ICMI Algebra Conference
University of Melbourne
December 9-14 2001

Where Next?
some comments on implementation

Hugh Burkhardt
Shell Centre and Schools of Education
Nottingham and Michigan State Universities

This Conference has brought together many insights into the learning of algebra, and into aspects of teaching that can help students to understand and do it better. Here I want mainly to comment on what else may be needed to turn these insights into actual improvements in learning in school classrooms around the world. Some of these comments are specific to algebra; others apply more widely, to other aspects of learning mathematics, and learning how to use it to solve problems of importance to the solver. The kind might call it "Generaliser evaluates solution"; others may think that "Real problem evaluates solution" describes it better.

The Conference

I begin with a few comments on the conference discussions. The papers presented include some of the best thinking from around the world on the difficulties that students find in learning algebra, and reports on some new and promising efforts at helping them. The working groups analysed these and other inputs, and synthesised them into interesting and useful reports. Building on years of research, much of the discussion was along expected lines. I will mention two of the aspects that surprised me, before going on to issues of implementation.

- Some important issues seemed to fall between working groups. For example, the potential of computers and calculators, and some proven achievements, was the theme of some groups; others recognised the central importance of technology but did not consider in detail what this might mean for their topic. The "Why Algebra? What Algebra?" group, for example, did not address the specific roles of technology in its suggested outline of a minimum "algebra for all". There is general recognition that these technologies are transforming the way algebra is done, outside and ultimately inside schools. Systematic links need to be established, showing evidence on how technology can support the learning and doing of different aspects of algebra.
- The power of algebra as a tool for solving problems from outside mathematics received little attention. Applications were mainly illustrative, with the focus on understanding algebraic concept and skills. There was little reference to learning the modelling skills needed to 'mathematise' practical problems with algebra, yet in the world for which we prepare students most algebra is done for such purposes, not as a pure study. (There was even a reference to "modelling as a bridge to algebra" – to me, a startling reversal of priorities)

Perhaps these omissions can be considered in the writing of the book.

These and some other mismatches arose from the balance of those taking part. The dominant focus was on gaining insights into learning and teaching through research, with little attention to the practical problems of how these insights might lead to better teaching and learning of algebra in the various countries around the world that form ICMI. But ICMI is much broader than a research group. What are the further steps needed to help countries bridge the gap between research insights and improved practice?

Making a difference in practice

I aim to address both meanings of "in practice", that is to look at the *practical* problems of achieving improvements in teachers' classroom *practice* in helping students around the world learn algebra more easily and effectively.

Research insights are a key input to this process but, obviously, they are not enough – unless you believe that most teachers read the research literature and, then, can and will turn its insights into substantial improvements in their practice. Neither is plausible, let alone true. What else is likely to be needed? There are standard answers.

A curriculum design will be needed, embodying policy decisions in the light of research and available resources. This design may be set down in 'standards' or other policy documents.

Curriculum materials embodying the results of research need to be designed and developed, so that work they well with typical teachers and their students in realistic circumstances of personnel and support. If done well, this demands more than research – notably a great deal of creative design input, and systematic refinement through trialling.

Professional development support will be needed, because the range of classroom strategies and skills required for this kind of teaching is much broader than for traditional 'teacher-centred' mathematics teaching. This too may need design and systematic development of materials to support the leaders of these activities.

Pressure will surely be needed to encourage all those who need to change to do so. They are all busy people with well-grooved professional practice; few will make the effort that such changes demand without compelling reasons. Pressure is less expensive to provide than the various kinds of support just outlined, and thus attractive to politicians and the public, particularly when labelled "accountability". It follows that the assessment used must be well-aligned with the curriculum goals; otherwise, many teachers will only teach the aspects that are assessed.

Who is responsible in each education system for developing these elements? What might be done by the ICMI community to help them? How far has this conference contributed?

I shall look at these questions starting from the focus of the conference's work by looking at the contributions of research.

Different styles of research

In education, the question "What is research?" has long raised controversy, even "paradigm wars". I want to take a broader view, looking across fields at the different meanings and traditions and asking how each can contribute to making a difference in the learning of algebra in the ways outlined at this conference. Any such review will see that strength in research requires a variety of approaches, tailored to the problems in hand.

I will start with the definition used in the UK Research Assessment Exercise (RAE), in which all UK university departments were rated during 2001 on their work over the previous five years.

