Alsop, G., & Tompsett, C. (2007). From Effect to Effectiveness: the Missing Research Questions. *Educational Technology & Society*, 10 (1), 28-39.

From Effect to Effectiveness: the Missing Research Questions

Graham Alsop
Faculty of Computing, Information Systems and Mathematics, Kingston University, Surrey, KT1 2EE, UK
g.alsop@kingston.ac.uk

Chris Tompsett
Learning Technology Research Group, Kingston University, Kingston-upon-Thames, Surrey, KT1 2EE, UK
c.p.tompsett@gmail.com

ABSTRACT

For many researchers and developers in information and communication technology in education (ICTE), the transfer of knowledge from research into educational practice is slow and limited. For most researchers concerned with changing practice, the failure to make a significant impact is attributed to those who practice in education, whilst those in practice see technical research as irrelevant to education. This paper argues that a comparison between research in healthcare and research in education, at one stage disparaged, is informative. Research in healthcare is expected to pass through a number of distinct stages, from small-scale, laboratory-controlled experiments to large-scale trials. Research cannot be integrated within healthcare delivery until these stages are complete. This paper uses this model to argue that intermediate stages between research in ICTE and changes in educational practice are currently omitted or ignored. A 'road-map' is provided to characterize the distinct research questions that should be expected at each stage. Without completion of all of these stages, those in educational practice might argue, justifiably, that there is no warrant to change how education is delivered.

Keywords

Evidence-based practice, Quality of evidence, Effectiveness, Technical innovation, Education-care

Introduction

This paper is concerned with exploring the gap between research in information and communication technology in education (ICTE) and its implementation within educational practice. The paper draws on models and theories of evidence-based practice to suggest that there are a number of distinct phases to be completed before 'research' could be considered sufficiently robust to validate a change in practice. Rather than presuming that practitioners in education are resistant to change, the paper questions whether the evidence that is produced by researchers in ICTE can be seen as sufficient to warrant a change in practice in education.

The paper starts by identifying a set of issues that should underpin the application of evidence-based practice to education. A systems approach is used to draw an analogy between healthcare and (analogously) 'education-care'. There is no presumption that precise standards from healthcare could, or even should, be applied, but key principles are argued to be directly relevant. This approach is then extended to compare the introduction of a new therapy in medicine and the development of a new application of ICT in teaching and learning. This analogy raises questions that are accepted as obvious in healthcare, but which are seldom raised in ICTE, (cf. flip-flop technique in grounded theory, Glaser & Strauss, 1967). Once raised, the validity of these questions does not depend on the analogy itself. If the questions are relevant in the new context (i.e. transfer from research to practice in education), then this is all that is required. From this point the analogy offers a guide to potential answers to these questions, which must be assessed within the new context.

Drawing on central principles from evidence-based policy and practice in healthcare, this paper argues that improving the quality of research itself will remain insufficient to justify the transfer of current research to a change in practice by professionals. Instead each 'new idea' must be taken through additional stages, each of which is characterised by a related, but distinct, research question. Each of these questions must be addressed before researchers have the right to claim that professional practice should change.

From evidence-based policy to evidence-based practice

The first clear exposition of evidence-based policy in healthcare was published by Fletcher and Sackett (1979) as a reassessment of the annual health check in Canada. This study provided the first attempt to collect and assess all the knowledge that could be relevant to designing a programme for wide-scale adoption. The most immediate conclusion of this study was that most of what had previously been treated as 'evidently' true was seldom backed by evidence of sufficient quality to justify the conclusions. Faced with the volume of potential evidence that could be considered, a systematic approach was developed to search for what could be relevant, to assess the relevance of the published research and determine the strength of the recommendations that could be made. In doing this, Fletcher and Sackett identified principles to report the conduct and outcomes of any research on which the decision to license any new treatment is made (Moher *et al.*, 2001; Cochrane Collaboration, 2005).

In education in the UK, as elsewhere, there is a similar view that government policy should be based on evidence (e.g. Blunkett, cited in Taylor, 2004) though there are few who would claim that this is so (e.g. DfES, 2002; OECD CERI, 2002). From this perspective e-learning could not claim to be better. This has revived the debate concerning whether the same standards could ever be applied to education (Hargreaves, 1996; Hammersley, 2002; Simons, 2003) or e-learning (Oliver & Conole, 2003). The debate is not necessarily helpful and can be seen, too easily, as a simple conflict between 'good' and 'relevant' research. For many researchers in education this is treated as sufficient grounds to reject, out of hand, any attempt to compare models of research in education with those in healthcare.

