Understanding the prospects for transformation

by David Nicol, Steve Draper

Introduction

In 2005 the Scottish Funding Council (SFC), then SHEFC, launched the e-Learning Transformation Programme (hereafter, ELTP). The authors of the present paper are part of one of the six successful project teams (Re-engineering Assessment Project). The SFC based the programme on a report (Tripp et al. 2003), and both they and the report's authors were clearly influenced by the Pew programme in the US and by the vision that the sector could be transformed by the application of ICT (information and communications technology) to learning and teaching.

Such visions and initiatives have a long history. For instance Papert (e.g. 1980) preached the transformative power of computation for conceptual development i.e. for a qualitative increase in educational quality. While widely influential as an inspiration, no such widespread increase in educational quality has been manifest. There have been and continue to be many attempts to apply ICT in education. Most have cost more, and most have led to no significant learning improvements, although a few have achieved this. Those that have achieved educational improvements have tended to be led not by technology but by a specific pedagogical idea (Draper 1998). An example of the latter we shall use below is Mazur's Peer Instruction (Mazur, 1997).

Somewhat in contrast to this experience, our society is pervaded by a belief that ICT is a driving force for reducing costs in every field, surely including education (just as first steel, then the railways, then the automobile, and now ICT have driven the economy forward and transformed innumerable aspects of society). For the first time, the Pew programme has offered a concrete demonstration of this presupposed benefit in the tertiary sector by reporting that all 30 of its projects achieved cost reductions while maintaining, and in most cases measurably improving, learning outcomes. Funders cannot afford to ignore the possibility: widening participation, and hence social justice, depend upon it. However most academic instinctively, and on the basis of almost all published literature to date, are reluctant to believe this. One counter argument to the applicability of the Pew programme for UK developments would be that its selection of the 30 most promising out of perhaps hundreds of thousands of HE course teams in the USA represented the tiny number already in a position to change, and does not apply to the vast majority of teaching.

So what are the real prospects for transformation, and what are the features of the Pew programme that might have been important in its reported success? The first step in exploring this is to compare in some detail the Pew and SFC funding programmes. To help in relating them to the most successful past developments, we will also discuss the case of Mazur's Peer Instruction. The general question is what, if any, inferences can be drawn from the Pew programme that would apply to the UK context. The more specific issues are what are the crucial causes or enablers of improvements in learning in teaching (in tertiary education), and hence what policy initiatives might be most effective.

Part A: Comparison and analysis of the Pew programme

Following a brief overview of each approach, we compare and constrast them with respect to a series of issues.

Overview of the three elements

Pew

The Pew programme, directed by Carol Twigg and funded by a charitable foundation, ran 1999??-2002 ar gave $6M to 30 course teams across the USA to introduce redesigned courses. They spanned many disciplines, and institutions of many kinds from community colleges to private 4-year universities. All 30 showed significant cost reductions, and 25 showed significant measured learning outcome improvements. The process was both highly selective, and highly interactive and supportive, so that successful applicants
had by the time of the final award a detailed implementation plan ready to go that satisfied both their own aims and those of the funders. [www.thencat.org]

**SFC**

The SFC E-learning transformation programme [web ref??] runs 2005-7 and awarded £6M to six project teams, each multi-institutional and roughly equally divided between FE and HE. This was run in a way much more typical of of past and indeed current public funding initiatives, with an open call to all institutions and all disciplines, with a detailed specification of selection criteria, but no iterative feedback or proposals. Since it is still in progress at the time of writing, no outcomes are available, while in contrast th Pew programme can now report not only on the funding period but on what happened in the following years to each funded course (and in some cases learning benefits only clearly appeared after the first year).

Whereas the requirements in Pew were the same for all projects, the ELTP had three categories: "The FE-HE transition", "Support for students and promotion of effective learning", and "collaborative delivery of specialist content".

**Mazur**

Mazur’s "Peer Instruction" is used here as an example of successful innovation driven primarily by a pedagogical idea, rather than by funding or technology, that has been adopted in numerous other institutions. Other examples might be Problem Based Learning (PBL) in medical schools, or Just in Time Teaching (JITT) for first year science classes (Novak et al., 1999). The essential idea is to use class time for peer discussion (even in very large classes), structured by a specific recipe based around well designed "ConcepTests" designed to test understanding, and which tend to provoke different views. (See Mazur, 1997; and Nicol & Boyle, 2003 for more details.)

