Practically Perfect?
Excellence in practice-based educational research

A Discussion Paper for Seminar 3
in the
ESRC TLRP Seminar Series
Quality in Educational Research

December 1st 2005, Oxford

Presented by
Stephen Steadman
Sussex University School of Education

This paper draws on research funded by
The Teaching and Learning Research Programme
of the
UK Economic and Social Research Council
Introduction

This paper has two main sections. The first provides a picture of the Early Career Learning at Work (LiNEA) Project that started in January 2001 and ended in June 2005. I was a co-grant holder the project manager, and a field researcher in two of the three professions being studied. The first section, which is based on four recent papers prepared to describe the research methods used in the project (Furner and Steadman, 2004; Steadman et al., 2005a, 2005b, 2005c), gives a fair amount of detail in order to give readers an adequate purchase to form their own opinions of the contents of the following section. The second section examines the LiNEA project with reference to the kind of frameworks and criteria that are being brought to bear to assess quality in applied and practice-based educational research. Because this paper rests so heavily on experiences in a large scale, collaborative research project, I need to make it clear that the opinions expressed in this paper are mine alone.

Section 1 Describing the Early Career Learning at Work (LiNEA) Project

A brief over-view of the project

This research was a longitudinal study of the learning of newly qualified nurses, graduate engineers and trainee chartered accountants during their first three years of employment. The accountants and engineers were formally contracted trainees, for whom employers have developed systems of organised training support. The engineers started with relevant degrees, e.g. in engineering or computer science, and most were seeking the advanced meta-qualification of Chartered Engineer. In contrast, the accountants’ degrees were rarely in relevant subjects and all received formal out-sourced training to prepare them for the professional examinations of the Institute of Chartered Accountants. The nurses provided another contrast. Their initial training programme allocates 50% of its time to working in practice settings and concludes with a professional qualification. Thereafter, newly qualified nurses are expected to take advantage of a variety of in-service training opportunities to broaden and update their knowledge and skills, but communication failures and work pressures often mean these options are not fully used.

Previous research has shown that important components of what is learnt are often carried as tacit knowledge and skills that are drawn on in working situations to address problems without the user being conscious of their utility. Conceptual clarity was essential in the project and much time was given to associated discussions. Eraut’s (2000) paper on ‘Non-formal learning and tacit knowledge in professional work’, provided a sound starting point. In that paper Eraut defined the ‘personal knowledge’ that workers use as what individual persons bring to situations that enables them to think, interact and perform. This definition includes non-codified, or unwritten, personal knowledge and a far broader concept of knowledge than that implied in academic performance. For example it includes, not only personalised versions of public, formal, codified knowledge, but also everyday knowledge of people and situations, know-how in the form of skills and practices, memories of episodes and events, self-knowledge, attitudes and emotions. It focuses on the value of knowledge in use. Unless one stops to deliberate, this kind of knowledge is usually available in an integrated form and ready for action. Much of this personal knowledge is, of course, taken for granted and often difficult to explain – especially to a researcher. Even formal learning on courses or in workshops, although readily recognised as learning, may be forgotten after a while. So the project had to devise a methodology to access what people need to know at work when most learning at work is informal and therefore unlikely to be readily acknowledged or scarcely remembered without some pertinent prompting.
The central research questions were:

- RQ1 What is being learned?
- RQ2 How is it being learned?
- RQ3 What are the main factors affecting this learning in the workplace?

The methods the project used to develop responses to these questions are described in more detail below. In this brief overview, it is necessary to skip ahead to the project’s outcomes. In theoretical terms, these mirrored the main research questions (Eraut et al. 2005; Maillardet et al. 2005):

- A progression typology of learning trajectories was proposed in answer to RQ1.
- An analysis to distinguish different work and learning processes from the learning activities workers employ is used in answer to RQ2.
- For RQ3, a ‘double triangle’ framework is offered to analyse both the factors which affect learning (Challenge and value of the work; Feedback and support; Confidence and commitment), and the contextual factors (Allocation and structuring of work; Encounters and relationships at work; Individual agency and expectations of performance and progress) which influence the learning factors.

The project also produced three Interim Sector Reports. The interim findings they contained led to action research with some of the companies and organisations that had been the project’s partners, as the original research proposal had envisaged. More detail on this is given in a paper in Montreal at the AERA conference (Caballero et al. 2005). A third layer of outcomes is represented by the project’s various conference papers that describe project’s work and its findings.

**General features of the research design**

The project combined visits, observation and interviews, with managers and mentors as well as the new employees themselves, over three years to gather its data. The resulting data set comprises field notes on observations in the workplace, and transcripts of interviews with learners together with transcripts of interviews with their managers, mentors and colleagues. We also had access to a variety of documentary sources such as work logs, training programmes, etc.

The central research questions were originally developed for a previous project on mid-career learning in the workplace (Eraut et al. 1998a and b, 2000). That research collected its data from two interviews held some months apart. Key changes in the design of the LiNEA project were the addition of a period of one to two days observation prior to interviewing, repetition of the observations and interviews over a far longer period, and the addition of data from interviews with managers, mentors and other work colleagues. These changes in the research design reflect the attention that was given to the methodological problems of eliciting tacit knowledge. For example, the observations not only provided the evidence of the field-notes, they also enabled us to use workplace documents and activities as starting points for conversations about the use of embedded knowledge and its acquisition that would otherwise have been impossible. This was an important development, but other challenges remained.

The scale of the project, both in its scope and its time span, brought problems. Table 1 summarises the number of main informants from whom data in the form of recorded and transcribed interviews has been collected in this longitudinal study. It was expected that numbers would dwindle over the three years and the table shows that they did – to the greatest extent in nursing, as was also anticipated. The table also shows that the project had difficulty recruiting partners in the accountancy sector. Several explanations are possible, but there are no clear favourites. In certain circumstances it was possible to record the views of other people in the work place, and the project collected a further 107 interview transcripts. Another source of data is the corpus of field notes written up for each visit.
Table 1: Numbers of main informants and interviews

<table>
<thead>
<tr>
<th></th>
<th>Newly Qualified Nurses</th>
<th>Graduate Engineers</th>
<th>Trainee Accountants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviewed in first round of visits</td>
<td>40 (6M)</td>
<td>36 (10F)</td>
<td>14 (3F)</td>
</tr>
<tr>
<td>Interviewed in final, third year</td>
<td>21 (4M)</td>
<td>34 (10F)</td>
<td>11</td>
</tr>
<tr>
<td>Total participant interviews</td>
<td>91</td>
<td>121</td>
<td>53</td>
</tr>
<tr>
<td>Interviews with managers, team leaders</td>
<td>40</td>
<td>92</td>
<td>22</td>
</tr>
</tbody>
</table>

Practical problems are bound to arise when seeking to observe and interview engineers liable to be work station-centred or on site, accountants on client’s premises and client’s time, and nurses who work changing patterns of shifts. Yet we needed to maintain relationships with organisations and with our informants over several years while personnel changed and some of our informants moved to other jobs. The nature of the changing relationships themselves affected the way observations and interviews were conducted. The focus of the interviews had to change with time, partly because background became filled in, and partly because we needed to collect adequate samples of the learning domain. It was also important to balance the competing needs of consistency in the application of our procedures across the whole project against the need for sensitivity to sector differences. No small part of this was the challenge of developing a reliable common system of coding and cross-checking for different researchers working in different sectors. In fact, the whole business of data handling, reduction and analysis has been crucial in a project designed to amass very large quantities of qualitative data.

Time and timings

Time on a project is invaluable. Timing matters too, but it is not always within one’s control. In the summer of 2000 ESRC/TLRP approved the project, with a starting date of January 2001. We were informed of this in summer 2000, but the associated contract did not arrive until the turn of the year. Understandably, without a firm contract our universities did not allow us to advertise for the three research fellows we required, let alone appoint them. It took three months, in one case longer because of the need to re-advertise, to appoint the three research fellows. This was a serious delay that upset a training schedule for the research fellows, and a sequence of piloting. Both had been set out in our proposal.

There were other knock on consequences. For example in the nursing sector, until we had named researchers we could not submit our proposals to the various ethics committees at NH regional and Trust levels. We were relatively fortunate in that the additional information we were asked to supply for consideration did not lead to further delays, but there is no hierarchy whereby approval at some ‘senior’ level clears the path at other levels.

