Teaching and Learning RESEARCH BRIEFING

June 2003 Number 1

Towards Evidence-based Practice in Science Education 1:

Using diagnostic assessment to enhance learning

Much research has been carried out on students' understanding of key science ideas, but this has not led to marked improvement in teaching and learning. As a means of improving practice, banks of diagnostic questions, based on research, were developed for several core science topics. These were used to monitor students' understanding of key science ideas, and to explore how the provision of research-based materials of this sort can influence teachers' practices and students' learning.

- Science teachers' practice, and students' learning, can be significantly enhanced by providing teaching materials that embody research findings and insights.
  → If findings and insights from research are 'translated' into specific practical implications, or teaching materials, the likelihood and scale of their impact on practice is greatly increased.

- Carefully designed probes, based on research, can provide quality information on students' understanding of key science ideas, and inform subsequent action.
  → Tools for quickly 'measuring' understanding of key ideas can help focus learning activity and indicate levels of understanding across a class. More should be developed and made available to teachers.

- The level of students' understanding of many fundamental science ideas is low, and increases only slowly with age.
  → Current teaching approaches in science do not result in widespread understanding of many core ideas. Levels of understanding of a few key ideas should be monitored systematically over time, to inform curriculum decisions.

The research

The EPSE Network

This project is one of four undertaken by the Evidence-based Practice in Science Education (EPSE) Research Network. The Network is a collaboration involving the Universities of York, Leeds, Southampton and King's College London. Its overall aim is to explore ways of enhancing the impact of research on practice and policy in science education, by improving our understanding of the interface between researchers and practitioners. The EPSE Network has developed and evaluated several examples of evidence-informed practice, and has explored practitioners' perceptions of the influence of research on their practice. Whilst focussing on science education, the findings and outcomes may also illuminate the research-practice interface in other subject areas.

Background and rationale

Over the past 30 years, a great deal of research has been carried out in many countries on learners' ideas about the natural world. This has helped identify commonly-held ideas which differ from the accepted scientific view, and has shown that these are often very resistant to change. These findings have clear implications for the pace and sequence of instruction in many science topics, particularly those which involve understanding of fundamental ideas and models. Yet whilst many teachers know of this research, it has had little systematic impact on classroom practices, or on science education policy in the UK.

One response might be to try to communicate more effectively to teachers the *findings* of research on science learning. This, however, seems rarely to be an effective strategy for changing educational practices. Instead, this project sought to influence teachers' practice by providing them with *tools* derived from research – a bank of diagnostic questions, based on the kinds of probes used by researchers. This would enable them more easily to collect data on their own students' learning of key ideas, and so adjust their teaching in the light of this evidence of their own students' learning.

There is also another sense in which this approach can be seen as an example of evidence-based practice. A review by Black and Wiliam (1998) of research on the use of formative assessment showed that this can lead to significant learning gains, through clarifying learning objectives, and providing informative feedback to learners. A possible obstacle, however, is the difficulty of generating good formative assessment questions and tasks. By providing teachers with a large bank of diagnostic questions, we can explore the extent to which this facilitates change in practice, leading to better learning.

		Q2 current the same	current gradually used up	other	TOTAL
Q1	current the same	65	46	4	115
	current gradually used up	11	63	4	78
	other	6	7	4	17
	TOTAL	82	116	12	210

Table 1 Models used by Key Stage 3 pupils in two questions on electric current

What we did

The first stage of the project was to develop banks of diagnostic questions. This was done in collaboration with a 'partnership group' of practitioners, which included several primary and secondary teachers, LEA advisers, and writers of teaching materials. This group chose the science topics for this work: electric circuits, forces and motion, particle models of matter, and life processes (digestion, respiration and photosynthesis), and acted as a critical 'expert' review group throughout. Diagnostic questions were trialled in the partner teachers' schools, with follow-up interviews with some students to probe their understanding further.

Questions with a two-tier structure (Figure 1) were found to be particularly useful. The first 'tier' asks for a prediction about what will be observed – and the second then asks for the best explanation for this. A correct answer combination is a good indicator of understanding. Other answer patterns indicate common misconceptions. Questions were also developed with more open response formats, and to stimulate small-group discussion.

The banks of diagnostic questions produced were then used in two ways:

- to evaluate students' current understanding of some key ideas at the ends of Key Stages 2, 3 and 4;
- to monitor how a panel of teachers used diagnostic question banks in their own teaching, and the impact on classroom activity and student learning.