In contrast, a less divisive characterisation of the research in this field has been proposed by Oancea (2005). She separates the educational research community into 'intellectuals' and 'technicians'. Intellectuals are engaged in understanding the complexity of education and in providing tools and results for other researchers in the field to use. The benefits of this research are not intended to be directly applicable to practice and may well seem contradictory, e.g. conflicting definitions of assessment and learning. In contrast, technicians are those who are concerned that the results of their research should be applicable to practice. For technicians it must always be pertinent to ask whether the evidence that they present is sufficient to justify any recommended change in practice. If the research base in a field is insufficient to justify the development of evidence-based policy, then the need for practice to be evidence-based becomes focused on the individual practitioner rather than the institution or government. It is to the technician researcher that this paper is addressed.

The analogy is first drawn at a system level, and is outlined following a soft-systems approach (Checkland, 1999). This allows for the objectives of different stakeholders in the system to be recognised. This analogy is then extended to compare two parallel activities: the replacement of one part of a course with an ICT-based element and the introduction of a new drug. Within this analogy we revisit two key issues raised by Fletcher and Sackett in 1979, and explore two questions that researchers need to address in order to present research that could warrant technical innovation in educational practice. The first is termed the 'quality of evidence'; the second draws critical distinctions between phases of research.

Table 1: Components of CATWOE description.

The customers:	Those who receive the service
The actors:	The people within the organisation that collaborate to deliver the service/achieve the transformation
A transformation:	The change that is created by the system
A Welstanschauung:	The world view that justifies the purpose of the system
The owners:	Those who direct/finance the organisation
The environmental constraints:	The controls that exist outside the system that limit what is possible

The analogy

In this section it is argued that the introduction of any new approach to teaching and learning, on the scale that is expected by researchers in ICTE, is parallel to the introduction of a new therapy within healthcare. Practitioners in healthcare, and their patients, should expect that any new therapy should be demonstrably better than what is

currently used, through well-founded research. Practitioners in education, and their students, equivalently, should expect the same standard of research practice before any new approach is adopted for teaching and learning.

The analogy between the two systems is first established using Checkland and Schole's CATWOE analysis (Checkland & Scholes, 1991, p. 37). This identifies six key components within a system (see table 1).

These are outlined in parallel for the healthcare and education-care systems (see table 2, below).

Table 2: Comparison of Healthcare and Education-care systems.

	Healthcare system	**Education-care system**
Customers:	Members of the general population who present themselves with a specific condition.	Members of the general population who present themselves with a specific educational need.
Actors	The doctors, nurses and other healthcare staff - technicians, administrators, etc.	The teachers, lecturers and other support staff - technicians, administrators, etc.
Transformation:	Restoring the customers to the best possible health through the use of the best possible therapy	Providing customers with additional knowledge through the use of the best possible courses
World view	Each individual wishes to be able to participate in society unimpeded through loss of health that could be improved by available therapy	Each individual wishes to participate in activities in society unimpeded through lack of knowledge that could be gained from suitable courses
Owners	The government who control the overall current finance and future investment	The government who control the overall current finance and future investment
Environmental Constraints	The total cost of the system; therapies that are known to work; the differing effect of applying therapies on individuals; the ability of actors to apply therapies.	The total cost of the system; courses that are known to work; the differing effect of individuals studying different courses; the ability of actors to deliver courses.

(Note: The term 'Customer' is used here as the recipient of the service)

The analogy is primarily established as an equivalence of the 'owners' of each system. In each case the government would claim the right to expect that the overall transformation will match expectations.

Table 3: Extended analogy.

	Healthcare System	Education-care System
Expert Actor	Doctors and consultants from relevant healthcare groups	Teachers/lecturers and tutors
Support Actors	Junior nurses and technicians, (radio-technologists, etc.) …	Librarians, technologists, ICT specialists, support teachers
Resources	Drugs, scanners, databases	Course materials, ICT systems, libraries, repositories …
Options for Expert actors	A range of therapies for patients to follow that are appropriate for treating a specific condition.	A range of courses for students to follow that are appropriate for a specific educational need.
Expected Customer Response	The regime that is expected to be followed by an ill person to optimise the effect of the treatment.	Active participation in teaching sessions, completion of teaching materials, assessment etc.
Owner Controls	Government regulations that control the range of therapeutic treatments that can be used (e.g. NICE guidelines), government targets	Inspection to ensure that the standards achieved meet Government expectations on behaviour and exam performance including, in some cases, conformance to a limited number of 'correct' ways to teach, the national curriculum

Three issues could be viewed as potential differences between these systems. Firstly, customers in healthcare typically elect to be treated, whereas much of early education is compulsory – although attendance does not, of course, imply participation. Secondly, healthcare is typically focused on an individual, whereas delivery is more often group-based in education. Thirdly, healthcare can be seen as a service that restores individuals to, or as close to as possible to, 'normal' health. Education is more typically viewed as supplementing the ability of individuals. None of these differences is as clear-cut as stated and each begs further questions in themselves. Although these differences are noted at this stage, they will remain pertinent only if they invalidate the questions that are raised by the analogy. For now, some patience is requested of the reader to follow the main argument through. Issues that relate to a transfer from medicine to education are considered next in defining the analogy at a more detailed level.

This general context needs to be extended to cover distinctions between the actors and the decisions for which they are responsible (see table 3). This table allows for a more detailed discussion of the resources that are available, the range of choices that are allowed to actors and the control mechanisms that are used by the owners to constrain the behaviour of the actors.

At this level the role of the expert actor matches Eraut's generic model of decision making in professional practice (Eraut, 2003). Eraut identifies three sources of knowledge that a professional will integrate to determine the appropriate choice in any sequence of decisions. These include: relevant knowledge from research; specific, case-related information; and professional expertise gained through experience and previous practice. When evidence-based policy is available, it could be argued that this mediates between the volume of knowledge from research and the practitioner. However, this does not reduce the relevance of the other components. Neither does it remove the need to reassess any evidence within the specific context of each decision - as Jadad notes with specific reference to evidence-based practice in medicine (Jadad, 1996).

Table 3 makes no assumptions regarding the extent of ICT integration within any course. It can cover the range of technology-free systems, blended learning, open learning and, potentially, systems that are devoid of human intervention in their delivery (cf. Atkinson's argument, 1993 on the nature and location of a library resource). The replacement of one part of a course with an ICT-based element is parallel to the introduction of a new drug. Just as there can be no presumption that a drug is 'better' because it is new, there can be no presumption that any specific level of, or novelty in, ICT must be better. The next stage in the analogy extends the context to a framework in which a practitioner is expected (by the researchers) to change practice.

Interpreting the evidence-base

In order to provide the necessary link between research in ICTE and change in educational practice, we revisit two specific issues that are taken from the Canadian study. We discuss the first of these, levels of evidence, in terms of principles, rather than the precise standards that have evolved. The second is the distinction between efficacy, effectiveness and efficiency. Both of these issues are critical for the transfer of research into practice. The first addresses the extent to which conclusions could be determined by the context of the original research. The second addresses the issue of whether research is sufficiently mature to warrant a change in practice.

These two issues are considered in turn, firstly as they apply to the healthcare system, and then as they might apply to education-care.

Reliability in the research evidence

Fletcher and Sackett (1979) developed the concept of evidence-based practice as a systematic approach to evaluation of research and the formation of governmental policy on the yearly health check. Within this framework, each research report or evaluation is categorised in terms of the level of evidence that is inherent in the 'design' and conduct of the research, and the strength of recommendation that it would support. Five levels of evidence are normally recognized in medicine (see for example, OCEBM, 2004). These range from level 5, the lowest level of support - expert opinion, to level 1*, the highest - a meta-analysis of homogenous randomized controlled trials.

Few would suggest that these same standards of evidence can be transferred directly into education. A limited number of experiments have been conducted at level 1 (see for example, Toroyan *et al.*, 2003), but randomly controlled trials in education are difficult to construct for a range of philosophical, ethical and pragmatic reasons. Although some of the philosophical limitations might only affect the 'intellectualist' model of research (Oancea, 2005), ethical and pragmatic limitations are hard to avoid (e.g. providing additional resources, testing a 'better' model, crossover between subjects etc.). There is, anyway, a danger in treating the explicit levels of evidence, as more critical than the conceptual understanding of the issues that gave rise to them (cf. Wenger's comments on communities of practice in Wenger, 1998). Technician researchers in e-learning must still account for the same issues, whichever model of research is applied.

Levels of evidence and the management of potential bias

The seminal work that underlies the levels of evidence is almost universally attributed to Fisher's conceptual model of experimental design and analysis – magic squares, the randomised controlled design (RCT) and statistical significance (Fitz-Gibbon, 2003). Randomisation is, of course, central to the statistical analysis of outcomes against a null hypothesis.