In current implementations this is often supported by technology in the form of Electronic Voting System. It is mainly used in physics classes, and has spread widely in that teaching community. It has been shown (Crouch & Mazur, 2001) to lead to large learning gains compared to previous methods of teaching those classes.

**Grant awarding selection process**

**Mazur**

Mazur’s Peer Instruction approach of course had no selection process: instead it depends (as many e-learning innovations have over the years) on self-selected enthusiasts. A course team leader decides to adopt or invent a new approach, and invests the time, money, and equipment at their disposal in this personal initiative. An important advantage is that that they do not have to have evidence already in existence to convince others, and do not have to spend resources on writing grant proposals.

**Pew**

There have been three rounds of Pew funding with 10 applications funded per round. The first round was conducted a little differently, but in the subsequent two rounds, applicants participated in the three-stage application process shown in Table 1. (Numbers shown are approximately representative of the second round.)

1. The call was open to all 3,400 Higher Education (HE) institutions in the USA, or rather all the (5,000?) teams delivering large courses in all those institutions.
2. 135 applications were received.
3. 40 were selected by institutional readiness criteria. [Twigg ref] For each of these, a three person team (including senior staff from the HEI) attended a one-day workshop on how to produce the next stage of the application: how to redesign pedagogically, how to plan for cost savings and measure results.
4. 40 stage 2 applications were submitted.
5. 20 of these were selected on the basis of course readiness criteria. A core team from each attended a further one-day workshop on developing plans and budgets for the next stage. In addition, considerable individualised assistance on the proposals was given iteratively.
6. 20 stage 3 applications were submitted.
5. 10 were finally selected for funding, with complete plans in place for what would be done, and how outcomes would be measured.

This process is clearly highly selective, structured, and required a lot of unpaid design work by the applicants, but also offered (free) significant amounts of consultancy expertise and work on the designs. required from applicants active and visible involvement by senior management as well as course team leaders.

E-learning transformation programme
A single round application process was used. Open to all 43? Scottish institutions, but with cross institution collaboration required. Considerable detail on the selection criteria was given, including delivery costs e.g. in the embedding and sustainability criteria. Institutions were not required to identify and prove in advance the cost-savings they might achieve (even though the SFC would have been more than happy to see such information in advance).

Commentary
The highly selective nature of the Pew programme, particularly in terms of "readiness criteria", means that any successes are demonstrations of what is possible, but represent little information about whether they could then apply to other institutions. In contrast, a strong underlying theme of the SFC's ELTP is a visic of embedding whereby any successes would be expected to spread both within and across institutions. The self-selected nature of Mazur's approach also applies to its rollout to other institutions: again, others reading about it will have to self-select, although they will be able to do so on the basis of his success (and the published evidence of this), not merely on the basis of faith in their own ideas.

Measuring gains
The two important types of possible gain are cost (reductions), and learning output (gains).

Mazur
The existence of a widely used standard test (the Force Concept Inventory), independent of his course, for the relevant subject area has allowed Mazur and his team to test students both at the start and end of his course; and he did so both before he introduced his Peer Instruction method and for some years afterwards (Crouch & Mazur; 2001). They report a normalised gain now triple what it was with non-PI instruction, using the FCI test.

A disadvantage of standard tests are that his (MIT) students are already scoring highly on them at entry (67-71% pre-test class means; 78% post-test for non-PI, 85-92% as the method iteratively improved), but his calculation of normalised gains attempt to get over that. He does not report costs: presumably the cost of EVS equipment is the only difference.

Pew
Teams were required to report costs on a standardised accounting tool. They also had to agree, as part of their design, how learning outcomes would be measured. This varied between teams. Some used standard tests like Mazur. For others, a measure of student dropout was their focus, or evidence about a shift in quality of learning (e.g. depth of understanding) was collected. An important point however is that the teams had identified what their measure would be in advance, considered whether it aligned with their main aims in the redesign, and agreed it with an evaluation consultant.

