a particular research genre. I now want to illustrate this with respect to Furlong and Oancea's dimension of 'value-in-use' (Section 6.2).

Vignette 4. Researching student disaffection from learning (1997–2000)

The Norwich Area Schools Consortium (NASC) was organized as a research consortium and consisted of secondary schools in the Norwich area, a team of educational researchers based at the University of East Anglia, and officers from the Local Education Authority. It was funded by the Teacher Training Agency (now the TDA) and aspired to engage teachers *in* researching the 'curriculum and pedagogical dimensions of disaffection from learning'.

The project (see Elliott, 2004, pp. 264–290) proceeded through two phases. In the first phase volunteer teachers in schools, either individually or in small groups, planned and carried out small-scale pieces of research into issues relating to disaffection that they wished to investigate. These pieces of research were written up, circulated and discussed within the consortium as whole. At a midway point in the life of the project the teachers came together to generate themes from their case studies and research reports, as a basis for research across all the six (originally seven) schools in the consortium. Four main themes were identified and these provided the focus for action research in the second phase.

Many of the phase 1 case studies and reports—13 were produced by the end of this phase—depicted 'student disaffection' as a difficult phenomenon to identify, describe and explain. The teachers' experience of the actual 'value-for-use' of their research was often couched in terms of it helping them to ask better questions in their search to understand 'disaffection' as a phenomenon in their classrooms. The evidence they found most powerful in enabling them to do this invariably came from talking with children.

> The amount of insight we got … in half an hour. … You would think 'Well, this person is underachieving, why …?' They were shy initially, but once you hit upon the key issue. The key issues were different for everyone … we got loads and loads of stuff… lesson observations have been good but all that has been able to confirm is 'Yes, this is a (quietly) disaffected child and they behave in this way'. The patterns in which they behave are quite similar … the reasons behind the behaviour are all unique. That's what you need to know. You can identify the fact that the child is under-achieving but (need) to find out why. (See Elliott, 2004, p. 275)

Many teachers experienced this search for understanding as an unending quest that prevented them from generating a tidy set of findings about how to deal with disaffected students. Yet they came to feel at ease with it in terms of its 'value-for-use'.

However, within the TTA, and the education policy community more widely, there was a tendency to assume that the questions that need to be addressed by teachers can be clearly defined in advance of the research on the basis of existing data: about attainment levels in relation to national curriculum targets, about truancy and absenteeism, about exclusions and persistently disruptive behaviour in classrooms. All these data it was assumed provided indicators of disaffection from learning. Having identified levels of disaffection in these terms it was assumed that teachers are in a position to ask clear questions as a basis for action research in their classrooms, e.g., 'what strategies work to raise the attainment levels of under-achievers, or to prevent disruptive behaviour in classrooms, or to reduce absenteeism from lessons?' Underpinning such questions is an instrumentalist and objectivist rationality that renders action research a mode of technical reasoning. 'Value-for-use' in this context is a matter of gathering evidence about the most effective means of bringing about some measurable improvements that can be pre-specified in advance of the research. Such a view of teacher research is very different from the mode of action research that is illustrated in the vignettes.

The NASC teachers' experience of the potential 'value-for-use' of their action research can be linked to Criterion 3, inasmuch as it implies that the evidence that teachers gather should enable them to ask better questions in seeking to understand the complexities of the pedagogical problems they experience. Such questions are those that have the potential to challenge existing conceptions of the problem wherever they prevail, e.g., within the teaching profession itself, or the policy community, or the community of educational researchers.

The distinctions Furlong and Oancea draw in Section 6.2 of their report, between *short-term* and *long-term* impact and *actual* and *potential* impact, are important ones. The kind of evidence that many NASC teachers gathered during Phase 1 of the project has potential value-for-use beyond the confines of the project. During phase 1 of the project we had two striking examples of case studies that had potential 'value-for-use'. One was a piece of collaborative action research between my colleague Kim Brown and teacher Alan Fletcher, about the problem of disruptive behaviour in his classroom. From their research they produced a case study entitled *Disaffection or disruptive engagement?* In this study they gathered evidence that called into question the assumption that disruptive students are disaffected from learning. In summarizing their 'findings' they write:

> The blanket description of pupils as disaffected is challenged by the findings in this study. What emerges from the interviews with pupils and an analysis of the video recordings of a lesson is a sense of the potential that these pupils have to engage with their language learning. It appears that, at some level, there is a commitment on their part to engage with the teacher and the lesson he has prepared for them, but in the process of doing this, the pupils construct their own obstacles to success. It has been suggested (in this study) that poor social skills and individual needs for attention lead

> to what we have termed their disruptive engagement in their language lessons. (Brown & Fletcher, 2002, p. 189)

The question that emerged from this piece of action research has not, I believe, been captured in any of the published research into disaffection. Most of this research appears to assume, or at least does not challenge, the view that disruptive behaviour in classrooms is an indicator of disaffection from learning. This assumption is also prevalent in the policy community (see Elliott, 2000).

I am unaware of any actual impact that Brown and Fletcher's research has had on the way either educational researchers or members of the policy community think about the phenomenon of disaffection in our schools. Nevertheless, I would still judge the quality of this case study as having potential value-for-use in these contexts. I will now turn to my second example of NASC research that poses a challenging question.

In one of the schools, a group of teachers and the chair of the governing body embarked on a study of students they believed might be under-achieving (see Oakley, 2002). They showed no signs of visible disaffection regarding the normal indicators of disruptive behaviour and teachers' formal assessments of students' attainment. The evidence for their belief rested on a form tutor's observations of interactions between teachers and certain parents at a consultation event. A number of students from the teachers' point of view appeared to be making satisfactory progress whereas from the parents' point of view they were under-achieving. Follow-up interviews that the form tutor held with teachers, about how they made their annual assessments, suggested that their recorded assessments of students' attainments in relation to their potential could not be relied upon as an indicator of under-achievement. The observations at a parent consultation event and the interviews with teachers suggested the possibility that there were significant numbers of students who were under-achieving and quietly disaffected. Such students were called RHINOs (Really Here In Name Only). The problem arose as to how they might be identified, given their apparent invisibility to teachers in the classroom, and on the basis of what kind of evidence?

At this stage the form tutor (Oakley) and his colleagues with the chair of governors embarked on an observational study of a sample of students that figured in the conflicting interchanges between teachers and parents, and they combined them with interviews with the students themselves. Through the research they gathered evidence about the strategies the students employed to render themselves invisible to their teachers. The evidence suggested that the existence of RHINOs tended to be pedagogically relative inasmuch as it appeared to be linked to poor learning environments. The teacher–researchers then worked with some teachers to improve the quality of the learning environment in their classrooms, and with parents to secure their collaboration in identifying and engaging quietly disaffected students in learning.

The RHINOs case study had the potential to get teachers, policy-makers and educational researchers to reframe their understanding of disaffection in classrooms and schools. However, unlike my first example this case study began to have an immediate impact in both the policy and educational research communities. Oakley was invited to present the study at a national conference of teachers convened by the

TTA, and a shortened version of it was published by the DfES (Department for Education and Skills). Moreover, it influenced the development of a successful research bid to the ESRC (Economic and Social Research Council) by my UEA colleague Elena Nardi on 'The invisible pupil in the mathematics classroom' (see Nardi & Steward, 2003, pp. 345–367).

These two examples illustrate the capacity of good case studies to provide 'value-for-use'. Both are covered by Criterion 9 above with respect to assessments of their potential value-for-use. Indeed I would suggest that additional criteria to the ones I have already listed are not required for assessments of the potential value-for-use of action research. The list as it stands may be viewed as a specification of that potential. Earlier I argued in relation to my first example that lack of evidence of actual impact to date does not warrant the conclusion that it lacks potential value-for-use. It may well have impact in the longer-term. In my view *actual impact,* other than that which is an intrinsic feature of good action research and already covered by the criteria listed, should not feature in any criterial specification of quality for either this or any other kind of research. This is not to say that positive evidence of impact should not be taken into account. It is indicative of value-for-use but not definitive, and therefore should not be over stressed. Impact beyond the context of the research cannot be predicted with sufficient certainty, even in the longer term. It depends on contingencies that often make it difficult for those engaged in the research to anticipate, let alone predict, impact. This is so even when the research explicitly aims, as with educational action research, to generate practically useful knowledge.

For example, I recently received the following email.

> I am working on a major project on the use on rewards/sanctions in the juvenile system in the US and the UK funded by the Office of Juvenile Justice and Delinquency Prevention here in the USA.
>
> I find the consortium's work on rewards and sanctions in the Norwich schools fascinating and relevant to my own research. I see that a good bit of the consortium's thoughtful research has been published online and in journals. I plan to present what is publicly available at the British Society of Criminology's meeting during July in Glasgow in a paper titled, 'What juvenile justice researchers can learn from the Norwich School study'.

No-one involved in the NASC research had estimated, or perhaps could have estimated, the potential of this piece of action research, to have 'value-for-use' in the juvenile justice system.

Notes

1. Oancea and Furlong (this volume) provide a fuller discussion of Aristotle's concepts of both *phronesis* and *techne*.

References

Becker, H. (1998) *Tricks of the trade: how to think about your research while you're doing it* (Chicago, University of Chicago Press).

Brown, K. & Fletcher, A. (2002) Disaffection or disruptive engagement? A collaborative inquiry into pupils' behaviour and their perceptions of their learning in modern language lessons, in: J. Elliott & B. Zamorski (Eds) *Researching disaffection with teachers,* special issue of *Pedagogy, Culture and Society,* 10(2), 169–192.

Browning, L., Crowe, J., Ebbutt, D., Elliott, J., Hurlin, T., Jones, R., Pedlar, R. & Reay, T. (1974) *Implementing the principles of inquiry/discovery teaching* (Norwich, Centre for Applied Research in Education, University of East Anglia).

Carr, W. (in press) Philosophy, methodology and action research, *Journal of Philosophy of Education.*

Dewey, J. (1916) The nature of method, in: *Democracy and education* (New York, The Free Press).

Ebbutt, D. & Elliott, J. (Eds) (1985) *Issues in teaching for understanding* (London, Longman).

Elliott, J. (1974) Sex role constraints on freedom of discussion, *The New Era,* 55(6), 147–155.

Elliott, J. (1976/77) Developing hypotheses about classrooms from teachers' practical constructs, *Interchange,* 7(2), 2–22.

Elliott, J. (1983) A curriculum for the study of human affairs: the contribution of Lawrence Stenhouse, *Journal of Curriculum Studies,* 15(2), 105–123.

Elliott, J. (2004) Using research to improve practice: the notion of evidence-based practice, in: C. Day & J. Sachs (Eds) *International handbook on the continuing professional development of teachers* (Maidenhead, Open University Press).

Elliott, J. (2006a) Educational research as a form of democratic rationality, *Journal of Philosophy of Education,* 40(2), 169–185.

Elliott, J. (2007) *Reflecting where the action is: the selected works of John Elliott* (Abingdon, Routledge).

Furlong, J. & Oancea, A. (2006) Assessing quality in applied and practice-based research in education: a framework for discussion, *Review of Australian Research in Education: Counterpoints on the Quality and Impact of Educational Research,* 6, 89–104. Available online at: www.bera.ac.uk/pdfs/Qualitycriteria.pdf (accessed April 2006).

Macmurray, J. (1957) *The self as agent* (London, Faber & Faber).

Nardi, E. & Steward, S. (2003) Is mathematics T.I.R.E.D.? A profile of quiet disaffection in the secondary mathematics classroom, *British Educational Research Journal,* 29(3), 345–367.

Nussbaum, M. (1990) An Aristotelian conception of rationality, in: *Love's knowledge* (Oxford, Oxford University Press).

Oakley, J. assisted by Wind, C., Jones, D., Joseph, D. & Bethel, M. (2002) RHINOs: a research project about the quietly disaffected, in: J. Elliott & B. Zamorski (Eds) *Researching Disaffection with Teachers,* special issue of *Pedagogy, Culture and Society,* 10(2), 193–208.

Rorty, R. (1991) Pragmatism without method, in: *Objectivity, relativism and truth: philosophical papers. Volume 1* (Cambridge, Cambridge University Press).

Rorty, R. (1998) Derrida and the philosophical tradition, in: *Truth and progress, philosophical papers* (Cambridge, Cambridge University Press).

Stake, R. E. & Schwandt, T. A. (2006) On discerning quality in evaluation, in: I. Shaw, M. Mark & J. Greene (Eds) *Handbook of evaluation* (New York, Sage Publications).

Stenhouse, L. (1970) The humanities project: the rationale, *Theory into Practice,* 10, 154–162.

Stenhouse, L. (1975) *An introduction to curriculum research and development* (London, Heinemann Educational Books).