During the planning of the project we had been in negotiation with potential partners in each of the three sectors. The delay between approval and an actual start to the project affected the accountancy sector the most. Some organisations now proved extremely slow to allow us the entry and access that we thought we had already agreed. In one case personnel and/or responsibilities had changed. In other firms the project had lost priority and momentum. In many cases we had to restart negotiations almost from scratch, and we did not end up with the numbers we had hoped for. For a long while we diverted resources trying to better the situation, but with very limited success.
In the project proposal, beyond saying that we intended there should be four visit/interviews over three years, we had not specified the intervals between them too closely. This was just as well because this part of the project’s work did not run smoothly. On top of the problems caused by the delayed start, two further problems emerged. First, the preparation of the interim reports, based on the first round of visit/interviews, took longer than anticipated and, in some cases, prolonged the interval between the first and second round of visits at a time when one might expect trainees to be learning most rapidly. Second, we experienced difficulties in arranging subsequent visits and being able to rely on appointments. Although we had anticipated these problems, and to some extent allowed for them in our planning, the scale of the problem, particularly in the accountancy sector, and in nursing, was much larger than expected.

**Observations – intentions versus reality**

The place of observation in the research design, and its combination with semi-structured interviews, was of key importance because of the known problems in accessing informal and tacit knowledge referred to in the overview above. But we had to have a common approach to making the observations and it was agreed that we would focus upon the following seven areas:

1. what our key informants were doing;
2. what objects they were using, e.g. equipment, software packages, manuals, hand-over sheets;
3. work and work-related social interactions, e.g. consultations, meetings, clubs;
4. routines - daily or longer, or related to work patterns;
5. the nature of environment, e.g. general surroundings, inc. things like noise levels;
6. things taken for granted, e.g. dress codes, locking drawers;
7. reactions - including to us as observers.

We did not use schedules and researchers could, and did, decide for themselves what was relevant. The observations could indicate the knowledge, skills and understandings being deployed – how and where did our informants learn them? It was also possible to see people in situations where they were ‘forced’ to learn, e.g. two trainee accountants using a software package that was new to them.

Observation brings many advantages. Some examples are that observation:

- educates the observer/interviewer about the working context, necessary tasks and priorities, and thus enriches subsequent data gathering;
- provides ‘clues’ to the use of knowledge that must have been previously learnt;
- allows complexity to be appreciated;
- enables comparisons that suggest questions; and
- discourages the painting of ‘ideal pictures’ by informants when they know reality has been observed;

A main advantage is that, with observation as a starting point, an interview becomes a discourse of description, rather than a discourse of justification that can so easily result from asking what an informant has had to learn in order to do the job. Other benefits accrue from combining observation and interview data. The meaning of observation data alone is not always explicit. What should an observer infer from seeing a trainee accountant walk round an organisation, following the path of an invoice through the stages of its processing? But being able to discuss the events with the trainee, ask what was going on, and what the trainee found out, helps to decide the correct inference and brings out what was being learnt. Other examples of productive periods of observation included the handovers on hospital wards to new shift teams, and walking the streets where an upgrading of water main systems was in progress with the planning engineer, and being shown what the options for upgrading were. The civil engineer had herself found it essential to make the same field visits in order to understand the implications for the work gangs on the ground of what she was planning back in the office. Observation thus enriches subsequent data gathering during an interview.
The visits as planned in the project proposal, were supposed to take the following form. Day one, go along for an extended period of observation, stay out of the way and keep quiet, except at breaks and lunch time, but talk to the managers and mentors as you get the chance to do so. The next morning continue observing, and in the afternoon do the interview with the advantages of being able to refer to observed incidents and processes where the informant has either been learning or appears to have drawn on previously learnt knowledge, skills or understandings.

But, as noted above, observation often runs into practical problems that have to have negotiated solutions or, more accurately, compromises. In the nursing sector we found that nurses were so busy, finding the time to do the interview was a problem in itself, and by the end of their shift, they had often forgotten what they were doing at the start of it. This is a problem which previous researchers have noted. Being able to refresh a nurse’s memory from the observer’s field notes, or by other means, such as the digital photographs that Christine Fessey (2002) used in her very rigorous research, is absolutely necessary when this happens.

In the other two sectors different problems arose. Trainee accountants can spend significant parts of their working lives on the premises of clients. Although this could have raised significant problems, in practice, by allowing the accountants to negotiate access on our behalf, this did not prove to be the case. Most clients accepted that it was the trainees who were being observed, not themselves, and we were careful in our positioning and other actions and conversations on site to make it clear that this was the case.

Nowadays many engineers and accountants spend prolonged periods in front of computer screens, much more time than we had ever expected when planning this research. An observation period of up to six hours could potentially be spent looking at the back of someone’s head. So we negotiated with the trainee and manager what should happen in those circumstances and, if it was not profitable to prolong an observation period, we negotiated it down, or we rearranged how we spent our time.

Generally we found it helped to use tea and coffee and other breaks to raise some of the things we thought we may have noticed. This kept us on track, made sure we understood what the trainee was working on, and gave opportunities for more general conversation that often identified topics worth following up in the formal interview. In all sectors, being an observer brought problems of where one should place oneself to be able to see, yet be out of the way and definitely not in the line of sight of the person being observed. This proved awkward at times but we did not find this an insuperable problem.

Observation of an individual over many hours has many of the characteristics of ‘shadowing’ and there is the danger of becoming an oppressive presence, which raises some ethical doubts about using this technique. Presumably, a sense of oppression was felt by one trainee accountant who was sufficiently disconcerted by the first period of observation to wish to drop out of the project. However, subsequent negotiations with the trainee, and their training manager, found a compromise. The trainee agreed to be interviewed on each round of field visits, but not observed. Clearly, we agreed to this arrangement because we did not wish to lose all the data from this individual, and we hoped the decision would eventually be reversed when the trainee could see how the project was progressing. But we also agreed because we were still able to observe other trainee accountants in the same organisation and use insights from that set of data when interviewing the one person who objected to being observed. Thus the arrangement did not completely undermine our position on the importance of using data from both observations and interviews.

Very few others among our 90 trainees showed signs of this being a problem. However, we have found that, as nurses move into more senior positions, they are less happy about being observed and less happy about taking time out for an interview, even when the time requested is only 30 minutes. Oddly enough, in one engineering firm, it was someone totally unconnected with the project, and to this day not known to us, who found the presence of our observer unsettling and complained to the management. They quickly resolved the issue by explaining fully what was going on.
In practice, it did not seem to matter very much if the two days of a visit were not consecutive. We were still able to pick up on enough of what was going on. And the visits were not always two days in length because of the patterns of shift work, or the time taken in travelling between a work site and the office. The key consideration appeared to be, whether enough of the action was observed to provide sufficient material for meaningful interactions with the trainee and their managers. We, as observers, had to see enough of the trainee’s working life to gain an understanding of their present workload, and the knowledge and skills it required, so that our interviews could be enriched by reference to the shared reality.

**The interviews, how they developed over time, and sampling what was learnt**

Interviewing as a means of data collection is itself worthy of scrutiny. Saeeda Shah (2004) refers to the abundant literature on interviewing methods. She lists 26 notable sources, and still amply justifies her focus on cross-cultural aspects of this form of data collection – a topic of strong relevance to our work with the significant ethnic minority group representation in all three of our professional sectors. Doornbos, Bolhuis and Denessen (2004) have recently researched work domains and work-related learning in the Dutch police force. The researchers deliberately had a male-female pair of researchers conducting their single interviews in an attempt to minimise some of the known problems caused by cross-gender interactions in relation to information that informants may deem gender sensitive. Aware of these and other considerations, we used at least two observer/interviewers in each sector, although not interviewing as a pair because we valued the personal relationship over time that develops in successive visits and interviews.

The interviews have all been semi-structured. They had to be designed this way in order to allow the introduction of leads from our observations, and to allow the informants to talk about what has proved important to them. Our intention was to tape record every interview with each participant, and this was achieved with very few exceptions. Notes were taken at the same time, in case of accidents, and on the few occasions when equipment, or the researcher, failed to work as planned, the notes proved their worth. But over a period of three years the interviews inevitably had to change in their focus.

First round interviews concentrated on the trainee’s background, the transition from higher education into the working environment, how useful their HE studies were proving, the experience of induction processes, the different kinds of support provided and what the interviewees were most aware of learning. The interview also asked about the interviewee’s expectations – both of how their job would develop, and where they expected to be in future years.