E-learning transformation programme
Measures of learning outcomes are more or less required, but their specific nature is being developed during the projects.

Commentary
[Sneak in a self-ref to cost benefit papers? To triangle of cost-quality-time, or cost-quality-quantity?]
The level at which the intervention is focussed

Interventions to improve learning and teaching could be made at many levels: which might be most important?

1. National level: PBL was only introduced to some UK medical schools because of repeated pressur including threats of closure made by the General Medical Council.
2. HEIs. The E-university in England was a failure. The Open University in England introduced a sustained alternative model for a whole university, that is still popular with a large subset of learners. This is definitely an HEI level change: different in the whole organisation, but equally requiring markedly different learner habits. Phoenix University in USA is a big commercial success with a quite different model.
3. Departments. These are the natural level of power in many ways because HE the world over is organised around disciplines (and so departments). Often underdiscussed in e-learning, it is the locus of much that is vital in educational design. For example at Glasgow, dropout rates are much lower in medicine and law, where departments correspond to an integrated set of courses throughout a student's degree. Rea learning constraints often do not correspond to single modules, but to things like a common learning activity being repeated over a degree programme e.g. the essay, the "problem" in physics, learning to use the library and to read the research literature.
4. Course teams. The locus where the major learning design usually is done e.g. the number of lectures, tutorials, labs, exercises etc. is decided.
5. Individual academics. Personal delivery skills remain important e.g. being effective or not at e-moderating. Teaching fellowships and awards operate at this level. Also individuals often add things not prescribed by the team e.g. model answers, personal face to face feedback, ...
6. Individual students. Laptop initiatives: organising one's notes in a digital not written format. Being trained to use mindmaps. Being trained to use reflection in one's learning. Plagiarism. At least by their final year, students are spending perhaps 8 times as much time studying alone as with staff present: what they do in this time is surely much more of an influence on outcomes than what staff do.

Mazur

This style of innovation operates exclusively at the level of a course team or its leader.

Pew

This operated mainly at the level of a course team. Many of the cost reductions were through specialisation of labour within the team (using cheaper staff for some tasks), and major changes in learning activity such as the abandonment of lectures are decisions about a whole course, not about the delivery done by an individual. It should not be forgotten however that the selection process involved senior management: son of the Pew projects involved moving to the use of huge computer labs, and only institutional-level actions could make that possible.

E-learning transformation programme

The ELTP projects vary in the level of their focus. The REAP project revolves mainly around course team its degree of success (like Pew) will depend on how many course teams make successful redesigns. The TESEP project, in contrast, is acting mainly on individual staff, and supporting their individual changes of practice.

Commentary

From a policy point of view, the level of intervention is an important variable: e.g. do you invite students, staff, departments, or institutions to bid. Successful examples can probably be found, or created, at every level.

Partnership arrangements

A key requirement in the SFC transformation programme was that institutions worked in partnership with other institutions from the HE or FE sector. This was not required or allowed within the Pew programme.

One advantage of institutional partners is that the lessons learned at one institution can be shared across if other institutions as the project develops and that good practice can be shared. A second benefit is that partnership arrangements bring forward dissemination processes: the robustness and the transferability of the redesign models can be piloted at partner institutions before being rolled out to across the HE sector. third reason for partnerships is that this helps distribute risks; if one partner fails to deliver then this would
not end the project. Despite these potential benefits there are significant costs to the partnership approach coordinating cross-institutional activities. In effect, such coordination was done in Pew by the central team led by Twigg, while in ELTP significant amounts are done within project teams.

**Institutional or Course Change**

Another difference between the Pew programme and the SFC programme is the level at which transformation was sought. In the Pew programme change was implemented solely at the course team level. Institutions had already demonstrated significant transformational change at institutional level as defined by their readiness criteria. In contrast, a key focus of the SFC programme is on achieving institutional change. This was highlighted by the SFCs use of terms like embedding and sustainability an is evidenced through the reporting framework for SFC funded projects.