However, randomisation has a far wider significance, even if this is often unrecognised (Imbens, 2002). As Fisher points out (1990, p. 17-21), randomisation limits the influence of unpredicted and/or unknown factors on the experiment. Level 1 (randomised controlled trials - RCT) research provides a standard for research, which removes the need for the research to be repeated within the same context. Each of the other 'levels' of evidence (levels 2, 3, 4) of experimental design can then be seen as increasing the potential for the outcomes to be confounded by factors that are external to the experiment, or an inherent part of it (see table 4 below).

Table 4: Sources of bias in different experimental settings, levels 2 to 4

2	Cohort Study	No randomisation before study: unknown factors may influence allocation to control and intervention group
3	Case-controlled study	Cases are identified 'post-hoc': data collection is subject to recollection bias; control group is matched to the intervention group; data collection may be biased towards the intervention group.
4	Case series	No control group – data biased to individual cases

The inability to conduct an RCT does not invalidate research at other levels. The difference does not lie in the validity of the results per se, but in the scope with which the results can be transferred to different, future contexts. As noted above, evidence-based practice requires that evidence, even at level 1, must be assessed for validity within the context of each particular decision.

If research in ICTE by 'technicians' is expected to feed evidence-based practice in education, then an assessment of the limitations inherent within the context and process of the research is required, unless a methodology exists that is specific to education and inapplicable to healthcare that obviates this. Only the design research methodology (Collins, 1999) (previously called a design experiment, Brown, 1992; Collins, 1992) could claim to be relevant.

This methodology attempts to marry the two modes of research that are suggested by Oancea. Learning theory informs experimental design and evaluation of experimental design provides feedback to learning theory through a series of cycles. This rejects the controlled laboratory model of scientific research in favour of collaborative development between researchers, designers and practitioners, conducted within the complexity of social and educational interactions that take place in natural learning environments (as echoed in MacFarlane, 2004). Although the cyclical development has strong parallels with action research (Laurillard, 2004; Lewin, 1946), the end point is both a well-designed but adaptable solution, and a more-refined intellectual model. The process is focused on innovation, rather than problem solving and intended to be data rich, in order to create the volume of data that is required to build complex models and to establish an archive of data for others in the wider community to access. Barab and Squire (2004) raise concerns over a methodology which confuses the role of the researcher as both designer and evaluator of that design in practice. Their recommendation is that design science should follow good

practice in qualitative research in order to transfer richly contextualized knowledge to a more detailed, theoretical model. However, instances of this methodology are few and isolated.

Brown's initial design experiment (Brown, 1992) was extended with Campione into the 'Fostering a Community of Learners' for integrating groups of pupils across a number of years in primary education (Collins *et al.*, 2004, pp. 14-21). Joseph's 'Passion Curriculum' evolved over a similar length of time (Collins *et al.*, 2004, pp. 21-27) and is based within a similar mixed-age context. Other instances appear selective in the principles of design research that are followed (e.g. in language learning, Hoadley, 2002; 2004). Where practiced, design research highlights the need to distinguish the roles of designer, practitioner and researcher. It may orient some of the research outcomes towards the practitioner but it emphasizes, rather than removes, the responsibility of subsequent practitioners to assess the relevance of that work in the context of their own practice. Evolutionary design and extended testing in practice, without the ties to theoretical models, may be equally effective in providing suitable evidence for evidence-based practice (e.g. Language, Truth and Logic, Barwise & Etchemendy, 2002).

Irrespective of the model of research that is adopted, any factor that remains consistent within the research has the potential to restrict the validity of the outcomes in a different context. Similarly any conflict of interest, or loss of independence, between the researcher, the designer, and any practitioner involved in the research may restrict the range of contexts across which others might apply the outcomes. Whenever research is published as relevant to practice, it is incumbent on the researcher to acknowledge and address these factors.

Relevance to practice

The second component from Fletcher and Sackett's work is based on the distinctions described in terms of seven 'research priorities'. The distinctions between the first three will be taken into account, the remaining four are associated with issues that do not translate in the analogy that is being drawn. The first three, taken with minor adaptation from the original paper, were between efficacy, effectiveness and efficiency (1979, p. 1202):

- *Efficacy*. This area is reflected in the question: Does X lead to a better outcome *among those who follow the subsequent advice*?
- *Effectiveness*. The relevant question is: Does X benefit *those to whom it is offered*?
- *Efficiency*. The relevant question is: Is the effective maneuver [sic] being made available to those who could benefit from it with optimal use of resources?

(In the original article *X* was 'early detection of a condition or risk factor').