The second round interviews also had many aims. One was to fill in any gaps in the background demographic information. Another was to recheck topics in the first interview such as the usefulness of the HE, and who was proving useful to the interviewee in terms of support or otherwise. This was often not the person or people officially designated as mentor or team leader. But once again the main aim was to dig into what the interviewee had been doing since the last visit and what this meant s/he had had to learn. The engineers and accountants in our sample often had logs of their activities that provided useful aids to the trainees’ memories of what they had been doing. If they were compiling a record of relevant experience as required in preparation for becoming ‘chartered’, this too was a useful reference aid. Nurses’ memories over the medium to long term linked more strongly to their moves between wards, increasing seniority and courses they had attended.

In the third and fourth interviews we were concerned to ensure for any one individual, and for each of the three sectors, that we were sampling from the whole of the possible learning domain. If we merely followed what our informants were telling us, we were at the mercy of their memories. But our basic assumptions included the forgetfulness of informants, as well as the difficulties of accessing what they did not hold in conscious memory. True, we were being guided in our questioning by what we observed, but what we could observe in one or two days was too small a proportion of a whole working life for us to be comfortable about covering an adequate sample of things learnt.
We therefore devised a supplementary approach in these later interviews. By this time we had drawn up a draft of the ‘progression typology’ as one possible answer to our first research question. Each one of the eight sub-sections in the draft typology was turned into a prompt card, which was put on the table in front of the participant midway through the interview. They were then asked to look at each card in turn, and to talk about any areas of learning that they recognised as areas of accomplishment not previously discussed, even as areas to which they would liked to have given more attention. This proved a fruitful approach in reminding the informants, and in providing reassurance to the research team. It was clear from the interviewee responses that we were picking up areas of learning that we would otherwise have missed.

**Relationships and ethical considerations**

With this number of interviews extending over three years it was vital to maintain good working relationships with the trainees and there were many ethical issues to bear in mind. Issues of confidentiality were addressed in our early planning and in our written communications with potential informants when we sought their co-operation. These issues live on after the formal end of the project in two respects. First, we have not yet revealed the names of the organisations which have allowed us access for our research. We want them to be as fully informed as they can be about the nature of the project’s outputs before they decide whether to be named in our final reporting. Second, as we prepared to offer our data to the national archive, we met numerous practical problems in rendering the original data suitably anonymous. If the process of making texts anonymous is carried too far, the data becomes less useful to other social science researchers.

We found the process of seeking approval from the NH Trusts’ ethics committees onerous. Mainly this was because the questions to which we needed to provide answers were predominately posited on a (false) assumption that we intended to conduct some kind of experimental test of clearly formed, pre-existing hypotheses, preferably using control groups. Although this did not fairly reflect the aims or intentions of our project, our answers must have satisfied the committees, and we received the necessary approvals.

Nevertheless, we were aware that our methods had the potential to make trainees, and their close colleagues in the workplace, more aware of difficult circumstances, or personal short comings (theirs or those of colleagues) than would be comfortable, and we planned our working methods accordingly. One decision was to avoid any discussion with their managers of the progress of individual, named trainees. We had no wish to be seen as management or ‘centre spies’. Similarly, although it proved impossible to avoid occasional voluntary comments, we never sought negative comments about trainees, managers or their colleagues. On the contrary, we specifically asked trainees who they considered to be good role models, so that we could explore what it was the trainees still felt they had to learn.

In our approaches to organisations we made it clear that trainees should be volunteers, and we sold the project to the trainees on the strength of the stimulation involvement would provide for reflection on career progress over time. Initial reactions were good and were consistently favourable. Although it is standard practice to send informants transcripts for correction and verification, in our programme the trainees knew we would be back. This undoubtedly made a difference. Although trainees jibbed at reading through the long transcripts, often preferring to receive a summary, they appreciated the documentation we provided them with during the visits. We also took great care in our handling of informal interactions and in allowing for preferred means of making contact. To over generalise, in engineering we used email, in nursing we used the phone.

In another aspect of ethical conduct we took steps to ensure that all our informants had access to our papers and the interim reports. It was important for informants and participants at all levels to be able to appreciate what kind of research they were assisting and to receive as much feedback as we could provide. Putting as much as possible onto our project web site has been an important element of these
intentions. It would have been much harder in pre-web days, but we have still needed to tell people about the reports and draw their attention to the web site.

The rewards of our working to develop these relationships came in many ways indirectly. Trainees have been willing to self-report on their reactions to being observed and to offer comments on our research and the methods we used. Some contributed documentary evidence. Many spontaneously spoke of the benefits they gained by being involved. But perhaps one of the downsides is that, as we were seen more as colleagues in a joint enterprise, the surer they became that we would not mind if an appointment had to be rearranged, ‘just this once’.

Data handling and data reduction.

The project generated very large quantities of qualitative data and there are current, stimulating debates over what can be seen as appropriate and authentic ways of handling, reducing, and analysing qualitative data, especially in very large data sets. See, for example, the exchanges between Hodkinson (2004) and Hammersley (2005). We have well over 500 transcripts of interviews – most of which exceeded an hour’s duration – with our main informants and their work colleagues. Thus it was important to have well-organised procedures for handling this much data, let alone for reducing and summarising it in suitable fashion, and analysing it to best effect. This part of the paper describes how we approached these issues and Figure 1 provides a pictorial representation of the process.

Our semi-structured interviews with trainees were all recorded on tape and then transcribed by a team of transcribers. Transcription of the spoken word is not straightforward. Different conventions suit different purposes, and we had to make decisions about appropriate punctuation and other matters to suit our final intentions, including an obligation to offer our data to the ESRC Qualidata UK national archive at the end of the project. Our transcribers were not familiar with the jargon in each sector, so transcripts were first corrected by the research fellow who had conducted the interview, and then returned to the trainees who were invited to make additional corrections/amendments. We found that very few suggested amendments. However, this is an important step in data verification.

The transcripts, even with a reduced font size Times New Roman 10 and reasonable margins, usually exceeded 20 sides of A4. In some cases we produced condensed summaries of the interviews in which we were able to re-order the data, bringing information on the same topic from different times in the interview together in one place. We then included contextual information from the observations that preceded the interviews to produce an ‘account’ of a visit. Producing an ‘account’ may be regarded as a first step in interpretation (Miles & Huberman, 1994, Chapter 4). However, care has to be taken in the preparation of interview summaries and the final ‘account’ if it is to fairly represent the original data. Careless paraphrasing can alter the whole sense of an apparent learning episode described in an interview. It is often safest to use new text almost solely to introduce relevant extracts from the original transcript. Aware of this, ‘accounts’ were also sent to informants for verification.

Those analysing the data from the accountancy sector decided to convert their transcripts into ‘accounts’ for verification and further analysis, but did not abandon the creation of transcripts because of the project’s obligation to archive their data. This approach was not adopted by the engineers and nurses, mainly because of the extra work involved in using both transcripts and ‘accounts’.
Figure 1: A Representation of Data handling and Analysis in the LiNEA Project

Earlier research results

Field work, including observations & interviews

Transcribers

Research Fellows

Visit accounts

Informants

Verified transcripts &/or Verified accounts

Early transcripts and accounts

Code scheme

Nud*st N6 Qualitative data analysis package

Ongoing analytical discussion, using data as it becomes available

Circle is domain of analysis

Field notes

Raw transcripts

Corrected transcripts

Verified transcripts &/or Verified accounts

Visit accounts
One argument for ‘accounts’ was that they should be used instead of transcripts to save time, but there are several other arguments in favour that have been set out by Eraut (2005). Accounts can include observations from field notes, and provide a single document for each visit. An account’s shortness makes it much easier to spot connections between different parts if it and begin to theorise.

Similarly, reading ten or even twenty accounts to get a holistic picture of the nature of the evidence is more feasible than reading the same number of less structured transcripts, plus the field notes. Where respondent validation is more important than the authenticity of a text, accounts are more likely to gain responses from informants who are not confronted with their personal speech idiosyncrasies rendered into print. What the respondent is asked to scrutinise is the researcher’s interpretation. They appear more willing to do this than to challenge what was on tape. This helps to make the final research paper or report more authentic.

Transcripts and visit ‘accounts’ were both entered into Nud*ist (N6) as texts ready for coding. When interview transcripts, rather than visit ‘accounts’, were coded, contextual information, from field notes, was entered as foot notes to the transcripts. Issues associated with coding are revisited below.