One issue raised by these differences in intervention levels is the added burden and complexity associated with driving, coordinating and documenting change activities at two levels (course and institutional level) simultaneously rather than one (course level). In the Pew programme, institutions had to do this themselves, in advance and without assistance, in order to qualify to apply. Thus it required a team approach to course redesign with necessary leadership buy-in (e.g. the dean or head of department) rather than a ‘lone ranger approach’. Research shows that this approach is more likely to lead to institutional embedding although it was not a strong focus of the SFC approach.

**Pedagogy**

By "pedagogy" here we mean the approach to or method of teaching: what Clark (1983) called "teaching method". Clark's argument was that while there was ample literature showing that a change in teaching method could change learning outcomes, all the studies of technology in education either showed no improvement in learning or else had changed the teaching method as well as introducing technology. The are ample grounds for paying attention to this.

**Mazur**

Mazur's method is a pedagogical change: getting students to discuss with each other which answer to a question they think is correct and their reasons for this, as opposed to simply listening to a teacher. This turn is based on the pedagogical idea of learning for understanding as monitored by both learner and teacher with the aid of the questions that probe understanding.

**Pew**

The Pew programme did come with some general pedagogical ideas: promoting active learning, learner-centered provision (i.e. offering differently tailored resources for different students). The impression is, however, that these ideas would be promoted only if the course team did not already have a good idea in mind for their redesign.

**E-learning transformation programme**

The ELTP had as one selection criterion "impact on learning and teaching", although it did not require having a pedagogical rationale for the proposed changes. Different projects in the programme have different rationales. For example, the REAP project is centered around improving assessment practices, especially formative feedback, building on relatively recent ideas from the educational literature (Nicol & Macfarlane-Dick, 2006).

**Commentary**

Course teams often have a fairly clear idea of what they wish could be improved, depending on their circumstances: dropout rate, low exam scores, shallow student understanding, low lecture attendance, etc. They often do not have any idea how to achieve that, as indicated by grumbling but no belief that a solution exists. They certainly do not perceive that this or that equipment would solve their problem. Hence pedagogic ideas, in the sense of educational knowledge that links theory, symptoms of difficulties, and alternative teaching techniques to reduce the problems, is often the crucial missing ingredient preventing improvement.
The questionable role of money as a driver
The paradox here is that many funders expect new methods to save money, in which case surely teachers and institutions should adopt them without any other inducement: offering them money to do so seems a contradiction. So is money useful or effective in promoting worthwhile change?

Mazur
Mazur's technique raises learning outcomes but doesn't save money. It also hasn't required external funding like Pew and ELTP in order to spread. (The same is true of JITT.)

Pew
The Pew programme offered a considerable sum to the course teams selected, and left them entirely free to spend it as they wished. It also required cost reductions in course delivery. This was because at the time, almost no-one believed cost reductions (while raising not lowering quality) was possible, so a motivation for attempting it was required: it could be seen as paying for taking on the risk of change in the absence of good evidence. In the post-Pew era, Carol Twigg is engaged in a new programme for promoting course redesigns without direct payments, supported now by the evidence from the Pew programme that it is a realistic goal.

E-learning transformation programme
This programme also offered considerable money to projects. It will be spent on what each project believes in advance would be necessary. An interesting outcome from the programme, but possibly not easy to obtain, would be frank post hoc reports on what turned out to be worth spending money on in the light of experience. Because project plans were required in advance, projects will feel obliged largely to spend what was promised, and reluctant to assess whether this was useful. In contrast, the Pew programme attached no specific project plans to expenditure of the granted money, thus allowing flexible spending and promoting frankness about what turned out to be useful.

Commentary
Innovations like Mazur's peer instruction show that money isn't always the driver either of innovation or of its spread. In Pew, money was important at a time when there was a high perceived risk because of lack of good evidence. However Twigg is now working on demonstrating that course redesign can now be promoted based on the evidence rather than on financial inducements.

It is possible then that funding directed at creating convincing evidence may be most effective in the long run. This is consistent with some indications within our own project that supplying educational evaluation (as a service) to course teams is more welcome than just money. However the truth may of course be complex. E.g. perhaps money is best at getting teachers' attention, both privately and organisationally (which will gain departments' agreement to allow change); and this in turn then enables redesign to be considered even if it isn't actually essential to the implementation.