Understanding of key ideas

Rather than assess understanding across the sciences, as other surveys have done, we chose to probe students' understanding of a few ideas in depth. The samples of students tested came from a range of schools, and were above the national average for each age group (from predicted national test scores). Two important outcomes emerged.

First, the proportion of students able to demonstrate understanding of some very basic ideas in these topic areas was low, and often increased little with age. Fewer than 50% at age 14 understood that electric current is the same everywhere around a series circuit; the proportion at age 16 was almost identical. Ideas that build on this, about the relationships between voltage, current and resistance in simple circuits, were grasped by fewer than 20% of 16 year-olds. Less than a quarter showed sound understanding of the forces acting on objects moving at steady speed or slowing down. Only about half at age 14 (rising to about 60% by age 16) understood that mass is conserved in physical and chemical changes. Over 50% at age 14 wrongly classified some particle-level illustrations of mixtures and compounds, and of physical and chemical change. These results are not especially new or surprising – indeed they corroborate findings of studies that predate the National Curriculum. But they indicate that current approaches to the teaching of science do not lead, for many students, to a sound understanding of many basic explanatory ideas – and that this is not highlighted by current measures of student performance.

A second important outcome concerns how we measure 'understanding'. Most research studies, like examinations, probe each key idea with a single question. Instead, we used several, to explore the consistency of individual students' responses. We found surprising variation in the models used in different questions probing the same idea. Table 1 shows one example. Of the 115 students who used the correct scientific model in Q1 (in Figure 1), only 65 did so again in Q2. Either question alone is a dubious indicator of understanding. We found similar inconsistencies in answers on other science ideas. Clearly any measure of 'understanding' must include several questions on each idea tested, and state a criterion for 'understanding'. Agreed outcome measures are a pre-requisite for experimental comparisons of programmes or teaching approaches. Even for topics where targets can be readily agreed, developing such measures is a considerable task.

Helping teachers use diagnostic assessment

In the final phase of the project, we explored how provision of diagnostic materials might influence teachers' practices and student learning. Twenty teachers, including some from the project partnership group, were given a bank of diagnostic questions for a chosen science topic, with some suggestions on possible ways of using them. Interviews with these teachers explored how they used the materials and its impact on their teaching.

Not surprisingly, some teachers' initial reaction was *'oh good, questions – we're so short of useful questions'.* Several chose from the banks to make up new end-of-topic tests: *'we needed a new test … so I looked in the pack … and just pulled a selection out'.* A few talked of using the banks to develop pre-tests for a new topic, *'to find out what their initial ideas were before we then moved on'.* This opens up the possibility of more responsive teaching, taking students' existing ideas into account.

Many teachers were hesitant about increasing the amount of testing: *'we do so many tests. I wanted a more flexible way, that would enable me to see where the problems were in their thinking and move them forward.'* Other comments strongly reinforced this viewpoint. One teacher commented that she had initially seen the banks as test items but quickly came to the view that: *'the use for me is opening up discussion, thinking about how they're actually perceiving things.'* Strikingly, more than half of the group highlighted: *'discussion with the pupils … these have prompted a great deal, which wouldn't have been there without them'.* For some, this was a development of their teaching style. One acknowledged that, *'if **I** was the sort of teacher that was always prompting discussion, then it probably wouldn't have been a necessity, I wouldn't have needed that. But I did and it's helped, without question. I'm having more discussions in class than previously.'* For another it was a matter of seeing how to use: *'a much more interactive, discursive approach … a style of teaching I prefer. … perhaps I haven't been confident in physics before to risk it. This has given me an impetus.'* An NQT felt that the question banks helped clarify her own understanding, and highlighted key points to teach.

Almost all were positive about students' engagement in lessons using diagnostic questions, with several commenting that multiple-choice questions prompted more lively discussion than open-response ones, as they obliged students to evaluate alternative viewpoints: *'What I got, from one EPSE question, was an entire unplanned lesson, with pupils fully engaged and making real progress with their thinking.'*

The majority of teachers involved in the study reported changes, of varying degree, in their teaching styles, or their approach to the science topic in question. All planned to continue using the materials and teaching approaches they had begun to explore. All were convinced that students' understanding – and in some cases also their own understanding – of the ideas that really matter had been significantly improved. This exploratory study demonstrates the potential for influencing practice through the provision of teaching materials based on research, and developed through collaboration between teachers and researchers.