A second classification of similar issues is now conventional in the context of developing and testing a new drug (Jadad, 1998). Jadad identifies four phases (see Table 5 below).

Table 5: The four phases of research mapped to research priorities

Phase	Statement	Summary
1	Can a new drug be shown to have an effect?	(Effect)
2	Can a new drug be shown to have an effect within an RCT study with a selected population that takes the drug properly?	*Efficacy*
3	Can a new drug be shown to have an effect within an RCT study within a typical population that behaves normally?	*Effectiveness*
4	Are there any previously unobserved outcomes post release (case series)?	Side-effects

Although the two approaches are not identical, the similarities are clear across phases 2 and 3 (italicized) and this parallel is central to the discussion that follows. This represents a critical phase in validating the transfer from 'correct usage' to 'real life'. In phase 2 it is expected that any treatment will be applied *in the optimum conditions* in order to emphasize any potential distinction between the two experimental conditions. The conduct of the experiment will, typically, have additional resources and may well have 'committed' participants (such as those discussed in Parsons, 1974) or, in ICTE, provide little separation between the design of resources and the use of materials (cf. the definitions in Koper *et al.* (2004) of levels of reuse). These factors limit the potential to generalize results to normal practice. The key function of phase 3 research is to demonstrate that efficacy can be generalized to conventional

practice. If this is a critical step in warranting a change in healthcare practice, there seems to be no justification to eliminate this in education-care practice. To ignore this must imply either that such experiments will always succeed, or that they will always fail!

For the remaining phases, the first phase would appear to be a simpler level than considered by Fletcher and Sackett (1979). On the other hand, efficiency is not explicitly represented in the four phases, although it is a key consideration in any soft-system analysis. Even if a change is effective for all participants in the system, efficiency is often assessed in distinct ways. In healthcare, efficiency for a drug company is not the same as efficiency in healthcare policy. In education-care, on-line communities may be more efficient than face-to-face discussion groups from an institutional perspective, but are not necessarily more efficient from a teacher/lecturer or learner's viewpoint.

Table 6 (below) shows the combined set, integrated into a quasi-linear scale. Alongside each one is the appropriate healthcare question, and an analogous question for education-care. Since 'efficiency' in healthcare is not included within Phases 1 to 4, the education-care questions are labelled A to E.

Table 6: Distinct research questions in healthcare and education-care

	Healthcare System	**Education-care System**
Effect:	Phase 1: Can a new drug be shown to have an effect?	Phase A: Can a new ICTE part of a course be shown to (1) have an effect (2) have an effect on learning, within a limited number of students in advantageous conditions?
Efficacy	Phase 2: Can a new drug be shown to have an effect within an RCT study with a selected population that takes the drug properly?	Phase B: Can a new ICTE part of a course be shown to have a positive effect on learning across a suitably large, selected range of students who study properly?
Effectiveness	Phase 3: Can a new drug be shown to have an effect within an RCT study within a typical population that behaves normally?	Phase C: Can a new ICTE part of a course be shown to have a positive effect on learning across a suitably large range of students where no control is maintained on how it is used?
Efficiency	The relevant question is: Is the introduction of an effective drug being made available to those who could benefit from it with optimal use of resources*?	Phase D: Does the introduction of a new effective ICTE part of a course with a limited set of resources, for a specific group of students represent the best use of resources?
Side-effects	Phase 4: What otherwise unknown side effects result from full-scale? use of a drug	Phase E: What otherwise unknown side effects result from full-scale use of a new ICTE component in a course?

What is gained?

The analogy has led us to identify two issues relating to the quality of research that could claim to justify a change in practice in ICTE. The first is the need for a proactive recognition of the limiting contextual factors in reporting research outcomes. The second is the categorisation of research into a number of phases. If a change in practice within healthcare is not warranted without a sufficient volume of contextually relevant evidence at phase 3, then the implication is that changes in practice in education-care, or more specifically in the use of ICTE in education-care, are not warranted until sufficient evidence at phase C is available. In order to clarify this, we outline the key distinctions between the first four phases in education-care. (The final phase, phase E, would be expected to follow a 'case series' model.)

In characterizing each of these phases below, the intention is to distinguish the goals, the context and the key participants within the research and to remain inclusive in terms of the research that could be classified. These definitions anticipate that the research methodologies will change from phase to phase, removing some artificial conflicts but without adopting a prescriptive view. Each description is structured as a series of constraints on the original CATWOE description complemented with two additional categories: 'D' identifies the designer(s) and 'R'

identifies the active researcher(s). In each case the research outcomes are expected to be directly relevant to the transformation (T) and reflect the Weltanschauung (W). Distinctions between phases A to C are given prominence.

Proposed Definitions

Phase A can be characterized as a demonstration that some technology has a measurable effect on learners, either indirectly in terms of some behaviour that might be considered as leading to improved learning (A.1) or as measured in terms of existing learning outcomes (A.2). This work is likely to be small scale, with a restricted set of well-motivated students (e.g. within a single course module). The research is likely to be conducted by one or two self-motivated teachers/lecturers who are either the researchers/designers or are closely related to the research process. This research is expected to be conducted under 'research positive' conditions: the research is likely to be funded explicitly, or indirectly supported in terms of additional time available and/or personal commitment to establishing positive results, etc. (see table 7).

The distinction between A.1 and A.2 is made to highlight research that can demonstrate improved learning. Phase A research is dominated by the introduction of new technology (e.g. interactive video disks, hypermedia, Internet, reusable learning objects), even though, all too frequently, the educational content and the learner's experience remains unchanged. It is accepted, as one reviewer noted, that the introduction of new technology is almost certain to produce some indirect effects, but removing A.1 from the classification would exclude a considerable volume of research in ICTE that is already published.

In addition, if the practitioner is to be convinced that they should change their own practice, it is essential that the impact on learning (A.2) is established rather than presumed (or artificially created by changing the criteria by which learning is established, MacFarlane, 2004).

Table 7. Summary of research in phase A

Phase A: Effect	
C	A selected set of students (e.g., in a single course module)
A	A small set of well-motivated teachers/lecturers/technologists related to the research
T	Integration of technological components in an educational environment
W	Introducing educational technology is reliable
O	Owners provide funding to cover the costs of conducting the research and development of educational materials
E	In constrained conditions
R/D	May be/include any one (or more) of the actors

Table 8. Summary of research in phase B

Phase B: Educational efficacy	
C	A selected set of students on a number of courses
A	A small set of well-motivated teachers/lecturers/technologists
T	Introducing a specific e-learning component into a well-defined set of courses
W	The introduction of (new) e-learning, under suitable conditions, leads to an improvement in educational outcomes in a clearly-defined range of educational contexts
O	Provide additional resources to cover the costs of conducting the research and development of educational/support materials
E	A well-defined set of conditions
D/R	Should not be one of the actors

There remains an inherent risk, in research classified as phase A.2, that improved learning results from a shared understanding between the designer and actor(s), rather than from factors that are inherent in the design of the resources when used appropriately. Phase B research must address this issue and demonstrate that educational

improvements can be reproduced reliably within a wider, but well-defined, educational context. This phase must be able to establish the relationship between design, good practice and educational outcomes, subject to conditions defined by a suitable set of contextual variables. Care must be taken to separate use 'as considered appropriate by the practitioners' in this phase from use 'as required by the designer', which should only be included in phase A. It is expected that the context will still be conducted in resource positive conditions, but the allocation of resources should be tracked and quantified where possible, in order to identify the conditions for use in subsequent trials where additional resources are not available (i.e., phase C). It is expected that the students will be selected from those that are most likely to take advantage of the resources that are available (see table 8).

At the end of this phase it should be rational to claim that new resources could be of benefit, across a well-defined range of contexts, as long as those who are using them are suitably motivated and/or rewarded. Successful research at phase B validates research for phase C.

Phase C must demonstrate that local 'efficacy' can be reproduced as effective practice once the advantages of research in earlier phases have been removed. Since the intention is to explore the reliability of the relationships that are established in phase B, it is critical that the practitioners and the students are a fully representative sample of those who would be expected to use the resource within its anticipated target context. All aspects of the design, including support materials, should be fixed at this stage. Although limited resources may well be necessary to support the research, their use should be to enhance evaluation rather than any aspect of 'normal' use. It is likely that the degree of educational improvement will vary to a greater extent than in phase B. Although it might be tempting to presume that minor changes in the materials will lead to further improvements, such an assumption cannot be validated without further trials (see table 9). It is more rational to recognize such variations as limitations on future use and to investigate the variation in phase D.

It is only on successful completion of research at this stage that it becomes plausible to suggest that other practitioners could replicate these outcomes in a matching context.