> “ **'Research' for the purposes of the RAE** is to be understood as original investigation undertaken in order to gain knowledge and understanding. It includes work of direct relevance to the needs of commerce and industry, as well as to the public and voluntary sectors; scholarship; the invention and generation of ideas and, images, performances and artifacts including design, where these lead to new or substantially improved insights; and the use of existing knowledge in experimental development to produce new or substantially improved materials, devices, products and processes, including design and construction.”

The breadth of this definition may surprise people. It arises from taking seriously four different traditions, characteristic respectively of the:

Humanities, Sciences, Engineering, and the Arts

If you look for a fundamental measure of quality in research across all these fields, it is difficult to go beyond:

Impressing key people in your field

The balance of qualities which achieve this varies. What balance would be most beneficial for Education, and how far does it differ from current criteria? I believe that all of these traditions have important contributions to make in Education, but that currently the balance of effort and of ‘academic credit’ is far from optimal. Let us look at each in turn, the nature of the activity and the forms of output, and their potential impact on students’ learning in typical classrooms.

The 'humanities' approach

This is the oldest research tradition, based on scholarly acquisition of knowledge and critical analysis of it, and of other people’s work. From the RAE definition it is “original investigation undertaken in order to gain knowledge and understanding; scholarship; the invention and generation of ideas..... where these lead to new or substantially improved insights” Note that there is no tradition of empirical testing of the assertions made.

The key product is:

critical commentary

There is a lot of this in education. The ideas and analysis, based on the authors’ reflections on their experience, are often valued. Without the requirement of further empirical testing, a great deal of ground can be covered. However, since so many plausible ideas in education have not in practice led to improved large-scale outcomes, the lack of empirical support is seen as a weakness. How should you distinguish reliable comment from plausible speculation? This search for “evidence-based education” has led to the dominance of the ‘science approach’.

The 'science' approach

This approach to research is focussed on the development of better *insight*, of improved understanding of "how the world works", through the analysis of phenomena, the building of models which explain them, and the empirical testing of those models. In the RAE definition, it is *“original investigation undertaken in order to gain knowledge and understanding; scholarship; the invention and generation of ideas..... where these lead to new or substantially improved insights”*. Note that this is the same wording but now there is an essential requirement for empirical testing of the assertions made, which are now called hypotheses or models.

The key products are:

assertions
evidence-based arguments in support
evidence-based responses to key questions

where the evidence is expected to be empirical. These commonly appear in the forms of:

research journal papers
books
conference talks

This is the approach is now predominant in the research in science and mathematics education and, in the context of the learning and teaching of algebra, in the work of this conference. Such research provides insights, identifies problems, and suggests possibilities. However, it does not itself generate practical solutions, even on a small scale; for that, it needs to be linked to the 'engineering' approach.

The 'engineering' approach

This is directly concerned with practical *impact*, with helping the world "to work better", not only understanding how it works, by developing solutions to recognised practical problems. It builds on science insights, insofar as they are available, but goes beyond them. In the RAE definition is it is “the invention and generation of ideas.... and the use of existing knowledge in experimental development to produce new or substantially improved materials, devices, products and processes, including design and construction”. Again there is an essential requirement for empirical testing of the products and process, both formatively in the development process and in evaluation. The importance of science-based insights varies from field to field, depending how far the ‘theory’ is an adequate basis for design. I’ll come back to this.

The key products are:

tools and/or processes that work well for their intended uses and users
evidence-based evaluation and justification
responses to evaluation questions

With these elements, development *is* research. However, in the academic community it is often undervalued – in some places only insight research in the science tradition is regarded as true research currency.

Of course, definitions are man-made – social constructs that are partly arbitrary. However, the effects of lowering the status of ‘educational engineering’ include:

- lower standards of materials and processes, since the imaginative design and rigorous development that good engineering demands are not expected;
- lower practical impact of important results of insight-focussed research, since designers feel less need to know or use the background research;

- pressure on good practitioners in universities to produce insight research papers rather than use engineering research methods to improve their practice.

All this leaves a hiatus between insight research and improved classroom practice which is, to say the least, unfortunate. Society's priorities for education are mainly practical – that young people should learn as effectively as possible. The perceived failures of educational research to deliver the goods in practical terms is reflected in the low levels of support for it. If politicians have a problem to solve, is their first move to call a researcher? Not often.

This status pattern, where the pure is valued far more than the applied, is common; it is not general at any level of research. For example, the two people who have won two Nobel Prizes in the same field are:

- John Bardeen , the physicist, for the *transistor*, and for the *theory of superconductivity*
- Fred Sanger, the biologist, for the *3D structure of haemoglobin* (a first in this application of X-ray crystallography) and for *the procedure for sequencing DNA*.

At least two of these are engineering in approach, being "the use of existing knowledge in experimental development to produce new or substantially improved materials, devices, products and processes, including design and construction". With examples like these, education need not fear for its respectability in giving equal status to engineering research. There is a key difference in education – there is currently no industry with high standards of research-based development that takes forward prototypes developed in universities, turning them into finished products. A greater responsibility thus falls on the academic community to do this kind of work.

When built on good 'science', good 'engineering' enables us to draw much more definite conclusions about the outcomes. This has been clear in the contributions to this conference, for example those from the Freudenthal Institute, where the work has been taken beyond the small-scale study to implementation and evaluation in practice.

The 'arts' approach
This may be seen as related to the humanities approach rather as engineering is to sciences. In the RAE definition is it is "the invention and generation of ideas and, images, performances and artifacts including design, where these lead to new or substantially improved insights".

I will say little about this because, though it enriches education and could do more, it is not central to my strategic concern here, which is primarily concerned with translating insights into practical impact in classrooms and school systems.

Let me stress that this is not a plea for the abandonment of insight-focussed 'science' research in education. It is essential but not enough. Rather, it is an argument about balance – that there should be much more impact-focussed 'engineering' research and that it should receive comparable recognition and reward. *The different styles can and should be complementary and mutually supportive.*

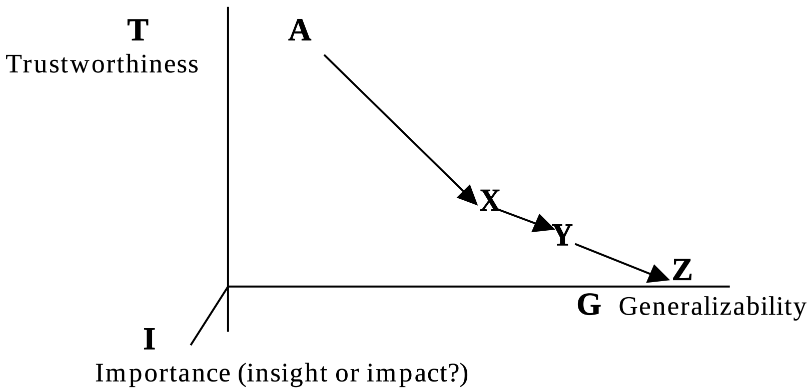
Schoenfeld's dimensions
In a recent paper Alan Schoenfeld has suggested three dimensions for classifying research outputs:

Generalizability; To how wide a set of circumstances is the statement claimed to apply?

Trustworthiness: How well substantiated are the claims?

Importance: How much should we care?

Typically, any given paper contains assertions in different parts of this 3D space. (From the above I believe that Importance, a key variable, could usefully distinguish 'insight' from 'impact') The graph below focusses on the other two variables above, G and T, say. A typical research study looks carefully at a particular situation, often a specific treatment and student responses to it. The results are high on T, low on G – the zone A on the graph.



The conclusions section of the paper then goes on to discuss the 'implications' of the study, often much more wide ranging but with little evidence to support the generalisations involved, which are essentially speculative (or perhaps in the humanities tradition). These are illustrated as X, Y and Z.

Only large scale studies, or perhaps metanalysis, can avoid this problem. The work of Alan Bell, Malcolm Swan and the Shell Centre team on 'Diagnostic Teaching' illustrates this well. (The approach is based on leading students into making errors, then getting them to understand and debug them through discussion) The early work showed comparable learning gains through the teaching period (pre- to post-test) but without the fallaway over the following 6-months that the standard 'positive only' teaching approach showed. The first study was for *one mathematics topic*, with *one teacher*, and *one class*. Only five years later, when the effect was shown to be stable across many topics and teachers could one begiin to make reasonably trustworthy statements about 'diagnostic teaching' as an approach. Even then, there remain further T2-type questions about its accessibility to typical teachers in realistic circumstances of support.

The general point is that much research is about treatments, not about principles; to probe the latter one must check stability across a range of variables (student, teacher and topic in this case). This typically **needs time and teams** beyond the scale of an individual PhD or research grant. Other subjects arrange this; if it were more common in education, the research could have high G *and* T and, if the importance were enough, be worth taking more seriously

My final point on research styles returns to the overall challenge – establishing a sound research-based path from insights to large scale implementation. The argument is summarised in the table below. Note the different research foci, R in the third column.

Four Levels of R&D

L	**Learning level**	*variables:* student task	*eg:* R: concepts, skills strategies, metacognitioı D: learning situations, probes data capture
T_1	**Teaching level**	instruction student task	R: teaching tactics +strategies D: classroom materials f some teachers
T_2	**Teacher level**	teacher instruction student task	R: representative teacheı with realistic support D: classroom materials f most teachers
C	**System level**	system school teacher instruction student task *etc*	R: system change D: 'Tools for Change' *ie materials for:* • classrooms • assessment • professional development • persuasion • ....

Note the crucial difference between T_1, which is about teaching possibilities, usually explored by a member of the research team, and T_2, which is about what can be achieved in practice by typical teachers with available levels of support. Currently, nearly all research is at L and T_1 levels. A better balance across the levels is needed, if research and practice are to benefit from each other as they could. As noted above, T_2 and C -level research need larger research teams and longer time-scales

The status and roles of "theory"

Finally, some related comments on "theory", which is seen as the key mark of quality in educational research. I am strongly in favour of theory. (Indeed, in my other life, I am a theoretical physicist) However, in assessing its role, it is crucial to be clear as to how strong the theory is. From a practical point of view, the key question is:

How far is this theory an adequate basis for design?

Again it useful to look across fields. In Aeronautical Engineering. for example, the theory is strong; those who know the theory can design an aeroplane at a computer, build it, and it will fly, and fly efficiently. (They still flight test it extensively and exhaustively) In Medicine, theory is moderately weak, but getting stronger. Despite all that is known about physiology and pharmacology, most development is not theory-driven. The development of new drugs, for example, is still mainly done by testing the effects of very large numbers of naturally occurring substances; they are chosen intelligently, based on analogy with known drugs, but the effects are not predictable and the search is wide. However, as fundamental work on DNA has advanced, and with it the theoretical understanding of biological processes, designer drugs with much more theoretical input have begun to be developed. This process will continue – indeed there is now work, for example, on cancer drugs tailored to individual tumours.

Education is a long way behind medicine (100 years?), let alone engineering (350 years?), in the range and reliability of its theories. By overestimating their strength damage has been done to children, for example by designing curricula based largely on behaviourist theories. The current dominance of constructivism is similarly inadequate, though less dangerous. (Its incompleteness is more obvious, since it is impossible to design a curriculum built only from constructivist principles and a list of skills) It is not that behaviourism or constructivism are wrong; indeed, they are both right in their core ideas but they are incomplete and an inadequate basis for design. Physicists would call them 'effects'. The harm comes from overestimating their power.

Let me illustrate this with an example from meteorology. "Air flows from regions of high pressure to regions of low pressure" sounds and is good physics. It suggests that air will come out of a popped balloon or a pump, and that winds should blow perpendicular to the isobars, the contour lines of pressure on a weather map, just as water flows downhill, perpendicular to the contour lines of a slope. However, a look at a good weather map in England shows that the winds are closer to parallel to the isobars. That is because there is another effect, the Coriolis Effect. It is due to the rotation and of the earth which 'twists' the winds in a subtle way, clockwise around low pressure regions. (Like water down the plughole, they go round the other way in the Southern Hemisphere) In education there are many such effects operating. We have identified some of them (behaviourism and constructivism are two) but, as in economics, it is impossible to predict just how they will balance out in a given situation. Thus design skill and empirical development are essential, with theoretical input providing useful heuristic guidance. The essential point is that **the design details matter** – they have important effects on outcomes and are not determined by theory.

Empirical development is essential, usually *through several cycles* ("Fail Fast, Fail Often"), in *realistic circumstances*, mostly with *small samples* with *rich, detailed feedback* through observation and analysis . But such **care** is not enough; quality also needs **flair –** outstanding design skill, which is rare and develops slowly. This is an **engineering research approach** to educational design. If effective implementation of the research presented at this conference is to happen, this approach will be needed, across the levels L to C.