Reference

Black, P. & Wiliam, D. (1998). *Inside the Black Box. Raising standards through classroom assessment.* London: King's College, School of Education.

Major implications

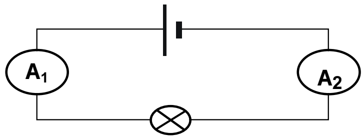
This work has shown that teachers' practice can be significantly influenced by making available teaching materials based on research findings and insights. In this project, the practices of a group of teachers were significantly enhanced (in their own view and that of the researchers) by access to banks of diagnostic questions informed by the findings of research on concept learning, and by researchers' analyses of subject matter and experience in probing understanding. These helped teachers to identify more precisely, and to focus teaching more strongly on, the key ideas that are at the heart of an understanding of these science topics, and which provide a basis for further learning. Teachers valued how the questions enabled them quickly to assess the understanding of all the students in a class, rather than sampling a few individuals. They also found structured diagnostic questions particularly useful for stimulating small-group and whole-class discussion, and reported high levels of student engagement, and lively debate about ideas and explanations, often providing clear evidence of student learning. Several felt that diagnostic materials helped them to teach science topics outside their specialist area in more interactive ways, and with a clearer understanding of which ideas to emphasise.

Similar banks of diagnostic materials are needed for other science topics with similar conceptual demands, perhaps drawing on the experience of current national projects in New Zealand and Australia to develop web-based banks of assessment resources for teachers. Ways of encouraging more systematic formative assessment using these materials also need to be further explored, perhaps by restructuring our materials or by providing more targeted CPD. Linking our 'provision of materials' approach to other current projects seeking to develop formative classroom assessment might be a means of increasing impact on practice.

The large banks of diagnostic questions produced are in themselves a major outcome of this project. They embody a huge amount of accumulated knowledge from research, Their 'translation' from research tools into teaching materials involved significant creative input; the outcomes are new artefacts. Articulation of findings in the form of specific practical implications or usable tools needs to be more generally and widely seen as a necessary and critical part of the process by which educational research might influence practice. The development of diagnostic question banks required detailed analyses of subject content in each science topic; learning objectives then had to be operationalised as tasks providing differential diagnosis (of students who do/do not understand a point). This is a severe test of the way ideas are stated and sequenced in curricula and teaching schemes. In our work, it highlighted significant problems in the sequence of ideas in several areas, particularly mechanics (forces and motion). Here key ideas in the learning sequence (like that all forces arise from interactions and so always come in pairs) are unduly delayed, and some critical ideas (like identifying clearly which object each force is acting *on*) are insufficiently emphasised. By drawing infelicities of this sort to attention, and by clarifying expectations of what students should be able to *do*, the development of diagnostic questions covering the whole science curriculum could lead to significant improvement of practice and outcome.

Finally, this work has shown that the level of understanding among students of many key ideas in science is low, that in many cases it increases only very slowly with age, and that well-known misconceptions are prevalent. It seems clear that current approaches to the teaching of science do not develop widespread understanding of many core ideas in the science domains studied. This, however, is not inevitable. Teachers, when using our diagnostic question banks, noted that many students actually engage much more strongly with these basic ideas, if given time and space to think about them, than with their normal science diet. It also seems clear that current measures of student performance in science do not detect and highlight these gaps in fundamental understanding. Whilst it is not possible to collect exhaustive data on every idea that might be deemed important, some systematic data collection should be undertaken, and maintained over time, to monitor levels of understanding of a few key ideas as students progress through their school careers, as a means of monitoring the curriculum and its effectiveness.

In this circuit, the bulb is lit.



(a) What can you say about the readings on the two ammeters?

Tick ONE box (✔)

- ❑ The reading on ammeter A_1 is bigger.
- ❑ The reading on ammeter A_2 is bigger.
- ❑ The readings on the two ammeters are the same.

(b) How would you explain this?

Tick ONE box (✔)

- ❑ The current is the same all round the circuit.
- ❑ **Some** of the current is used up by the bulb.
- ❑ **All** of the current is used up by the bulb.

Figure 1 A two-tier question on electric circuits

Further information

Further information on the project, including full text of several articles and conference presentations and a sample of the teaching materials produced, can be downloaded from the EPSE Network website (address below).