Table 9. Summary of research in phase C

Phase C: Educational effectiveness	
C	Typical cohorts of students selected from courses within well-defined subject areas
A	Normal teachers/lecturers/technologists, with no additional support
T	Introducing specific e-learning components into a well-defined set of courses under typical conditions
W	The introduction of e-learning is an effective approach to education
O	Provide additional funding restricted to cover the costs of evaluation
E	In a conventional setting augmented with proven technology
D	Should be neither an actor nor a researcher
R	Should be neither actor nor designer

Phase D research covers the transition from 'effectiveness' to efficiency and, from a systems' perspective, must account for the separate viewpoints of all the stakeholders. The systems approach would anticipate that each of the stakeholders would seek to exploit the new resources as efficiently as possible, but that this may be to the disadvantage of other stakeholders. From this perspective, the system will never be both stable and efficient for one stakeholder, if that requires an unacceptable disadvantage to other stakeholders. By implication, research in phase D requires completion of research at phase C.

Table 10. Mapping between factors and relevant phases

Factor	**Influences**
Self-selection /required attendance	Should be addressed directly – phase C.
Individual/group	a. As group-based learning – introduced in phase A2 b. As efficiency – phase D
Repair/improve	a. In compulsory education – phase C b. In non-compulsory education – phase D

With these definitions, it is possible to return to the three factors that were identified in establishing the analogy earlier in the paper. Each of these factors may reflect a distinction between the healthcare and education-care systems, but each can be addressed within a specific phase(s) of research (see table 10). Separating relevant research into phases focuses these factors into distinct questions, rather than weakening the analogy.

Conclusion

Research in ICTE that is relevant to practice in education has been characterized as falling into five phases, where each phase addresses a distinct set of questions. It is argued that unless research at phase C – effectiveness – has been completed successfully, rational practitioners, acting on the evidence-base, should not be expected to change their practice. Demonstrating that the introduction of novel technology produces some changes in student behaviour, or researching efficiency gains from a management perspective without prior research to ensure effectiveness within the learning process, is irrelevant to evidence-based practice. Even if the research establishes efficacy in a small set of advantageously selected contexts, then adoption for use with typical students, and no additional resources, whilst of potential interest, remains unjustified.

For the technician researcher, associating a phase with each research project brings a number of advantages in addition to understanding why practitioners are justified in ignoring research at levels A, B. Suggestions that any single method of research should be paramount become illogical, or at the very least, localized to each phase. Arguments regarding the relative merit of different research methods can be replaced with consideration of the relationship between, and interdependence, of the research questions. In a similar vein, whenever research is intended to be relevant to practice, due attention should be paid to clarifying the factors that constrain or potentially confound the applicability of any outcomes (cf. BMJ Publishing Group, London; The Lancet, Elsevier, London). If a practitioner is to change their practice, then the identification of contextually relevant research is paramount. Failure to acknowledge such constraints within published research invites the evidence-based practitioner to infer that the researcher was not aware of some factors that are significant to practice.

If the research community wishes to research into the failure of practitioners to change their practice, it may first be important to ensure that the community of 'technician' researchers in ICTE changes their own practice. If this were the case, then the rational practitioner might have a contextually relevant evidence-base from which to proceed!

Final Comment

Some might still argue that it is impossible to transfer expectations from medicine to education and that the field should not aspire to such standards. For those it is perhaps sanguine to compare the standards that are now expected for research in healthcare with the assessment from Fletcher and Sackett of published research at the time:

> "Most of our recommendations have been … based on grade III evidence (professional consensus or the opinion of experts). Even evidence from cohort studies and case-controlled studies was infrequently found, and many of the reports concerned uncontrolled series and at times were just case reports. Opinion or evidence from nonexperimental studies is a much less satisfactory basis for recommendation". (1979, p. 1202)

It would be encouraging to believe that we could see the same improvements, across a similar period of time, in linking research in ICTE and practice in education!

References

Atkinson, R. (1993). Library functions, scholarly communication, and the foundation of the digital library - laying claim to the control zone. *Library Quarterly, 66* (3), 239-265.

Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences, 13* (1), 1-14.

Barwise, J., & Etchemendy, J. (2002). *Language, Proof and Logic A logic courseware package*, Stanford, California: CSLI Publications.

Brown, A. L. (1992). Design Experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2* (2), 141-178.

Checkland, P. (1999). *Soft systems methodology: a 30-year retrospective*, Chichester: Wiley.

Checkland, P., & Scholes, J. (1991). *Soft Systems Methodology in Action*, Chichester: John Wiley and Sons.

Cochrane Collaboration (2005). *Cochrane Collaboration statement on registering clinical trials prospectively*, The Cochrane Collaboration, retrieved 15 December 2006 from http://www.cochrane.org/news/articles/2004.07.26.htm.