The choice of an analysis package for large qualitative data sets is no easy matter. The Nud*ist web site provides a great deal of guidance and references to previous studies that have used the package. We considered using the newer, alternative data analysis package from the USA, Qualrus. This is designed to handle qualitative texts and video material. But finally we chose Nud*ist (N6) mainly for pragmatic reasons. It was available as standard software in both of our universities. This meant no additional costs would be incurred and, at that time, we did not have the RAM memory capacity that Qualrus requires on our project laptops. It would have been helpful to have had, at the time we were facing the decision, the kind of comparative overview of the most important computer-assisted qualitative data analyses software packages, ordered by product functions that is available on the web today. (CAQDAS, 2005)

The thinking behind the coding and analysis.

The focus in this part of the paper is on how developments in the team’s thinking influenced the production of a workable coding frame and the application of that frame in practice.

As already stated, we came to the LiNEA research project armed with the findings of the previous investigation of learning at work among people in mid career. We further developed our theoretical position in a series of papers and in-house discussions that took into account recent and contemporary work of other researchers. Then we drew on our initial examination of transcripts and other data from our first round of visits, observations and interviews, and all this work led to the production of our initial coding frame.

Possible approaches to the analysis of ‘raw’ qualitative data span a wide spectrum. At one end the well-documented ‘Grounded Theory’ approach, in its purest form, treats every ‘explanation’ of data as a hypothesis that should be rigorously tested by a search for counter examples. Coding in this paradigm is thus a tentative enterprise to start with, subject to false starts, recasting of the codes and readjustments as the research progresses. At the other end of the spectrum some researchers are able to approach their data from a well-established theoretical position which allows them to decide in advance how their data will be coded and interpreted.

Our position was somewhere in between, perhaps nearer the latter end, but we were still engaged in an interaction between the data we were collecting and the development of our theoretical thinking. Our initial coding frame reflected the interface between an individual worker and the organisation they worked with, considered the transitions needed to become a highly skilled worker and the support they required to do this, and included a focus on what is learnt and how this happens – two of our main research questions. As our work progressed we obviously had to make adjustments to our initial coding frame. The coding frame we started to use was actually our third draft. And rather late in the
day in two of the sectors, a major amendment was to include the eight main headings from the
progression typology as additional codes to cover what was being learnt in more detail than had been
the case in the earlier code categories.

Coding

Early group sessions in which team members had together inspected and coded hard copy transcripts
showed expected divergence in the way coders used the project’s coding frame. Not only did different
coders assign different codes to the same piece of text, coders even differed in which pieces of text
they elected to code. Lack of between coder agreement is not unexpected, especially when using
complex coding frames, and our coding frame at that stage had ten main categories and 42 sub-
categories available for use. Few coders used anywhere near the full 42 subcategories because of the
nature of the transcripts, but it is a reasonable supposition that use of a coding frame with fewer
categories would produce greater apparent agreement. The guidance for coding at that time was quite
general in nature, being as follows:

‘Try to keep to these principles:

• Code fairly general themes, not detail.
• Do not try to code everything.
• Remember it is permissible to multi-code any one segment.
• Give priority to themes we expect to appear
  (These are implied in the code scheme anyway).
• Do at least two sweeps through a text: the first to identify and code the obvious, the second to
  see what can be coded in what remains.
  (A third sweep is better after several texts have been examined).’

But it was clearly necessary to practice more, and with better guidance. So we developed a system of
‘focussed’ coding in our treatment of the data that deliberately highlighted the project’s three main
research questions. In this system first we focussed primarily on identifying examples of what had
been learnt. At times this required us to infer what had been learnt, and making these inferences with
confidence often required specialist knowledge, which explains why we needed our research fellows
to do the coding. The coding then focussed on how that learning happened and, finally, on contextual
influences. Thus all three research questions were emphasised. The guidance for focussed coding was
as follows:-

‘Focus on what is being learnt.

• Look for obvious examples of what is being learnt.
• Look for indications of how the learning came about and code these.
• Find passages that illuminate these examples and code them appropriately. This will bring in
  the relevant contextual factors.

  Repeat. But now start by looking for what we can reasonably infer is being learnt.’

Although it was expected that this procedure should bring greater agreement on what and how to code
in respect of what was learnt, our first trials were not a sparkling success. The first step of looking
only for obvious examples of what was being learnt had the effect of restricting the number of codes
under consideration to seven. But this was not the problem. Despite vigorous encouragement to code
only the most obvious examples of learning, some of us continued to make numerous inferences.
Applying inference was meant to come during a second look at the data but, as one research fellow
commented, what is inference to one coder may be blindly obvious to another who is more familiar
with the sector.
Subsequent debate within the team brought out the importance of several points:

- having what is learnt clearly identified in the transcript’s text
- having clear evidence of learning (this might be elsewhere in the transcript text).
- not wanting to code everything that might just possibly have been learnt.
- being ready to apply a test of possible quotation – would this extract, if quoted, convince a neutral reader?
- understanding the meaning of the codes.

**Understanding the codes!**

The codes on paper do not, in themselves, form the code frame, because they do not carry the full meanings accorded to the main code headings or the sub-categories that have been decided in the project team’s discussions. The printed code frame merely represents one part of the ‘construct’ which is the code frame in action. In order to apply the code frame, we have to understand the basis on which the main and sub-categories have been decided, and subsequently labelled. To be manageable the frame cannot be too long, which means there has to be abbreviation of the headings and labels to the categories. Applying the code as intended, therefore also means remembering the process by which the code frame was agreed, and thereby remembering the meanings and intentions behind the labels. This implies that we must keep in mind the aims of the project. In addition, since coded data can be interrogated to answer different questions, we must also keep in mind the different purposes whenever necessary. This is not easy. Having some pre-set procedures to follow reduces demands on the memory, but it does not mean that memory is unimportant.

Therefore, a written coding frame can, at best, only summarise the theoretical understandings that underpin the construction of the coding frame. People who do not share the theoretical understanding will apply the coding differently from those who do. The difference between high and low inference coding is also important in the way the codes are assigned. We felt high inference methods appropriate because of the qualifications of our research fellows in the work sectors they were researching, and because of the theoretical backing we believe we have for the coding frame. And, of course, the person who observed and interviewed the trainee is the person doing the coding.

**Consistency, reliability and validation**

At this point an informed reader will have in mind several questions of the following kind:

- Would the same interviewer on a different occasion have collected the same data?
- Would a different interviewer have collected the same data?
- How much is changed/lost in transcription or the construction of a visit/interview summary?
- How much is this affected by who constructs the visit report?
- Would the same coder code the same way on a different occasion?
- Would different coders code the same data the same way?

These are important challenges for any social/educational research project. Mainly, but not solely, they are questions about the consistency and reliability of the procedures used in the research. In our project we opted for legitimate differences between sectors. We expect those with expertise in a sector to be able to interpret the sector’s data differently from non experts, and to be able to apply inference in ways that are unavailable to non experts.

How then can we lay credible claim to inter-coder reliability, and to consistency of interpretation in the coding? Part of the answer is that we place our faith in following agreed procedures such as we have developed in the guidance for coding shown above. But that alone is not enough. Alert to these issues, we have scrutinised our three interim reports to examine the differences between them and their significance. Some basic statistics show real differences in the way each report was presented to its readers (See Table 2 below).
The contrasts in presentation are marked. The engineering sector report has by far the most tables, while the nursing sector report uses nine observational vignettes which are not used at all in the other two reports. The way in which recommendations are summarised for consideration differs from one report to another. And yet anyone who reads all three reports will see that they have a common focus on the learning needs of the newly qualified workers and the support that these people need and rely on to make good progress. In comparison with nurses, engineers have a different appreciation of what counts as data and are more comfortable with information summarised in tables. The presentation of each report made allowance for the expectations of those in the sector who we wanted to read the reports and take action as a result.

Table 2: Differences in presentation of the Interim Reports

<table>
<thead>
<tr>
<th>LiNEA Project Interim Reports</th>
<th>Differences in presentation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nursing Sector (37 pp.)</strong></td>
<td></td>
</tr>
<tr>
<td>Quotations from interviews:</td>
<td>Most?</td>
</tr>
<tr>
<td>Observational vignettes:</td>
<td>9</td>
</tr>
<tr>
<td>Tables:</td>
<td>2</td>
</tr>
<tr>
<td>Figures:</td>
<td>0</td>
</tr>
<tr>
<td>Action points:</td>
<td>5 for Trust SMTs, 11 for Ward level</td>
</tr>
<tr>
<td><strong>Accountancy Sector (34 pp.)</strong></td>
<td></td>
</tr>
<tr>
<td>Quotations from interviews:</td>
<td>Fewest? (But fewest trainees)</td>
</tr>
<tr>
<td>Observational vignettes:</td>
<td>0</td>
</tr>
<tr>
<td>Tables:</td>
<td>2</td>
</tr>
<tr>
<td>Figures:</td>
<td>2</td>
</tr>
<tr>
<td>Conclusions:</td>
<td>Unnumbered, 4 discussion headings</td>
</tr>
<tr>
<td><strong>Engineering Sector (35 pp.)</strong></td>
<td></td>
</tr>
<tr>
<td>Quotations from interviews:</td>
<td>Many</td>
</tr>
<tr>
<td>Observational vignettes:</td>
<td>0</td>
</tr>
<tr>
<td>Tables:</td>
<td>13</td>
</tr>
<tr>
<td>Figures:</td>
<td>2</td>
</tr>
<tr>
<td>Action research topics:</td>
<td>6</td>
</tr>
</tbody>
</table>

We also intended that the interim reports would provide the basis for action research by our partner organisations, supported by the project. We believed this to be a sound mechanism for testing the project’s findings as a partnership enterprise in working contexts. In this we have had most success in the health sector where our interim report was warmly received and its messages welcomed. A National Health Trust has formally initiated an action research programme designed to provide more support for newly qualified nurses, and we appointed an additional team member to facilitate and help evaluate the process.

In accountancy, action research saw us helping a firm to evaluate its system of appraisal for trainees. In the engineering sector, however, despite that sector’s interim report being positively received by all our partners, it proved difficult to negotiate an agreed focus for action research. Then an IEE ‘accreditation visit’ to one company resulted in a firming up of mentoring procedures along lines strongly indicated in our engineering sector interim report. This yielded an appropriate opportunity for some action research. It also provided a powerful, if indirect, validation of our findings. The production of the interim reports and our follow-up with our partner organisations thus cemented our relationships in a number of ways. Our claim here is that our partner organisations have validated the
interim reports by their acceptance of them and the level of their willingness to make interventions based on the reports’ findings.

**Consistency in coding**

It is tempting now to ask, do differences between interviewers and coders matter with these outcomes for users of the project’s findings? But the answer is still, yes. They could matter a great deal. Given this is true, it is surprisingly difficult to find previous work that addresses the problems of inter-coder reliability when very large quantities of qualitative data are being handled in a software package such as Nud*ist. We were pleased to find an account by Jenny Cannon (1998) who describes the processes and outcomes of the attempts of her team of five coders to make sense of 10Mb of unstructured text data from 163 interview transcripts. We have over 500 transcripts, whereas Cannon quotes Lee and Fielding (1996) as claiming the median sample size in qualitative research is 40 scripts. But we were horrified to read that at one stage her team had more than 700 free nodes (coding categories) - eventually reduced to a still staggering 380 – and disappointed that the article, while helpfully informative about index construction in Nud*ist, said very little about inter-coder reliability.

Sylvain Bourdon (2000), on the other hand, addresses the computation of an index of agreement directly, and describes how to use the command structure within Nud*ist 4, together with a word processor, to do so. But the method is suitable only for pair by pair comparison, using very few coding categories on identical transcripts in an experimental setting. So we have had to tackle the problem ourselves.

We elected for a robust approach, using transcripts rather than working within Nud*ist. Without sight of first results, the same researcher may code a fresh copy of the same transcript after an interval of time – we tried about a week. After re-coding in this way, an index of self-consistency can be computed by counting how many codes in the first attempt are preserved in the second attempt, and then dividing this number by the average number of codes assigned in both attempts. The resulting proportion gives an indication, an index even, of stability of coding – an index of self-consistency in coding.

If different coders code the same transcript, an index of inter-coder consistency can be computed by counting how many codes agree, and then dividing this number by the average of the total number of codes assigned by each coder. The result provides the index value. However in both cases, self-consistency and inter-coder consistency, this is a crude approach partly because the total number of codes assigned will affect the value and meaning of the index. This is something Bourdon does not fully recognise. The value and meaning of the index will also, less noticeably, depend on the number and variety of codes the coder has available for use, and attempts to use. So there is still some way to go before a satisfactory means of reporting is devised. There has been remarkably little research on these issues, probably because, until relatively recently, qualitative data has not been amassed in large quantities with any real prospect of its satisfactory analysis. The aid of computer packages has changed all that and we now need to devise sound ways in which teams can co-ordinate their analyses. The process in relation to coding data is not dissimilar to that required in the field of examinations for training assessors to make reliable assessments that demonstrate a good level of inter-assessor agreement. This problem has been much better researched than that of inter-coder agreement; and it indicates both that training similar to that described above is necessary, and that further cross-coding activities are required to maintain a reasonable level of agreement after such training has finished.

**The development of the project’s models**

In order to understand the way the LiNEA project developed its three models in response to the three main research questions, it is essential to appreciate that LiNEA followed an earlier study of the learning of groups of professionals in mid career (The ‘Development of Knowledge and Skills in Employment’ project in the ESRC ‘Learning Society’ research programme 1995-7, Eraut et al.)
So, the theoretical thinking in the LiNEA project did not start from scratch, and although the LiNEA project is heavily qualitative in the research methods it has used, it has not been a small-scale, grounded theory style enterprise. But it is evident from the account just rendered, and Figure 1, that the LiNEA team developed the models as data was being gathered. This interactive process has not yet been fully described in published papers and space here permits only an outline.

The first to emerge were the double triangular relationship models, shown here (Figures 2 and 3) in their most basic form as we envisage them at present. Between them, the two triangles illustrate the personal and the surrounding contexts in which learning happens in the workplace. We started with a well established triangular relationship between Challenge, Support and Confidence derived from the earlier work, and it was the data from the first round of visits that threw up the significance of feedback, both immediate and on general progress, for the learning of the people we were shadowing. In similar fashion it was from the data that we came to realise that those making progress in their learning were those who appreciated the value of the work they were doing, for both their organisation and its clients, and were committed in what they did. However, it was only as data came in from the later visits that we recognised the need to incorporate the influence of a learner’s own personal agency on success in their learning. This last addition was actually agreed after the end date of the project. If one looks back at the sequence of our reporting each year to conferences, it can be seen how the model has evolved as we have seen more data and assimilated it. Those papers are also important because they provide extracts from our data in quotations, vignettes and brief case studies to illustrate the ways our models apply within sectors, and to support our interpretation of the data.

Coding the data for entry into Nud*ist has been a great help because, once the data is entered and coded, it is possible to call up reports that collate the evidence from interviews and visits to illuminate any chosen aspect. In effect, this provides an ability to raise a working hypothesis and check whether the data contradicts or supports it. Although it has to be admitted that, being humans, one is usually looking for support. Of course there is a danger of circularity in this process. But, we believe that with the coding frame having been developed in the way described above, it was developed on a sufficiently separate basis to avoid this pitfall.

Figure 1 – Learning Factors

Challenge and Value of the work

Feedback and Support

Confidence and Commitment

Personal agency

Figure 2 – Contextual Factors

Allocation and Structuring of work

Encounters and Relationships at work

Individual participation and expectations of their performance and progress

The second triangular relationship is completely a LiNEA creation, born out of analytical discussions of what was being observed and what our informants told us.
Our analysis of **what was being learned** also developed from a typology produced by the previous project on mid-career learning. Additional inputs came from our new data and re-examination of the literature on generic competences in management (Boyatzis 1982) and other occupations (Spencer and Spencer 1993). However, in contrast to the common Yes/No or single measure approach to competences, we agreed to treat each category in our new typology (Table 3) as a learning trajectory. The two successive projects enabled us to use many categories for both early and mid-career learning, and this convinced us that the concept of trajectories of lifelong learning fitted our data much more closely than that of a set of competences.

**Table 3 – What is Learned? A Progression Typology**

<table>
<thead>
<tr>
<th>Task Performance</th>
<th>Role Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed and fluency</td>
<td>Prioritisation</td>
</tr>
<tr>
<td>Complexity of tasks and problems</td>
<td>Range of responsibility</td>
</tr>
<tr>
<td>Range of skills required</td>
<td>Supporting other people’s learning</td>
</tr>
<tr>
<td>Communication with a wide range of people</td>
<td>Leadership</td>
</tr>
<tr>
<td>Collaborative work</td>
<td>Accountability</td>
</tr>
<tr>
<td></td>
<td>Supervisory role</td>
</tr>
<tr>
<td></td>
<td>Delegation</td>
</tr>
<tr>
<td></td>
<td>Handling ethical issues</td>
</tr>
<tr>
<td></td>
<td>Coping with unexpected problems</td>
</tr>
<tr>
<td></td>
<td>Crisis management</td>
</tr>
<tr>
<td></td>
<td>Keeping up-to-date</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Awareness and Understanding</th>
<th>Academic Knowledge and Skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other people: colleagues, customers, managers, etc.</td>
<td>Use of evidence and argument</td>
</tr>
<tr>
<td>Contexts and situations</td>
<td>Accessing formal knowledge</td>
</tr>
<tr>
<td>One’s own organisation</td>
<td>Research-based practice</td>
</tr>
<tr>
<td>Problems and risks</td>
<td>Theoretical thinking</td>
</tr>
<tr>
<td>Priorities and strategic issues</td>
<td>Knowing what you might need to know</td>
</tr>
<tr>
<td>Value issues</td>
<td>Using knowledge resources (human, paper-based, electronic)</td>
</tr>
<tr>
<td></td>
<td>Learning how to use relevant theory</td>
</tr>
<tr>
<td></td>
<td>(in a range of practical situations)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Personal Development</th>
<th>Decision Making and Problem Solving</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self evaluation</td>
<td>When to seek expert help</td>
</tr>
<tr>
<td>Self management</td>
<td>Dealing with complexity</td>
</tr>
<tr>
<td>Handling emotions</td>
<td>Group decision making</td>
</tr>
<tr>
<td>Building and sustaining relationships</td>
<td>Problem analysis</td>
</tr>
<tr>
<td>Disposition to attend to other perspectives</td>
<td>Generating, formulating and evaluating options</td>
</tr>
<tr>
<td>Disposition to consult and work with others</td>
<td>Managing the process within an appropriate time-scale</td>
</tr>
<tr>
<td>Disposition to learn and improve one’s practice</td>
<td>Decision making under pressurised conditions</td>
</tr>
<tr>
<td>Accessing relevant knowledge and expertise</td>
<td></td>
</tr>
<tr>
<td>Ability to learn from experience</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Teamwork</th>
<th>Judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collaborative work</td>
<td>Quality of performance, output and outcomes</td>
</tr>
<tr>
<td>Facilitating social relations</td>
<td>Priorities</td>
</tr>
<tr>
<td>Joint planning and problem solving</td>
<td>Value issues</td>
</tr>
<tr>
<td>Ability to engage in and promote mutual learning</td>
<td>Levels of risk</td>
</tr>
</tbody>
</table>

As described earlier, in developing this model, we presented it in ‘prompt card’ style to our key informants during later visits. Although this was done only in accountancy and nursing, all the categories and sub-categories were recognised as valid areas of progression.
Table 4 - A Typology of Early Career Learning Processes and Activities

<table>
<thead>
<tr>
<th>Work processes with learning as a by-product</th>
<th>Learning Activities located within work or learning processes</th>
<th>Learning Processes at or near the workplace</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation in group processes</td>
<td>Asking questions</td>
<td>Being supervised</td>
</tr>
<tr>
<td>Working alongside others</td>
<td>Listening</td>
<td>Being coached</td>
</tr>
<tr>
<td>Consultation</td>
<td>Observing</td>
<td>Being mentored</td>
</tr>
<tr>
<td>Tackling challenging tasks and roles</td>
<td>Getting information</td>
<td>Shadowing</td>
</tr>
<tr>
<td>Problem solving</td>
<td>Learning from mistakes</td>
<td>Visiting other sites</td>
</tr>
<tr>
<td>Trying things out</td>
<td>Reflecting</td>
<td>Independent study</td>
</tr>
<tr>
<td>Consolidating, extending and refining skills</td>
<td>Locating resource people</td>
<td>Conferences</td>
</tr>
<tr>
<td>Working with clients</td>
<td>Giving and receiving feedback</td>
<td>Short courses</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Working for a qualification</td>
</tr>
</tbody>
</table>

Our analysis of how learning was taking place (Table 4) is based on two distinctions: one between working processes from which learning was a by-product and processes with learning as their object; the other between the processes themselves and activities such as asking questions, which can be observed within many different kinds of process.

This way of organising our understanding of our data on ‘How is it learnt?’ crept up on us. Again we had leads from the earlier project, but only one member of the team was common to both projects, and this did not strongly influence most of the team debates. For a long time we discussed, and had prose descriptions of, various ways of learning we had observed or been told about on visits. Much of this discussion was initially sector specific, but it has been a great strength of the research design that we have been able to compare and contrast across three very different sectors when considering how to explain the data and offer models to assist in this. Gradually this led to guarded generalisations, and separation from specific contexts. It was the explicit recognition of the two distinctions, and the casting of the result into tabular form that finally crystallised the presentation.

**Personal thoughts on LiNEA before Section 2**

The LiNEA project team did not embark on its research looking to formulate a ‘Grand Theory’ of learning. What we were trying to do was to explore a good range of learning episodes, intending to improve on how the influences on learning were conceptualised. We were also seeking explanations, and we needed to test any relationships we might suspect were present. We did hypothesise as we worked. This was in order to sharpen the way we interrogated our data. But before we started, like any newly formed research team, we had to go through the ‘storming and norming’ processes to see where we stood in relation to the existing Grand Theories of learning. I am not going to expand on this here because I now need to turn to looking at LiNEA through different goggles.

There are many problems in providing an adequate picture of such a large enterprise in a limited number of pages, or in a restricted time slot at a conference, and I am not referring to this series of seminars where there is more space in paper and time than is normally the case at AERA, BERA or EARLI. The problem is exacerbated because the available conventions of educational research do not provide a standard framework for writing up qualitative, practice-based projects. LiNEA is also an unusually large scale qualitative project. There are no standard frameworks because the range of possible research modes is very wide. But where does that leave LiNEA? Taken at face value, and starting from a tradition I was trained in at Manchester in the 1960s, the description provided allows us to identify a number of interesting features worth ruminating.

Starting with Table 1, that gives the numbers of main informants and interviews conducted, the project team discussed ways of summarising the numbers of organisations, informants, interviews and visits. All of these are countable with relative ease. The table covers two of these categories, and two interesting questions are: what is the appropriate unit of analysis, and what is actually being sampled?

Then there is the question of how our key informants were selected. As often happens, even in school-based research, we asked for volunteers, but if the organisation agrees, the employees can find it
difficult to opt out. So we certainly had plenty of the self-selected, and we probably had a sprinkling of the press ganged, although they could have removed themselves from the cohort after a while, and the evidence is that, with very few exceptions, we only lost people when they moved – and not always then. Whatever the mix, it was not random, but do the contrasting sectors provide “control groups”?

A number of trade-offs are partly buried in the account. Trade-offs occur when, in order to gain access or in order to get some rather than no data, compromises are made. One was when we decided not to be ‘centre spies’ – a term I remember being coined on the evaluations of Schools Council curriculum development projects. This decision was absolutely necessary if our key informants were to agree to work with us, but it meant that judging the rate of progress of our key informants lacked an obvious yardstick that managers could have provided. It was one of the reasons we chose professions with some external mile stone markers: in nursing the grade system, in the other two, progress towards becoming chartered.

A different kind of trade-off is implied in our approach to both interviewing and observing. A similar project designed to check on our findings – now there is a thought – could decide not to use semi-structured interviews and loosely framed observations that did not employ schedules, choosing instead to employ more standardised approaches in the expectation that this might facilitate comparisons across the sectors. We went the way we did because our research design depended on the main fieldworker in each sector having experience in the sector and, therefore, being able to interpret situations and textual data in a high inference mode.

But it is the areas of data management, handling and analysis that a project like LiNEA meets its greatest difficulties in reporting its work. Textual data cannot yet be summarised as numerical data often can be. Its treatment, as it is summarised in accounts or coded inside an analysis package, despite honest efforts at communication, remains largely invisible and has to be taken on trust. Perennial questions surround the use of illustrative material and case studies. Question marks also hover over the assessment of reliability in coding unstructured data. And when it comes to the mysterious spiriting of meaning from the data, how may we compete with the apparent and psychological clarity of a null hypothesis tested at the p < 0.01 level?

However, in my opinion the biggest threat to the quality of the project’s research turned out to be the failure of the funding body to issue the contracts to our two universities in a timely fashion. The knock-on effects on our planned sequences were considerable, and although we had a six month extension, we never fully recovered, especially in the accountancy sector. What makes this so unpalatable is that this inefficiency undermined the project as it had been planned and approved, so the sponsors themselves were working against any value for money they publicly aspired to achieve.

Section 2 LiNEA within Frameworks of Applied and Practice-based Research

Part 1: Classifying the project

The first question is: in what respects may the Early Career Learning at Work (LiNEA) Project be fairly classified as an example of applied or practice-based research? As Furlong and Oancea (2005) point out in their introduction and appendix, there is much variation in the way the links between research, policy and practice are conceptualised. Before criteria can be applied we must be sure the LiNEA project is in its correct categorical box.

The project sits in Phase II of the TLRP, one of the schemes recognised as promoting applied and practice-based research. It is also practice-based in that the research was conducted in professional workplaces, although this alone would not be conclusive. There is other evidence in the project’ proposal to ESRC/TLRP. That made it quite clear that the project’s main preoccupation was not with the various forms of formal training and assessment used by the three professions. Its focus was the
informal, personal and often tacit forms of learning that take place on-the-job. The intention was to
demonstrate that, by increasing our understanding of this kind of learning, rewards could be reaped in
bringing young professionals into higher levels of expertise at a faster rate than was possible hitherto.
In doing so there were likely to be other benefits, better retention in the nursing profession for one.
Two of the three reasons given for the choice of nursing, engineering and accountancy was that they,
‘… play key roles in the UK economy and public services, [and] they use contrasting approaches to
professional formation…’ (LiNEA, 2000:1) And applicable outcomes from the project were
emphasised elsewhere in the proposal. For example, in relation to nursing:

‘For them [the newly qualified nurses], however, there is a danger that their still substantial
learning needs are being neglected by their employers. The United Kingdom Central Council
for Nursing, Midwifery and Health Visiting (UKCC) has picked this up in its recent document,
Fitness for Practice (1999), by proposing that ‘all newly qualified nurses and midwives should
receive a properly supported period of induction and preceptorship when they begin their
employment' (4.64:44). The nature of that induction is very important; the proposed research
will contribute evidence for decision making about its content and emphasis, and also provide
data about longer term learning needs. This is significant in a context where qualified nurses are
now prioritising professional development support when seeking employment. Employers who
fail to deliver this may also fail to recruit, so market factors are likely to force the pace. Hence
our employer partners in the health care sector are also interested in the action research
component.’ (LiNEA 2000:1)

An action research component was also built into the proposal:

‘The action research component of the study involves using early findings from the first cohort
of informants to plan improved learning support for later groups and evaluating the impact of
that change with a second cohort. This will involve an even closer relationship with our
employer-partners than that required for the longitudinal study. Indeed this relationship and our
partnership with two professional associations with responsibility for the qualifications which
trainee accountants and engineers are seeking to obtain is a central feature of the proposal.’
(LiNEA 2000:2)

The first section of this paper provides enough evidence to judge how far the reality matched the
proposal in these respects. However, the most room for debate may well be on the nature of the
project’s outcomes. Are the models too theoretical to qualify as practice-based? Where can one see
their usefulness in policy terms? To reply from a project standpoint would turn this paper into an
exercise in justification rather than a starting point for discussion, so I will resist the temptation.

At first consideration, LiNEA appears to fit the OECD Frascati Manual definition of applied research
quoted in Furlong and Oancea (2005:6), being an ‘original investigation undertaken in order to
acquire new knowledge..., directed towards a specific practical aim’. But it may trip over the
’specific’ wording.

It is much easier to place LiNEA within Pasteur’s Quadrant (Figure 2, below). – or is it? There is no
question about the research being inspired by considerations of use. It was. That places LiNEA
horizontally on the right. Was there a quest for fundamental understanding? Who would want to say,
‘No’? I think that some might sit LiNEA on the horizontal fence between Yes and No. But wait a
minute, the example given of pure applied research is that of Edison. I feel that keeps LiNEA above
the line, but it doesn’t feel quite right. It seems to depend on what one judges is meant by
‘fundamental’. (Incidentally, should the fourth and empty quadrant contain ‘practitioner-led, practice-
based research for the purpose of gaining a higher degree’, or is that too frivolous a suggestion?)

If the meaning of a key word is a matter of judgement, so too would be decisions on the level at which
the LiNEA project’s three research questions are pitched, the distance between its findings and
applicability, and the way in which autonomy and accountability between the project’s different
partners was managed. Judging the distance to application sounds very close to deciding whether the
The LiNEA project was overwhelmingly academic-led rather than practitioner-led, but it was clearly practice-based. However, a problem the project may find itself facing is how successful it can be in convincing others that, although the research took heed of context, and its two triangle model incorporates the importance of particular elements of context, its findings are nevertheless context transferable. Its transdiscipline credentials are respectable, and its research methods have been flexible, with the action research component allowing some crossing of institutional boundaries – at least in my opinion. If the reader does not readily agree with the last two claims, this signals an area of contention that would bear on whether the project has produced ‘Mode 2’ knowledge. I believe it is possible to claim social and economic accountability in the aims of the project and its work. This would strengthen the project’s claim to being applied and practice-based.

All in all, classifying even a project like LiNEA is not going to be a matter of ticking boxes. In these situations I am always drawn to Wittgenstein’s notion of a family of characteristics. Family membership is indicated not by possession of all the characteristics – large hands, a good head of hair, etc., but by either possession of the majority of the characteristics, or possession of a smaller number providing that they include some that are key to recognising the family. This implies that any one individual characteristic will have assigned to it a level of criticality in the decision-making. The notion of fuzzy sets may be similarly applicable.

### Part 2: Judging the quality

In this part of the paper the LiNEA project is examined against the dimensions of quality that Furlong and Oancea have suggested (Op.cit. pp. 11-16).

#### Methodological and theoretical robustness

This dimension has five subsections. The first is ‘Trustworthiness’ and under this term there is mention of both characteristics of the research itself; for example, reliability, groundedness, dependability, believability, plausibility, and authenticity – derived from different epistemic traditions; and characteristics of the researchers, e.g. honesty and integrity. The underlying common concern is judged to be with ‘the strength of the warrants for the relation between the research process and its representation to the world. For me, this comes down to methodological and theoretical transparency. It does not matter which experimental paradigm the researcher works within as long as s/he can communicate both the theoretical underpinnings and the methodological considerations and actions of the research in a transparent fashion.

The LiNEA project, in common with other projects that use predominately qualitative methods, has faced the problem of making its work transparent. The problems centre upon the ways in which such projects make the links between theory, data and conclusions. (These difficulties are quite separate
from issues surrounding technical matters such as sampling and the reliability of coding.) Thus one of the difficulties is how to make the processes involved in the generation and use of a coding scheme visible to those outside the project. This has been covered in Section 1, but has it been adequately transparent for the reader? Section 1 of this paper has relatively little to say on the LiNEA project’s theoretical stance, although some inferences can be drawn from the few references and from the ‘models’ themselves. This could have been rectified, but how long would Section 1 have then become?

In terms of the LiNEA project’s ‘Contribution to knowledge’ it can justly claim to be enhancing conceptual clarity in the field of learning at work on the basis of the three ‘models’ it has proposed, and it would equally claim ‘Propriety’ in its conduct of the research. Propriety, however, is not a black and white matter. It sometimes happens on projects – and I deliberately widen my references beyond LiNEA at this point – that bad things happen. For example, confidentiality can be breached, sometimes inadvertently, sometimes because of unprofessional careless behaviour, sometimes because someone holds particular knowledge. In my career I have come across examples of impropriety other than confidentiality. The standard by which to reach judgement has two necessary dimensions: deliberate or not; and degree of consequent harm.

When we turn to ‘Explicitness in designing and reporting’ we return to the problems of transparency where one of the additional problems for the LiNEA project is that the information it has provided appears in several publications. This year alone the project team will have produced 14 separate papers. Admittedly there is much overlap because similar messages are being given to different audiences, but projects that use qualitative approaches need more text to communicate adequately. They cannot briefly summarise data in a few lines, using mean, modes or medians and measures of dispersion. So either those who wish to conduct reviews must cover several publications to gather an adequate picture of a particular qualitative research, or some other arrangements must be agreed – a thought to which I return later. Thus we run close to needing ‘Paradigm-dependent criteria’ which we can now see have to make the links between theory, data and conclusions transparent, and also deal with the technical methods being employed.

Value for use – the technological dimension

It is heartening to have it proposed that the impact of research should be assessed not only in the short term, but also with consideration of its potential long term value. There is a need for research that has its eyes on the school desk or work bench, and for research with its eyes on the horizon outside the building’s window, but it is a tall order to expect both from the same piece of research. The LiNEA project can already demonstrate a useful short term impact in the uptake and official application of its messages in at least one NH Trust. Dialogue has also begun with professional institutions in engineering. But the project’s longer term influences are predictable only within extensive boundaries of uncertainty. LiNEA’s outcomes have ‘Salience and timeliness’ because they can affect how learners are supported in the workplace and the retention of staff now. I feel that it is the project’s messages about supporting learners that will have the longer term effects, but I realise that the lack of clear marker tags on the models and ideas will make it difficult to trace the take-up and use of the ideas in workplaces.

It is for the reader to judge LiNEA on the sub-dimension of fitness for purpose. ‘Purposivity’ is a word I find hard to empathise with. In respect of the other possible sub-dimensions, LiNEA’s planned action research, and even the way its models have been presented, show a definite ‘Concern for enabling Impact’. The models have diagrammatic or tabular representations that are readily comprehended by potential users outside the educational world. This is also a good mark for LiNEA in respect of the ‘Specificity and accessibility’ sub-dimension. The action research component of the LiNEA project may even qualify it under the ‘Flexibility and operationalisibility’ banner. While it is pleasant to obtain more than one gold star with the same piece of work, this may point to overlaps between the three sub-dimensions that require further thought. It is also hard to see exactly how one might judge ‘Flexibility and operationalisibility’ from outside the project.
Capacity building and value for people

This is a dimension in the judgement of quality that I am glad to see. The LiNEA project has not been, to my mind, highly interactive with its partners, not in comparison with the traditions of work that have aided curriculum reform. Nevertheless the project has some claims under this dimension. It has certainly proved its ‘Plausibility’ from a practitioner’s perspective, and it has stimulated ‘Reflection and criticism’ and ‘Stimulated personal growth’ among our key informants, the newly qualified nurses, the graduate engineers and the trainee accountants. We know this from their own reports.

However, I propose that this dimension should also take into account ‘value for people’ in a somewhat different sense. How far do the outcomes of the project directly help individuals, and enable people to run their lives in a more enlightened way? This kind of question might get closer to a focus on ‘the enhancement of ethically authentic action’. In a sense, what I am proposing is an explicitly moral judgement that some may prefer to avoid. But the authors stated that ‘... the roots of this dimension are in ethical concerns…’ (Op. cit. p.14 col. 1), and the prime value system operating appears to be a moral one, so why hide the fact? Such an approach would be a useful complement, or is that counter weight, to the final dimension that we now come to consider.

The economic dimension

This appears to have value as both a before and after dimension. Before a project can start, potential sponsors make judgements along this dimension, throwing most of the judgmental weightings, one would hope, onto ‘Originality’, ‘Feasibility’ and likely ‘Cost-effectiveness’. An experienced sponsor ought also to take into account ‘Auditability’. One would hope that the weighting assigned to potential ‘Marketability and competitiveness’ could never outweigh originality, feasibility and ‘Desirability’ – the missing sub-dimension. The LiNEA project obtained its funding against fierce competition. In this regard the judgements made on this dimension before the project started had to have been mostly positive.

However, the economic dimension is being proposed for use after a project has ended and reported. When reviewing externally funded researches, this implies that the reviewers set out to second guess the sponsors, at least to some degree. Of course reviewers have the benefit of hindsight. The research may have had problems, failed to complete, altered course, or in some other way changed its nature. Alternatively it could have turned out to be an astonishing success, having caught a wave of contemporary concern with its inherent ‘Plausibility’ and its recognised ‘Contribution to knowledge’. So ‘yes’, reviewers are entitled to revisit these areas, as the reader is entitled to judge LiNEA now. Well, not quite now, because we are still writing the book of the project. But when is the appropriate time to apply these judgements of quality?

There are also other question marks alongside ‘Cost-effectiveness’ and ‘Value-efficiency’. How will these be calculated or assessed? Cost-effectiveness is a commonly used term and a concept bandied about without much scrutiny. It is comparatively easy to measure cost, but effectiveness? How would that be anything other than a heavily subjective assessment? Likewise, knowing what one does about citation behaviour in the academic world, the value of present measures of ‘value-efficiency’ do not strike one as much more reliable than presently available measures of inter-coder reliability.

It is true that I have not jibbed up to now, and many of the proposed sub-dimensions for the judgement of quality in educational research are going to be subjective, within well organised systems to reduce individual biases. However, if these sub-dimensions are to be used, then a need for transparency in how the judgements of the reviewers are decided is essential. The publication of the UK Funding Bodies’ Guidance to Panels demonstrates in principle a willingness to be transparent, and the work of Oancea and Furlong continues this. In the on-going debates around these issues it will remain necessary to keep to this principle.
In conclusion

Having positioned the LiNEA project against the suggested multi-dimensional framework for judging quality in educational research and examined how the proposed dimensions might apply to it, what emerges that is useful for further consideration of the criteria that might be used?

First, deciding whether or not a project may be correctly classified as working in an applied and practice-based form of educational research is not straightforward. The decision on LiNEA is relatively easy, but it is not difficult to envisage projects that would be very hard to classify unambiguously. I am thinking of evaluation and curriculum development projects. Where would they fit? A question that leads me to wonder how far one might extend this kind of assessment of quality. Could one, should one, include official publications that are published and referred to as if they had some research basis.

There may be merit in the use of a flexible system of allocation based on ‘family characteristics’ or fuzzy sets.

Second, if judgements are to be made against the proposed criteria, projects must be transparent in three senses. They must achieve transparency in describing their working procedures and transparency in their accounts of the technical methods they have employed. Neither of these is problematic. But there is also a need for transparency in communicating the way the links have been made between the project’s theoretical stance(s), its data and its conclusions.

This last requirement is by no means easy. Those using predominately qualitative approaches have to find ways of convincing people that they did not start out knowing what their preferred findings would be. Whereas projects that employ predominately quantitative approaches, where it is quite de rigueur to have a precisely stated expectation of what you expect to find, have to guard against creating what I referred to earlier as an apparent and psychological transparency.

Without adequate transparency in these respects, the business of judging quality is seriously hampered. So an overall judgement of quality that seems called for, is one of the level of transparency in the project’s reporting.

A practical problem that the LiNEA project has – in common with most projects using qualitative methods – is that of needing more text to share an adequate description of its data. It is part of the problem that such projects tend to accumulate their findings over time so that reporting is extended across several or many publications. In the age of the web and the age of systematic reviews, there is no reason why projects should not deposit a standard set of information at a nominated web site if they want their work to be eligible for inclusion in systematic reviews. Given a range of agreed headings, projects could lodge information as it became available, extracting it from their published papers if that made life easier.

Third, there are probable, undesirable overlaps in the Technological sub-dimensions. And ‘Flexibility and operationalisibility’ appears to be more visible from within a project than from a reviewer’s seat.

Fourth, under Capacity building and Value for people, I believe there is a case for a more direct assessment of: How far do the outcomes of the research directly help individuals, and enable people to run their lives in a more enlightened way? It is an explicitly moral judgement, but why hide the fact in a dimension with an admittedly ethical substructure?

Last, I have argued for transparency from the researchers, I argue for a continuing commitment to ever greater transparency from the reviewers.
The Early Career Learning at Work (LiNEA) Project was funded from 1st January 2000 to 30th June 2005 in Phase 2 of the Teaching and Learning Research Programme of the UK Economic and Social Research Council. LiNEA stands for Learning in Nursing, Engineering and Accountancy. The project was originally known as the Learning During the First Three Years of Postgraduate Employment Project.

The full research team was:
Michael Eraut, Stephen Steadman and Judith Furner, 
University of Sussex School of Education;
Carolyn Miller, Claire Blackman and Catherine Caballero, 
University of Brighton Institute of Nursing and Midwifery; and
Fred Maillardet and Amer Ali, 
University of Brighton Faculty of Science and Engineering.

Correspondence to: s.d.steadman@sussex.ac.uk

References:
Project Web Site www.sussex.ac.uk/usie/linea/index
*References preceded by an asterisk are available on the project web site.

Interim Reports


Other references


LiNEA Project (2000) Proposal to the ESRC/TLRP


