Assessing action-research projects within formal academic programmes: using Elliott’s context-related criteria to resolve the rigour versus flexibility dilemma

Tim McMahon
University College Dublin

Elaine Jefford
Newcastle University

Publication details
Published version available from: http://doi.org/10.1080/09650790903093250
Assessing Action-Research Projects within Formal Academic Programmes: using Elliott’s context-related criteria to resolve the rigour vs. flexibility dilemma.

Tim McMahon and Elaine Jefford

aThe Centre for Teaching and Academic Development, University College Dublin, Ireland; bPhD Candidate – Midwifery, Newcastle University, Australia.

An action-research account of a successful improvement to practice presented for assessment within a taught MA in Education was given a fail grade. The student’s challenge to this decision led the university assessor to re-evaluate and change his own practice in assessment in much the same way as the student had in the original report. The narrative of this event is used to explore the central dilemma of assessing action-research reports inside academic programmes, namely, the competing needs of the action-researcher to follow the investigation wherever it leads and the need for the student to meet pre-set criteria for assessment. Following an introduction, Part 2 of the paper presents an extract from the original action-research account while Part 3 presents the story of the assessment and notes how it exemplifies some of the key issues relating to the way action-research can and should be assessed within credentialed academic programmes. Part 4 explores these issues further and suggests that Elliot’s (2007) formulation of quality criteria can resolve the central dilemma of marking action-research reports within credentialed programmes of study.

Keywords: action-research, quality, rigour, academy, assessment.

Part 1: Introduction

Presenting action-research reports in part or whole fulfilment of the assessment requirements of a recognised academic qualification is inherently more risky than presenting more traditional projects. There are, at least, three reasons for this. Firstly, the boundaries of action-research – the technical rules – have never been, and probably can never be, clearly defined. Secondly, what constitutes quality in action research remains a matter of some considerable debate (Heer & Anderson 2005: especially Chapter 4, Feldman 2007, Elliot 2007). Thirdly, action-research, by its
nature, is much less predictable and predictive than other forms of research. Consequently, it is all too possible that the research project will diverge, in form or content, from the formal requirements of the assessment system. The student thus faces the dilemma of either risking the integrity of the research by constraining it within predetermined safe boundaries or risking failing to meet the requirements of the Academy by following the action-research cycle wherever it leads. As most students will, understandably, want to ensure that they achieve the qualification, the big danger is of academic rules stifling effective action-research. As Sanger (1996) put it

The aspiration of collaborative, self-critical enquiry on the part of practitioners is constantly eroded by the need to be credentialed or otherwise supported by academic centres whose own agendas often lead to hegemonies in research approaches and relationships.

(Sanger 1996, 182)

Those of us within the Academy charged with assessing such reports also face a dilemma peculiar to the action-research paradigm. As academicians we have to apply fixed, published criteria to students’ work in a way that is fair and equitable; but, as scholars who are, most probably, action-researchers ourselves, we must educate our students in the inevitable unpredictability and frequently establishment-challenging nature of action-research. Should we encourage our students to challenge everything and put the goal of improving practice first, or should we advise constraint and safety and thus become an agent of the “hegemonies” identified by Sanger? Further, when we assess, how can we reconcile any apparent divergence between our responsibility to promote effective action-research and our responsibility to defend the rigour and fairness of public qualifications?
Describing this dilemma as peculiar to action-research perhaps requires some explanation. It is, of course, true that other forms of “alternative” or “new paradigm” research can present seemingly similar problems for students and teachers. What makes action-research work presented by students unique, however, is that action-research deliberately mixes discourses in a way that erodes the boundaries between knowledge-generation and action (Somekh & Zeichner 2009: 6). While this makes it “…uniquely suited to generating and sustaining social transformation” (Somekh & Zeichner 2009, 6)

It also bestows a concomitant uniqueness to the dilemma of reconciling the generic standards of the Academy with the standards applied to judging action-research as research.