Collins, A. (1992). Toward a Design Science of Education. In Scanlon, E. & O'Shea, T. (Eds.), *New Directions in Educational Research,* New York: Springer-Verlag.

Collins, A. (1999). The Changing Infrastructure of Education Research. In Langemann, E. C. & Schulman, L. S. (Eds.), *Issues in education research*, San Francisco, California: Jossey-Bass.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and Methodological Issues. *Journal of the Learning Sciences, 13* (1), 15-42.

DfES (2002). *Research and Development in England: Background report prepared for the OECD review*, London, Department for Education and Skills.

Eraut, M. (2003). Editorial. *Learning in Health and Social Care, 2* (3), 117-122.

Fisher, R. A. (1990). The Design of Experiments (8th Ed., 1971). In Bennett, J. H. (Ed.), *Statistical Methods, Experimental Design and Scientific Inference*, Oxford: Oxford University Press, 872.

Fitz-Gibbon, C. T. (2003). Milestones en route to evidence-based policies. *Research Papers in Education, 17* (4), 313-329.

Fletcher, S., & Sackett, D. (1979). The periodic health examination: Canadian Task Force on the Periodic Health Examination. *Canadian Medical Association Journal, 121*, 1193-1254.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: strategies for qualitative research*, London: Weidenfeld and Nicholson.

Hammersley, M. (2002). *Educational Research Policymaking and Practice*, London: Paul Chapman.

Hargreaves, D. H. (1996). *Teaching as a research-based profession: possibilities and prospects*, London: Teacher Training Agency.

Hoadley, C. (2002) Creating Context: Design-based research in creating and understanding CSCL. In Stahl, G. (Ed.), *Computer Support for Cooperative Learning*. Boulder, Colorado: Lawrence Erlbaum Associates.

Hoadley, C. (2004). Methodological Alignment in Design-Based Research. *Educational Psychologist, 39* (4), 203-212.

Imbens, G. W. (2002) *Fisher's Exact Test in Completely Randomized Experiments*, retrieved December 16, 2006, from http://www.econ.ucla.edu/hotz/e232d/ readings/chap4.pdf.

Jadad, A. R. (1996). Are you playing evidence-based medicine games with our daughter? *British Medical Journal, 347*, 274.

Jadad, A. R. (1998). *Randomised Controlled Trials*, London: BMJ Books.

Koper, R., Pannekeet, K., Hendriks, M., & Hummel, H. (2004). Building communities for the exchange of learning objects: theoretical foundations and requirements. *ALT-J, 12* (1), 21-35.

Laurillard, D. (2004). E-learning in the Knowledge Economy: the right context for innovation. *Colston Colloquium,* 23-24 March 2004, Bristol, UK.

Lewin, K. (1946). Action Research and minority problems. *Journal of Social Issues, 2* (4), 34-46.

MacFarlane, A. (2004) E-Pedagogy - Conference Review. *ALT-C 2004*, September 14-16, 2004, Exeter, UK.

Moher, D., Schultz, K. F., & Altman, D. G. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet, 357* (April), 1191-1194.

Oancea, A. (2005). Criticisms of educational research: key topics and levels of analysis. *British Educational Research Journal, 31* (2), 157-183.

OCEBM (2004) *Levels of Evidence, Centre for Evidence-Based Medicine*, retrieved December 16, 2006, from http://www.cebm.net/downloads/Oxford_CEBM_Levels_5.rtf.

OECD CERI (2002) *Educational Research and Development in England*, Paris: OECD Publishing.

Oliver, M., & Conole, G. (2003). Evidence-based practice and e-learning in Higher Education: can we and should we? *Research Papers in Education, 18* (4), 385-397.

Parsons, H. M. (1974). What happened at Hawthorne? New evidence suggests the Hawthorne effect resulted from operant reinforcement contingencies. *Science, 183* (March), 922-932.

Simons, H. (2003). Evidence-based practice: panacea or over promise? *Research Papers in Education, 18* (4), 303-311.

Taylor, R. (2004). Evidence-based policy: the role of randomised controlled tests. *Randomised Trials in Education Research*, Bonhill House, London: ESRC RCBN, Cardiff University.

Toroyan, T., Roberts, I., Oakley, A., Laing, G., Mugford, M., & Frost, C. (2003). Effectiveness of out-of-home daycare for disadvantaged families: randomised controlled trials. *British Medical Journal, 327*, 906-910.

Wenger, E. (1998). *Communities of Practice: Learning, Meaning and Identity*, Cambridge: Cambridge University Press.