Discipline specificity
Higher education is primarily organised into disciplines. Teaching is far more likely to draw examples from how the subject is taught in other organisations but the same discipline, than it is from the same organisation but a different discipline. Consequently an issue is: how likely is an idea or a programme to influence all disciplines or only one. Furthermore, the community that does generalise across disciplines writes in a language that works for it, but few working academics find it accessible. Cross-discipline dissemination seems almost to require a double translation: from the first into educational generalisation, and from that to the second discipline's terms of reference.

Mazur
Based in physics. Still mainly applied to that, with perhaps some spread to other science subjects.

Pew
Deliberately selected applicants across as broad a range of disciplines as possible.

E-learning transformation programme
Similarly meant to apply across the board.
Dissemination: Drawing the lessons from a programme

Mazur
Did his own dissemination. Probably crucial to the spread of his method is his writing a book for teacher besides some journal papers.

Pew
A special feature of this programme is that the programme director organised the writing and publication of reports in a standard format from all 30 projects. Furthermore, she drew and wrote up some conclusions about the patterns in the set of projects as a whole. This took some deliberate effort, since whatever their obligations, not all projects supplied immediately usable material for the standard reports, so significant effort in rewriting was required from the programme centre. The advantage, though, has been considerable. Firstly there are reflections across all 30 projects, where in other cases the only writeups would be from people looking only at their own project, with no-one adopting a wider perspective. This offers a better opportunity for general educational conclusions. Secondly however most projects did some of their own dissemination through their disciplinary channels. Thus both perspectives (discipline-specific, and general educational) were well supported, as can be gleaned from the web site, where each project has a page listing its own (usually discipline-specific) dissemination.

E-learning transformation programme
There is an independent evaluation team for the programme as a whole, which may provide an overview across projects. Writing up within projects is likely to be variable as to whether it is discipline-specific or targeted at the educational literature, depending on the project personnel. It seems unlikely that both will systematically covered. To some extent, project may do their own cross-disciplinary publishing.

Commentary
Dissemination of learning and teaching knowledge, as opposed to research, is an important but neglected issue. As well as the difficulties of cross-disciplinary communication, the drivers of research dissemination are absent from work on learning and teaching.

Part B: Inferences from the analysis of the Pew programme
The Pew programme has established as an existence proof that it is possible to redesign large courses to be delivered more cheaply while maintaining, and in 3/4 of cases significantly improving, learning quality. It did this in 30 out of 30 cases. These were across a wide range of disciplines, and widely differing types of institution from high ranking universities to community colleges. (However these were all large courses: there is as yet no similar demonstration for small courses, and since there are no large courses in FE in Scotland, the applicability for FE is not established.)

To achieve this existence proof, it naturally selected the 30 most likely to succeed at this, from the perhaps 5,000 large courses in the US. We therefore as yet know very little about how widely, across how many courses, similar benefits could be rolled out in the near future. Nevertheless, there are some conclusions we may draw.

It has brought out that almost universally, academics have assumed that this could not be done: that either significant learning benefits were not practicable or that substantial extra money was necessary to achieve them. It is clear that this is false as a general proposition, because now there is a set of counterexamples.

It also showed that you only get these benefits if the redesign is carried through with the desired benefits clearly in mind, and consistently moved towards. In the Pew programme, the redesigns were iteratively developed with a requirement for cost reduction, and a requirement for some particular kind of learning improvement. In the past, redesigns have almost always been carried out with no particular requirement for cost savings, and usually with no focussed requirement for learning improvements: instead many projects have been essentially exploratory: trying out some new method or equipment, but without a strong and specific goal of the effects desired. This was not inherently irrational. Exploration is one important type of research, while careful planning is most productive when applied to repeating established effects. Furthermore, the most dramatic cases of ICT revolution in recent years have seemed to show another story
the internet and web, and texting on mobile phones are both cases of unplanned revolutions where utility was discovered after design, implementation, and deployment had all been done. However the Pew programme establishes that for benefits in education, a deliberate and focussed approach is necessary, rather than deploying a magic panacea whose effects are manifest before the workings are well understood. That is, educational benefits from redesigning courses and employing ICT are not natural or spontaneous, let alone automatic and normal.

Given that it is now a clear possibility, even if the conditions for it are not yet well understood, what might take to promote beneficial redesign of courses? We briefly discuss five possible enabling causes, that could require special action to make beneficial redesigns actually happen.

1. **Motivation** to undertake significant changes and redesign. The Pew programme paid $200,000 per course team. This was probably an essential factor at the time. Now the existence proof of attainable benefits is there, it may no longer be necessary: course teams can rationally hope for benefits independently of outside motivators. Either the hope of saving money which they can retain, or of achieving their own educational benefits e.g. an improvement in student retention, and raising of learning standards, may now be sufficient to motivate a redesign and the substantial effort and risk it entails.

2. **Pedagogical ideas** about what kinds of change can be powerful, and for what effects are important. Perhaps even more important is the capacity to translate abstract educational ideas into the specific discipline and course context of the redesign. There are mixed indications here, but satisfying this condition in one way or another is likely to be an essential factor. It is the redesign of the teaching and learning approach, not the mere presence of technology, that determines learning outcomes, as technosceptics have always argued (Clark 1983), and as was assumed and acted upon in the Pew programme.

3. **Staff development** (rather than money or equipment) could prove to be a strategic factor here. While in very general terms this could be seen as an aspect of [2], the distinction is that [2] is about thinking up ideas to put in a course design, while [3] is about whether an academic has had personal experience of using specific delivery methods that may be required in redesigns e.g. running a course in which there are no lectures, supplying on demand interaction over email without being swamped, ...

4. Moving to a **culture of evidence-based teaching**. Our experience is that many teachers are interested in and responsive to evidence that is reported that is relevant to course design decisions, but do not generate any themselves. This means that their post hoc judgement of their own decisions often rests only on their perceptions of whether the volume of student complaints goes up or down, and whether the perceived quality of exam performances improves or worsens. To move away from staying with habit ("standard practice"), to be comfortable in maintaining and embedding new course designs etc. probably requires systematic defining of what would count as good evidence for and against the success of each given redesign, and embedding its collection into course designs. This in turn will progressively encourage further course redesigns because many more cases will come into existence with a clear evidential message.

5. **Personal dissemination.** In considering adopting some new method, one of the things a potential adopter most values if they can get it, is a direct talk with a person who has done it successfully. But why should such people give such talks? There are no careers in "course redesign" (though there are careers in educational technology: for people who usually do not teach, much less design, courses). The drivers of both supply and demand for research papers are absent for dissemination about teaching. Those who design courses are not required to cite the literature on course redesign to get a course approved (demand creation), nor are they required to write papers about it (supply creation).

**Conclusion**

In the light of the Pew programme we could view the history of ICT projects in HE as divided into three phases: each with characteristic properties, and each with characteristic funding programmes.

Phase 1: (enthusiasts). Driven by technology, and aimed at producing demonstrations of technology in learning and teaching. They showed learning outcomes, but seldom any more than the less technological...
approach they replaced (because these not aimed for). Typically the funding was for technology projects only, thus making technology the single dominant design and funding requirement. What they all got was the use of technology. Usually this led to "no significant difference" i.e. the learning is no worse, but not measurably better, than what it replaced. This approach is by no means dead today, with many funding programmes still requiring the use of ICT rather than requiring learning improvements using whatever methods work best. This naturally leads to projects that achieve the use of technology but not learning improvements, and also strongly suggests that funders do not believe that ICT methods are the best for education, or they would not need to require them. Dropping the topic of "e-learning" and returning to a focus on learning and teaching is the way to move beyond this.

Phase 2: Pedagogy-led. These projects usually improve learning measurably because that was aimed for, often by adding an appropriate technological fix for the identified problem. In this phase, the view is that quality improvements are possible and demonstrated, but cost reduction at the same time is assumed not to be possible. This view is consistent with the generally valid dictum that cost, time, and quality form a triangle, and you cannot improve one without worsening others.

Phase 3: Cost and learning gain led. Aim for cost reduction, and get it. Aim for learning improvement secondarily; usually get it. The implicit claim here is that new technology allows an escape from the cost, quality, time triangle (at its historical levels). The issue is going to be, learning how to use it advantageously: not "how to use it" as individual users, nor as technologists, but as course designers.

References