Despite a considerable body of carefully argued literature stretching from Guba and Lincoln (1981) through the seminal work of Carr and Kemmis (1989) to Elliot (2007), these questions remain a vexed and challenging part of the “ongoing conversation” identified by Herr & Andersen (2005) – a conversation typified by the debate between Heikkinen et al (2007) and Feldman (2007) in the pages of Educational Action Research. Elliott’s (2007) formulation (he would say distillation) of a set of quality criteria for assessing educational action-research, enables a resolution of the central dilemma that arises when reports which reflect the inevitable messy and unpredictable nature of action-research are presented for assessment within an academic programme where the maintenance of standards is paramount. The criteria that Elliot formulated are open enough to accommodate the contextual nature of value judgements but rigorous enough to enable such judgements to be tested against the universal
dimensions of quality formulated by Furlong and Oancea (2006). Elliott, however, states that his criteria should not, individually, be regarded as universally applicable to all action-research as they would not necessarily capture the full quality residing in research conducted in other contexts. Rather, he argues that within his distilled list will be found appropriate criteria which can be used to judge any given action-research project. This idea of selecting from a distilled list is also found in Winter, Griffiths and Green (2000) and would seem particularly appropriate for action research presented within the Academy. Taking this approach - bearing in mind the universality of the underlying dimensions formulated by Furlong and Oancea and used by Elliott in developing his schema - both enables, and provides a theoretical justification for, letting judgments concerning the quality of an action-research account as research to stand also for the standards-setting criteria of the Academy when these are based on selections from Elliott’s distilled list.

Elliott distinguishes between the following three types or class of action-research:

1. External researcher(s) helping those involved in a situation gather evidence so that an expansion of their understanding can enable them to change that situation for the better

2. Experimental teaching which attempts to improve practice

3. Collaboration across a networked learning community in an attempt to improve practice.

Regardless of type, action-research should be judged by the three principles of theoretical and methodological robustness, value-for-use and the potential to enable beneficial change; but within these generalities more context-specific criteria can be applied to each class. If this is true, then Elliott’s paper is, indeed, a long awaited
breakthrough for those of us in the Academy repeatedly faced with the ‘central
dilemma’ identified above.

Adopting Elliott’s own approach, this paper uses a vignette of practice to explore how
Elliott’s criteria can be used within the Academy. The vignette consists of two parallel
stories. The first is that of how critical reflection on two instances of assessment led a
nurse-tutor to improve her own practice - and subsequently document it as the
assignment for a module within an MA programme. The second story is how that
assignment caused the tutor responsible for marking it to revise his ideas about what
constitutes valid and legitimate action-research and about how academic criteria
should be applied to action-research projects. These stories are presented in sequence.
Part 2 consists of an edited extract from the action-research account presented for
assessment towards the MA. This gives details of how practice was not only improved
but made safer. Part 3 presents the story of the assessor and the difficult questions he
faced in deciding whether to pass or fail the assignment. Finally, Part 4 suggests the
main lessons to be learned from these narratives and suggests the adoption of Elliott’s
(2007) criteria as a theoretically sound way of applying rigour whilst recognising the
importance of context. Part 2 is written by Elaine Jefford; Parts 3 and 4 by Tim
McMahon.

Part 2: Extract from an action-research account presented for assessment towards an
MA in Education
The first incident occurred when I was working with student nurses as a teaching assistant for an area medical education consortium - a joint venture between a Health Trust and a University. This employment was part-time, the bulk of my professional practice being as a serving midwife. Within that role, I worked to professional guidelines which required me to ensure that any woman entering the maternity setting should be able to obtain individual fulfilment from their birthing experiences… In order to achieve this, the midwife was expected to facilitate empowerment of the woman concerned. Gibbon (1990) believes empowerment of the individual through self-care is a two-way process, the fundamental issue being communication and the exchange of information, resulting in the ability of the individual woman to make informed decisions. At the time, I firmly believed that individuality and empowerment were reflected in my practice and had become an integral part of my personal philosophy when dealing with any client or person. As the events of these incidents unfolded, however, I came to question whether my values and beliefs really were reflected in my actual practice.

I was asked by the University to assist in what was called the skill laboratory - a simulated ward environment where students could learn new, or consolidate existing, clinical skills in a situation that was considered, by the university-based tutors, to be unthreatening. My intended role was to act as a client, on whom student nurses could practice their aseptic techniques, whilst a university tutor assessed their competence. The model of competence-based assessment used was perceived by the university-based assessors to be congruent with the definition offered by Wolf (1995):

…a form of assessment… that so clearly states both the outcomes - general and specific - that assessors and students and interested third parties can all make reasonably objective judgements
with respect to student achievement or non-achievement of these outcomes; and that certifies student progress on the basis of demonstrated achievement of these outcomes. (Wolf 1995, 1).

On arriving at the skill laboratory for the first time, I was informed that there were not enough lecturers and asked if I would act as an assessor given that I was still grounded within clinical practice. I agreed and was shown to a workstation where there was a variety of nursing equipment and a new “client” (who was a part-time lecturer). She had just placed an artificial dressing on her arm to enable students to practise their sterile dressing technique. I was given a mark sheet that required me to record a numerical percentage mark for each student and left to get on with the process of assessment. I was not given any other criteria and I did not explore why the organiser assumed my competence to assess. Upon reflection, I believe the reason why I did not explore these issues was because I felt confident as a practitioner in my abilities to assess clinical practice.

The first of the students I was expected to assess arrived and proceeded to explain why she wished to redress the wound. Following this explanation, the “client” looked at the student and nodded her head. The student, assuming that the nod indicated informed consent, then changed the dressing in accordance with accepted protocols. This scenario of the nod followed by safe application of a new dressing was repeated by a further five students. I passed all of them as being competent.

Reflecting on the event afterwards, I realised that there were at least two possible scenarios that could cause the client to nod without having an intention to indicate informed consent. Firstly, the “client”, in my assessment situation, could have been demonstrating behaviour equivalent to that documented by Helman (1998) where
patients entering a hospital subjectively experience a substantial loss of control over the body, personal space, privacy and behaviour. In such a case, the nod could have indicated a relinquishing of control and responsibility for the actions about to be undertaken on her wound dressing. Alternatively, because the student used medical terminology likely to be unfamiliar to the lay-person, the nod might have indicated a simple failure to understand what was about to happen. Such occurrences were recognised as long ago as 1961 by Goffman, who suggested that hospitals are small societies each with their own unique cultures, rules of behaviour, tradition, rituals, hierarchies, all compounded within their own incomprehensible language.

In any event, my reflection led me to believe that I had failed to acknowledge the possibilities of these scenarios and thus, like the students, had failed to ensure that informed consent was being given. Further, I began to question whether I had subconsciously anticipated that the client would submissively fall into the medically-dominated mode of thought and accept that the dressing had to be changed and that the nurse was competent. I remain unsure whether my behaviour was governed by the culture of the particular hospital where I was working at the time or by the indoctrination that occurred on entering the medical setting at the young and impressionable age of eighteen, some twenty years previously. Nevertheless, after exploring this issue, I believe that my guiding influence was the ingrained belief that, in medicine, the predominant focus should be on the physical dimensions of an illness - in this case the wound. I collaborated with this medical ethos by allowing the student to reduce the client from an individual to a specific segment of an illness or condition. Engel (1980) suggests that this is the ‘mind-body’ reductionism of the medicalisation of illness, which ignores the contribution of personality and humanity. I had been
indoctrinated with non-humanistic theoretical modes of nursing both as a student and as a qualified midwife and this contributed to my dogmatic approach to assessing the students’ competence in clinical skills. I was endorsing the power-coercive role of the professional within the skills laboratory environment.

Looking back, it is incomprehensible to me that I could have accepted such blinkered authoritarian behaviour, reducing a complex human being down to an object of an assessment of a wound dressing, without personality, feeling, emotions or needs. Alongside this is the realisation that I blindly accepted someone else’s preconceived ideas on what an assessment is and their narrow interpretation of assessment criteria. These revelations have come through self-reflection where I was able to observe what happened and analyse my thoughts, feelings and actions to gain insight into their meanings, whilst simultaneously exposing my weaknesses.

A possible reason for my behaviour is offered by Powell (1989) who suggests that one reason nurses and midwives are not learning from practice may be that they are conditioned to ascribe a higher value to technical-rational knowledge than to learning from personal experience.

Later that same week I discussed this incident with a colleague and came to the conclusion that my own philosophy of care needed re-examining. Also that the assessment process I had utilised was either inappropriate or incomplete as it made no reference to the principles underpinning assessment or to any formal marking criteria. I could do nothing about the assessment because my involvement had been a one-off. (I was not, then, employed by the university or in any way part of the team delivering
and assessing the programme.) I had neither status nor power to intervene and it was unlikely that I would ever see those students again.

I could, however, learn from the incident and seek to ensure that if ever again I had to act as an assessor I would be fully prepared to deal with the ambiguities and moral dilemmas of observational assessment.

The challenge to this determination came just over a year later when I took up the position of a full-time lecturer at the university concerned. As part of that role I was required to assess students within the same skill laboratory. In the meantime, I had commenced study for an MA in Education and part of that programme required that my teaching skills be assessed by a designated professional who is recognised as a competent practitioner/lecturer/assessor. I decided, therefore, to invite this person, whom I will call Sabrina (not her real name), along to assess me when I was acting as an assessor to student nurses in the skills laboratory; she accepted.

Being a full-time member of staff combined with what I had learned about the theory of assessment on my MA programme gave me the confidence to insist on a thorough briefing before I started my clinical assessments. As a result of this insistence, I gained access to a minimum standards sheet and a ‘clinical skills level’ summary sheet. In the latter, the minimum standard to be achieved by the student was given as a checklist of behaviours or actions. However, the performance levels were somewhat ambiguous and very much open to individual interpretation. I was determined not to repeat the mistake of over-generosity (Quinn 1995) of my first clinical skills assessment.
Sabrina arrived and we talked through the impending assessments and the two assessment sheets. As I was assessing the first student it became apparent that his theoretical knowledge and clinical skills were not in harmony. His communication skills were excellent. He gave a full explanation and rationale for wanting to undertake the aseptic wound dressing and answered the client’s queries with a show of understanding. Yet the skills he demonstrated when applying a dressing would have carried an unacceptable risk to a real client.

Amongst a cocktail of conflicting emotions, I felt an element of compassion for the student (in that I was thinking of failing him) and an overwhelming sense of uncertainty for myself. When being assessed by another there is an enormous spectrum of possibilities that extend from inspectorial or judgemental to enabling and empowering. I, like the student, wanted to prove myself as competent, yet if I failed him how would that be interpreted by my own assessor? I even thought, fleetingly, that as Sabrina had been out of clinical practice for a long time, perhaps she hadn’t noticed the student’s errors. But if she had, then the trust between us would be severely damaged. More importantly, my own memory of the unhappiness I had felt after reflecting on my previous experience of clinical assessment made me determined to apply what I considered acceptable professional standards in making this, and all future, assessments.

For me one paralysing thought was uppermost, that I was being observed and assessed by the senior assessor. Should I fail this student or not? Should I permit the student to make mistakes within a ‘safe’ university clinical skills laboratory and, having given
appropriate feedback, still pass him? Was I applying my individual standards appropriately? Boud (1998), however, points out that, ultimately, it is the professional responsibility of the individual lecturer to interpret the expectations of the assessment criteria and determine whether to fail or pass. I decided that this student must fail. In communicating this decision, I felt it was important to help the student reflect upon, and learn from, his performance. From this starting point he should be able to develop his abilities to reflect on his practice and to subsequently progress along the novice to expert continuum (Benner 1984). I asked him, therefore, for a verbal self-assessment against the assessment criteria. The result of this process was that the student stated that he should fail and gave the reasons why. He went on to acknowledge that he had not really practised the procedure, but had just learnt the theory hoping that would suffice. He had not really grasped the reality of either the assessment criteria or the process. We explored the assessment criteria until I felt that he had a reasonable grasp of what would be expected for his re-assessment and he left the laboratory with a recommendation to reattempt the aseptic procedure once he was comfortable in both his theoretical understanding of what was a safe procedure and his practical ability to carry this out.

Sabrina and I also took time to sit and reflect on the incident. She agreed with my decision to fail the student, but stated that if another assessor had assessed him he might well have passed. This compounded my concerns about the validity and reliability of assessing clinical skills by observation when so much depends on an assessor’s subjective understanding and interpretation. Peshkin (1998) suggests that a person’s subjectivity is an intricate part of him/her self, which cannot be removed or switched off. It should, therefore, be acknowledged and considered when setting
assessment and marking criteria for observational assessments. A possible solution that occurred to me would be to have a second assessor, thus tentatively placing individual subjectivity within the realms of team objectivity. Sabrina suggested that in order to enhance the assessment experience I could do two things immediately. Firstly, I could create a document giving information regarding the procedure in a step-by-step format, and give it to the students prior to their assessment. This would have the benefit of limiting the potential for students’ individual interpretations of the purpose and process of the assessment by making explicit our intentions as teachers in this regard. This method could also act as checklist, which would prove beneficial in providing precise feedback. Secondly, I could raise the issue at the next clinical skills meeting of university tutors. I did both of these things. As a result, students now receive a document which gives them a briefing as to what is expected of them when undergoing assessment of clinical practice and which is used by assessors to inform their assessment decisions and to structure feedback. Resources do not permit the luxury of having two assessors present for each clinical assessment. However, the unavoidable subjective dimension of assessment by observation (Peshkin 1998) has, at least, been discussed and acknowledged by the team. Most importantly of all, I now feel confident in applying my own professional standards to clinical assessment. This has made me a better teacher and a better assessor. My colleagues in the teaching / assessment team report that they find using the new document to guide assessment and feedback a useful additional tool. The students from the first cohort to use the new document reported that it helped them to prepare for the assessed task.

Part 3: The Academic Assessment.
The above is extracted from an assignment for an MA module that was submitted prior to the publication of Elliott’s 2007 paper. In assessing the assignment, I could recognise evidence of critical reflection on an experience followed by a planned intervention that, on evaluation, was judged to be beneficial in that it led to:

1. a change in attitude on the part of the nurse-tutor concerned.
2. a change in her personal practice as an assessor in a particular clinical setting.
3. a change in the documentation given to students prior to clinical assessment.
4. a change in the professional discourse of the teaching team indicative of enhanced understanding - in that they acknowledged the existence of problems within the assessment process and had begun a dialogue aimed at addressing them.

Thus, the student had completed one full cycle of activity using action-research methodology. Her work could, therefore, be regarded as a classic demonstration of reflective practice producing beneficial change. My immediate concern as an assessor for a Master’s programme, however, was that I did not feel that there was any justification for calling it action-research. I felt that it did not demonstrate that the student had gone through two complete cycles of the action-research spiral – a condition for “true” action-research laid down by Kember (2000) and one that I had been applying when marking work for the module – a circumstance, moreover, explicitly signalled to the students in the written module guide and in classroom briefings. In applying the requirement for two cycles to be completed, I was not out-of-step with general practice in the Academy. Grogan et al (2007), in exploring appropriate structures for professional doctorates in Education, state that dissertations
would need to be based on at least two iterations of the action-research cycle in which the student diagnoses a situation using pre-understanding of that situation informed by theory; plans an intervention in the organization; implements the intervention; and then studies the results of the intervention. In coming to this conclusion they cite Coghlan & Brannick (2005).

My student, on the other hand, was fully convinced that she had completed a piece of action-research – albeit very small-scale. Her project contained a defence of this assertion (not included in the extract given above) based on her belief that she had met the three conditions Carr and Kemmis (1986, 23 - 24) laid down as being “individually necessary” and “jointly sufficient” for action-research to be said to have taken place, namely:

1. **The subject matter was a social practice susceptible of improvement** – in this case her own performance as a clinical practice assessor, that of her co-assessors and the performance of the students being assessed.

2. **The project had proceeded “through a spiral of cycles of planning, acting, observing and reflecting, with each of these activities being systematically and self critically implemented and interrelated”**. The student acknowledged that she had not completed two full cycles. She argued, however, that learning through action-research is a continuous process that is of a type even in its initial stages. “It enables one to reach an understanding and then apply it to transform experiences. I planned and implemented an improvement in practice. This is the basic purpose of action-research.”
3. **The project had “involved those responsible for the practice in each of the moments of the activity, widening participation in the project gradually to include others affected by the practice, and maintaining collaborative control of the process”**. Here the student argued that the collaboration with Sabrina and the subsequent discussions within the assessment team constituted this widening participation as did the changes to the documentation given to students before the assessment.

I had no problem accepting the argument in point one. I felt that her argument in point 3 was debatable (was this really a collaboration or had she simply caused a change of practice by pointing out where a relatively-easily implemented change could be beneficial?) but I was willing to give the benefit of the doubt. I was, however, unconvinced by her argument that it was unnecessary for her to have completed two full action-research cycles in order for her work to be valid. Carr and Kemmis (1986) had used the plural “cycles” in defining their requirements and Kember (2000) – a key text for the module - had specifically referred to the need for two cycles to be completed. Given this, it seemed to me that there had been insufficient observation of the effects of any change and that, therefore, there were doubts about whether this was or was not action-research.

The student concerned was not one to accept without question my judgement on this matter. In the tutorial in which I gave her my views on her project she made it clear that she didn’t really care that much about the criteria for action-research accounts postulated by others - except in so far as it impacted on whether or not she gained a
pass for her assignment. Both as a reflective practitioner, and as an action-researcher, her main concern was that she had succeeded in both improving her own practice and increasing her understanding of the context in which that practice took place.

As a result of this tutorial discussion – which can be regarded as the first stage of an appeal against the mark of failure I had awarded - I began to see that I had fallen into the trap of elevating models of action-research processes into an abstract discipline with a set of procedures that had to be applied to practice. I realised that I had become part of the process hegemony identified by Sanger (1995) and tyranny of method identified by Thomas (1998). I had raised a barrier to the recognition by the academy of a manifest and well-documented improvement in personal practice, and, in the words of Jean McNiff, had turned

…action research into a technology, an oppressive instrument which can potentially distort other people’s creative practice.” (McNiff, 2002, 52).

Thus I was faced with the dilemma identified in the opening paragraphs of this paper. On the one hand, I needed to apply fixed, published criteria to students’ work in a way that is fair and equitable, on the other, as an action-researcher I could not deny the justice of my student’s argument. Technically, reversing my decision to award a pass was not a problem. The assessment criteria merely required an action-research project of sufficient length and rigour. They did not define action-research nor did they specify how rigour was to be judged – leaving both of these to me as the tutor. In the teaching of the module, however, I had expressly told students that two cycles of the action-research would be necessary to achieve a pass and the printed material given to students also contained reference to this requirement. How then could I justify not
applying this criteria? Again, this was technically unproblematic. In teaching and in
the printed material distributed to students I had made clear that I would always
negotiate both my judgements and the criteria on which they were based where a
sound case for exceptions could be made. The formal module documents (which I had
written) also contained reference to this negotiation. Consequently, I had the authority
to award the pass. I also checked to see whether not completing two cycles was
crucial in the failure of any other candidate, knowing I would have to remark their
work if this was the case. There were no other cases. The work received the pass it
deserved and I had learned a very valuable lesson.

What matters here is that the incident brought home to me the need to ensure that
assessment criteria for action-research conducted within the academy is always
written to allow such flexibility. This is the issue discussed in Part 4.

Part 4: A Flexible Approach to Assessing Action-Research Projects within the
Academy.

At one point in her assignment (see Part 2 above) my student wrote

Looking back, it is incomprehensible to me that I could have accepted such blinkered
authoritarian behaviour, reducing a complex human being down to an object of an assessment of
a wound dressing, without personality, feeling, emotions or needs.

“Blinkered authoritarian behaviour” is, equally, a very good description of my
application of the two-cycle rule in the above instance. It was fortunate, and by no
means inevitable, that the way the assessment criteria were written allowed a change of mind. What, then, are the lessons for the assessment of action-research projects within formal, credentialed academic programmes?

In my view, the most obvious one is the need to ensure that assessment criteria are formulated as generalised global values such as Elliott’s theoretical and methodological robustness, value-for-use and the potential to enable beneficial change. The case for using global statements of this kind in assessment criteria is similar to that advanced by McMahon and O’Riordan (2006) and McMahon and Thakore (2006) when arguing that learning outcomes which are stated in global terms are more likely to develop critical thinking skills and to prompt deep learning. As was found to be the case with learning outcomes, it is likely that it is best to classify more detailed guidance on these criteria as being illustrative or advisory rather than exhaustive and definitive, to present these in student handbooks or their equivalent and to explore their meaning and possible application as part of the learning activities of the programme. Their use may well also provide an appropriate strategic response to Knight’s (2006) identification of three distinct forms of assessment on student work carried out in Higher Education and of the local and contextual nature of two of them – namely those concerned with judgment of evidence and appreciation of subjective phenomena. Whether Elliott’s more detailed criteria should be used as well as the overall global criteria or given as the more detailed guidance is a matter for the professional judgement of teachers who know (or, if they don’t know, should investigate) what is most likely to help their own students. My own experience suggests that while, as a teacher and action-researcher, I would be happy to use only the global criteria of theoretical and methodological robustness, value-for-use and the
potential to enable beneficial change as the published assessment standards, the majority of students are likely to react more favorably to the use of Elliott’s more detailed defining statements.

Whichever option is adopted, it is vital that explorations of the criteria form part of the learning activities of the students. Again, there is a parallel here to the findings of McMahon and O’Riordan (2006) who reported that learning outcomes expressed in global terms significantly prompted more deep learning and critical thinking than more constraining criteria only when they were discussed and explored in class.

The real lesson from the above vignette, however, is the need for flexibility in the application of the assessment criteria by the tutor when confronted by students with new perspectives and insights. This implies some element of student input into the assessment process such as the assessment tutorial mentioned in Part 3. Among other things, this helps ensure that the student is able to draw the assessor’s attention to merits s/he might have missed and / or to enable the assessor to give more focussed feedback. It also provides a structure that recognises, and takes advantage of, the fact that when there is a formative element involved, assessment should be a “collaborative act” (Yorke 2003, 496)

This call for flexibility is not new. In the 1980’s, Boud (1981 and 1986) was one among many who called for elements of negotiation and participation to be introduced into assessment in higher education. One way of doing this is specifically to build into the assessment regime an aspect of negotiation that allows - preferably encourages – students to put forward a case for a particular mark for their work. This is the kind of
self-assessment championed by Boud and others and not only do all the arguments that it would promote deeper learning and greater autonomy still hold good today but they are augmented by the theoretical perspectives developed by Torrance and Pryor (1998, 2001), Knight (2006) and, at least by logical implication, by Pryor and Crossouard (2008).

Unfortunately, such flexibility can be hard to put into practice because of a perceived need to apply the same criteria in the same way to all work in order to ensure fairness through reliability (the criteria are applied consistently) and objectivity (the criteria are applied without favouring some students over others). This is why Elliott’s criteria are so valuable. They provide defensible criteria which can be selected and adapted to context. (This, of course, implies also being able to justify not applying any given criteria.) Adopting them as the published assessment criteria (along with an explanation of how their application involves a debate on which criteria are applied and how) would solve the ‘central dilemma’ of academic rigour vs. action-research as critical investigation unbounded by antecedent conditions.

Had Elliott’s criteria been available at the time, then the assessment of the piece of work from which Part 2 (above) is extracted would have been unproblematic.

The extract given as Part 2, is long enough to show that the student had met, at least eight of Elliott’s (2007) sixteen criteria (Elliott does not suggest that all sixteen have to be met, rather that the relevant criteria are applied in context). She had certainly used critical reflection-on-action to resolve what she perceived to be a problem of practical concern (criteria 1) through the gathering of data from different points of
view (criteria 2) – namely, herself, her observer and the teaching team. Further, she had questioned tacit assumptions - her own, and other’s (criteria 3) and demonstrated that she had extended her own understanding in order to open up new possibilities (criteria 4) in a deliberative process that challenged her own teaching strategies - in this case with regard to assessment - and the aims to which they were directed - in the form of the tacit assumptions of the existing assessment team (criteria 5). Her narrative reveals an open exploration with others of the situation, including revealing and challenging her own prejudices (criteria 6). Her discussions with Sabrina show a determination to find out how others interpreted the situation while the critical reflection contained in her text reveal an honesty about her own motives and a respect for the rights and views of others (criteria 7). Finally, she increased her own ability to do her job in a more effective and fair manner (criteria 8).

The first four of these criteria were seen by Elliott as applying to the externally mediated class of action-research where an external researcher helps teachers gather evidence so that an expansion of their understanding can enable them to change that situation for the better. The second four criteria are seen as applying to experimental teaching. Given that my student’s reflections began when she was brought in as an external expert and led to further reflection culminating in a planned intervention when she became a new member of the team, I feel that both these sets of criteria are appropriate to the context of her research.

This shows how Elliott’s criteria can bring flexibility into the assessment system without compromising either standards or the principle of fairness. Such flexibility can have huge payoffs in terms of helping students to develop higher order cognitive
skills. Its main importance when it comes to assessing action-research accounts within formal, credentialed academic programmes is, however, that it provides a best fit between the methodology of assessment and the methodology of the research that is being assessed.

In the same way that the rules and maxims of action-research are rules of art not science that can never be more that rule-of-thumb guides to practice (Ebbut 1985) so criteria for assessment of action-research projects within formal academic programmes have to be seen as guidelines that inform negotiation between student and assessor.

Just as in the process of judgement one asks questions of a text or a person, the person or text must be allowed to ask questions in return. (Smith & Deemer 2000 p 889)

References:


24


Somekh B & Zeichner K (2009) Action research for educational reform: remodelling action research theories and practices in local contexts *Educational Action Research* 17(1), 5-21