More details of the study of teachers' use of diagnostic question banks can be found in: Millar, R. and Hames, V. (2003). Improving the research-practice interface. The impact of research-informed teaching materials on science teachers' practices. Paper presented at the Annual Conference of the National Association for Research in Science Teaching (NARST), Philadelphia, March.

A discussion of the data on students' understanding of key science ideas and its implications for measuring 'understanding' can be found in: Millar, R. and Hames, V. (2002). Using diagnostic assessment to improve science teaching and learning. *School Science Review,* 84(307), 21-24.

A TLRP 'gateway' book, in the *Improving Learning* series, is in preparation on the outcomes of all four EPSE Network projects and their implications for efforts to increase the impact of research on practice in science education. This will be published by RoutledgeFalmer, in 2004. Other articles for academic and professional journals on various aspects of the work are also planned.

The warrant

The diagnostic questions developed in the course of this work are based on an extensive body of research, augmented by the experience of a group of practitioners. These 'state of the art' tools provide significantly clearer and less ambiguous data than many previous studies. Trials of questions and analysis methods during the development process has led to guidelines for the design of test instruments, and criteria for improving the validity and reliability of measures of performance. The overall design of the banks of items is also based upon a clear and explicit analysis of the content areas probed.

Evidence of levels of student understanding of key science ideas are based on samples of over 200 students, for each science topic, at the end of each Key Stage. Other data collected on these students enables us to compare each sample to the national average. (Samples are all close to, or slightly above, national average performance.)

Evidence of the ways in which provision of research-based diagnostic materials can influence practice comes from case studies of 20 teachers in 8 schools. The consistency in the views and experiences of these teachers suggests that the outcomes reported are representative and are likely to be replicated more widely. In general, the very positive response of teachers to the materials and the approach adopted in this study lends weight to the general conclusion that research-based materials are a powerful way of influencing practice.

EPSE Network website:
http://www.york.ac.uk/depts/educ/projs/EPSE

EPSE Network grant holders: Robin Millar (co-ordinator), John Leach, Jonathan Osborne, Mary Ratcliffe

Project team (EPSE 1): Vicky Hames, Robin Millar

Project contact: Professor Robin Millar Email: rhm1@york.ac.uk
Telephone: +44 (0)1904 433469
Department of Educational Studies, University of York, Heslington, YORK YO10 5DD

www.tlrp.org

Teaching and Learning Research Programme

TLRP is the largest education research programme in the UK, and benefits from research teams and funding contributions from England, Northern Ireland, Scotland and Wales. Projects began in 2000 and will continue with dissemination and impact work extending through 2008/9.

Learning: TLRP's overarching aim is to improve outcomes for learners of all ages in teaching and learning contexts within the UK.

Outcomes: TLRP studies a broad range of learning outcomes. These include both the acquisition of skill, understanding, knowledge and qualifications and the development of attitudes, values and identities relevant to a learning society.

Lifecourse: TLRP supports research projects and related activities at many ages and stages in education, training and lifelong learning.

Enrichment: TLRP commits to user engagement at all stages of research. The Programme promotes research across disciplines, methodologies and sectors, and supports various forms of national and international co-operation and comparison.

Expertise: TLRP works to enhance capacity for all forms of research on teaching and learning, and for research-informed policy and practice.

Improvement: TLRP develops the knowledge base on teaching and learning and collaborates with users to transform this into effective policy and practice in the UK.

TLRP is managed by the Economic and Social Research Council research mission is to advance knowledge and to promote its use to enhance the quality of life, develop policy and practice and strengthen economic competitiveness. ESRC is guided by principles of quality, relevance and independence.

TLRP Directors' Team
Professor Andrew Pollard *Cambridge*
Dr Mary James *Cambridge*
Dr Kathryn Ecclestone *Newcastle*
Dr Alan Brown *Warwick*
John Siraj-Blatchford *Cambridge*

TLRP Programme Office
Dr Lynne Blanchfield *Lsb32@cam.ac.uk*
Suzanne Fletcher *sf207@cam.ac.uk*

TLRP
University of Cambridge
Faculty of Education
Shaftesbury Road
Cambridge CB2 2BX UK

Tel: +44 (0)1223 369631
Fax: +44 (0)1223 324421

T·L·R·P TEACHING & LEARNING RESEARCH PROGRAMME

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL