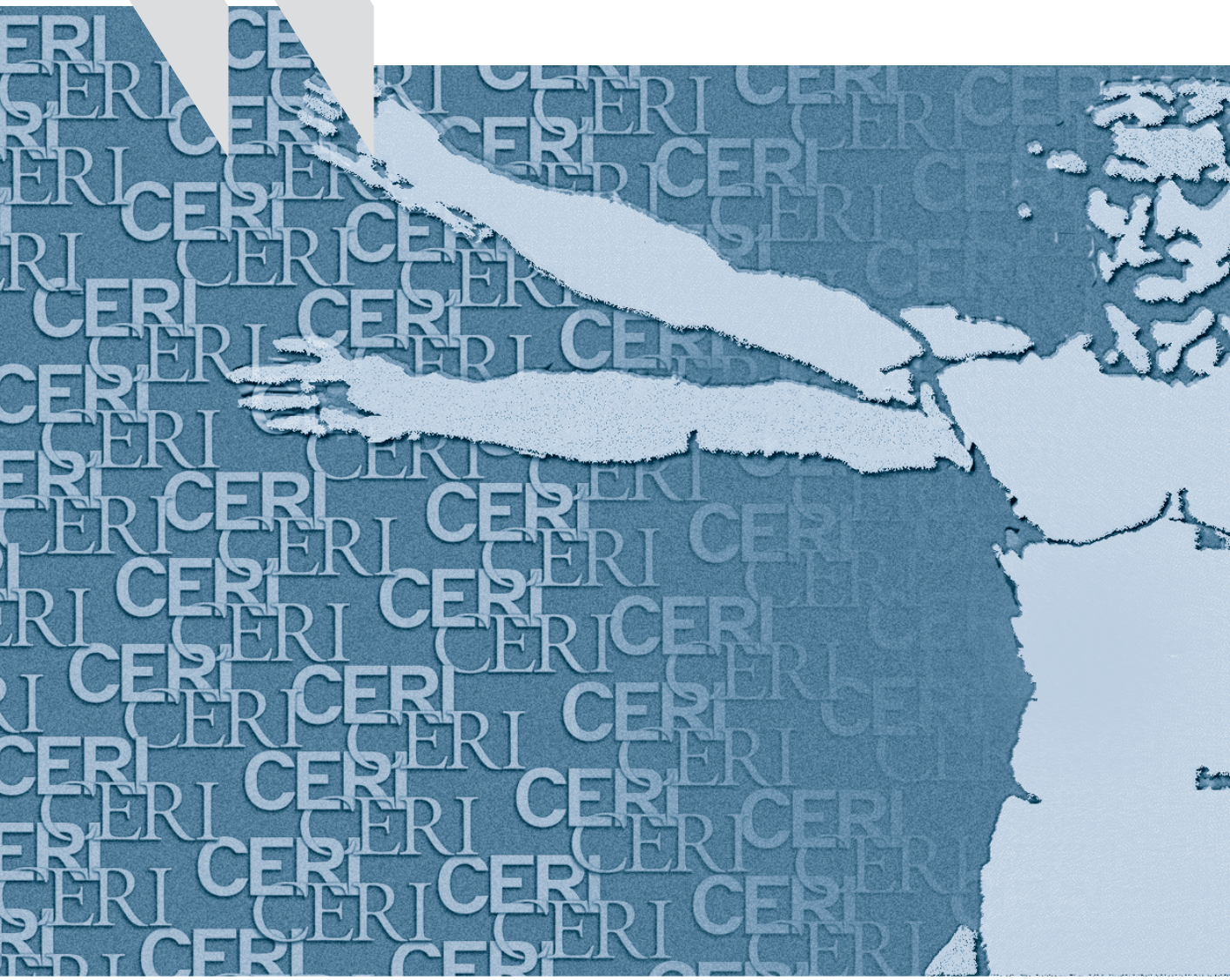
Knowledge Management

Evidence in Education

LINKING RESEARCH AND POLICY



OECD

Centre for Educational Research and Innovation
CERI
Centre pour la recherche et l'innovation dans l'enseignement

Centre for Educational Research and Innovation

Evidence in Education

LINKING RESEARCH AND POLICY

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 30 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

Foreword

The foundation for this volume was laid more than a decade ago in the 1995 OECD Centre for Educational Research and Innovation (CERI)'s report *Educational Research and Development: Trends, Issues and Challenges.* This report raised the question of why the role of governments in promoting and using educational research had emerged as a prominent issue.

Almost a decade later, CERI's work on knowledge management pointed to the key role of knowledge-based innovation in education. A series of country reviews of educational R&D confirmed that in most if not all countries the issues of effective relationships between research and policy makers, capacity-building within those domains, and importance of allocating scarce resources in the most efficacious manner remain as important as they were ten years ago.

The current project, labelled "Evidence-based Policy Research in Education", centred on a series of workshops held between April 2004 and July 2006 in the United States, Sweden, the Netherlands and the United Kingdom, which brought together key players in research and policy to exchange experiences and practices.

Evidence in Education: Linking Research and Policy brings together highlights of this workshop series. The publication looks at the issues facing educational policy makers, researchers, and stakeholders – teachers, media, parents – in using evidence to best effect. It focuses on the challenge of effective brokering between policy makers and researchers, offers examples from Canada, Finland, Singapore, and the United Kingdom, and presents the politicians' perspective.

Within the CERI Secretariat, this report was edited by Tracey Burns and Tom Schuller, with the assistance of Cindy Luggery-Babic and Delphine Grandrieux.

Barbara Ischinger,
Director, Directorate for Education

EVIDENCE IN EDUCATION: LINKING RESEARCH AND POLICY – ISBN-978-92-64-03366-5 © OECD 2007

ACKNOWLEDGEMENTS

This conference series would not have been possible without the support of the hosts of the conferences. We wish to acknowledge the support of the US Department of Education Institute of Education Sciences; the Ministry of Education and Research, Sweden; the Ministry of Education, Culture and Science, the Netherlands; and the Economic and Social Research Council's Teaching and Learning Research Programme and the Department for Education and Skills, the United Kingdom.

We wish to acknowledge the contribution of the following individuals in organising the conferences from the host side: Jon Baron and Phoebe Cottingham (United States); Mats Ekholm and Hans-Ake Öström (Sweden); Janneke Timmer (Netherlands); and Andrew Pollard and James O'Toole (United Kingdom).

In addition we would like to thank the participants from all OECD countries who took part in the conferences to make them such a success, including the contributing authors to this publication: Adrienne Alton-Lee, René Bugge Bertramsen, Bob Boruch, Satya Brink, Thomas Cook, Jane Davidson, Stephen Gorard, David Gough, Rebecca Herman, David Hogan, Bill Kilgallon, Hannele Niemi, Johnny Nilsson, Andrew Pollard, Rien Rouw, Hans Stegeman, Charles Ungerleider, Maria van der Hoeven, and Jerzy Wisniewski.

And lastly, our thanks to colleagues (past and present) Barry McGaw, Henno Theisens, Simon Field, Greg Wurzburg and Janet Looney for comments on previous versions of the Secretariat analysis.

Table of Contents

Executive Summary

There is increasing pressure across OECD countries for greater accountability and effectiveness in education policies and systems. Still, available information often does not provide the elements necessary for decision-making, either because the rigorous research relevant to policy needs has not been conducted, or the research that is available does not suggest a single course of action.

Evidence in Education: Linking Research and Policy explores the issues underlying the use of evidence in educational policy-making. It discusses what constitutes evidence for research in education, how that evidence can best be utilised, and possible solutions to challenges observed by participating countries.

Research is playing a newly important role for evidence-informed policy – but what counts as evidence?

The OECD Secretariat analysis argues that the recent re-emergence of this issue is due to a number of key factors, including:

- a greater concern with student achievement outcomes;
- a related explosion of available evidence due to a greater emphasis on testing and assessment;
- more explicit and vocal dissatisfaction with education systems, nationally and locally;
- increased access to information via the Internet and other technologies; and
- resulting changes in policy decision-making. These are accentuated by broader issues to do with the perceived legitimacy of policy-making in general.

At the same time, there has been a shift across most OECD countries to de-centralise decision-making in education, giving more responsibility and mandating to local authorities. Given greater information, less quality control, a more informed public, and a greater diversity of policy makers, the role of research for evidence-informed policy becomes newly important.

Part One also contains the contribution of two well-respected methodologists, Tom Cook from Northwestern University (United States) and Stephen Gorard from University of York (United Kingdom), who look at the need for quality research and the appropriate methodology in education research. This issue is tackled in the form of a dialogue in which each author first states his position in terms of the key concerns for education research methodology, and then exchange on what counts and should count as evidence. Cook and Gorard concur that experimental designs, and especially that of randomised control trials, should be given a stronger role, though they differ on their

exact place in the range of research instruments. For policy-making, is it a question of using the best available evidence, or should only the best evidence be used (which would imply that evidence that does not meet the gold standard could not be used)?

Brokerage agencies are key stakeholders in linking research and policy

Part Two looks at linking research to policy and the interface between communities of stakeholders. In doing so it provides a series of studies of existing and newly formed brokerage agencies. Chronologically they range from EPPI-Centre in the United Kingdom (formed in 1993) to the Knowledge Chamber of the Netherlands and the Danish Knowledge Clearinghouse, which were created as a result of the OECD/CERI workshop series that provided the basis for this publication.

The brokerage agencies are also distinguished in their goals and means, with New Zealand's Best Evidence Synthesis Programme providing an example of a brokerage programme embedded within the Ministry whereas the Canadian Council on Learning, although also federally funded, is separate from the provincial Ministries. The United States' What Works Clearinghouse works in collaboration with a number of other institutes and subcontractors, and also conducts consumer surveys and questionnaires to ensure that the service it provides is meeting the demands of the users (researchers, practitioners, policy makers, etc).

As part of the continuing effort to draw parallels between education and other subject areas and learn from relevant experience, Part Two also provides an example from the field of Social Care in the form of the United Kingdom's Social Care Institute for Excellence (*www.scie.org.uk*). From this contribution we can observe both the similarities of experience, goals, and efforts of a brokerage agency in this domain, as well as some shared challenges.

Implementing evidence-informed policy research: promising examples of national programmes

Several examples from the field, drawn from a number of different countries (Canada, Finland, Singapore, United Kingdom) are presented in **Part Three**. Canada's contribution looks at the launch of the National Children's Agenda, which focused on the long-term vision of fostering good Canadians by aiming for lifelong learners, productive workers, nurturing parents, and engaged citizens. Life as Learning (Finland) was set up as a national research programme aimed at encouraging the development of a research culture to support interdisciplinary and international research projects.

The contribution from Singapore is unique in that it is a non-OECD country with an unparalleled set of funding and research opportunities, including a strong database with data on the entire student population. Lastly the United Kingdom's Teaching and Learning Research Programme (TLRP) is an immensely complex initiative with assured medium-term funding (2000-2011) that aspires to improve the quality of education research in the United Kingdom as well as contributing new knowledge.

All of the contributions allow us a close look at the quality of the relationships between the primary stakeholders as well as how the project bridges the gap between the

evidence they can provide and the needs of the policy maker. This section thus serves as both a set of concrete examples interesting in their own right and as a way to observe the various theoretical issues discussed in this volume in practice. Similarly to Part Two, it includes an example from outside the Education field, in this case a programme governed by the Department of Human Resources and Social Development (Canada).

Politicians' perspectives

The evidence-based policy research debate has to large extent been focused on policy makers and researchers as the major players. In **Part Four** we hear the voice of the politician, charged with making decisions and evaluating the best way forward for the education system in their jurisdiction. Johnny Nilsson, the former Secretary of State for Education in Sweden, speaks of the experience of the past and some of the limitations and concerns of the politician in a period without the current emphasis on evidence-based policy. We hear also of the serious weight that has been placed on this process in the Netherlands, and how and why this work was so important to the Minister of Education, Science, and Culture, Maria van der Hoeven. The Netherlands has a number of initiatives that have followed from the OECD/CERI work and are looking to the future.

Jane Davidson of Wales uses her experience as the longest serving education minister in the United Kingdom to discuss the role of evidence in policy-making and the struggle she has engaged in to bring it to the forefront of the policy-making process. Finally we have a unique case study from Poland, which has recently brought together current and former Ministers and decision-makers in an on-going effort to raise the priority and profile of education research in policy-making in that country. Jerzy Wisniewski, former senior civil servant and CERI Governing Board member, contributes the challenges and aspirations of a country that is just beginning to focus on these issues.

Evidence in Education: Linking Research and Policy brings together international experts on evidence-informed policy in education from OECD countries. The report looks at the issues facing educational policy makers, researchers, and stakeholders – teachers, media, parents – in using evidence to best effect. It focuses on the challenge of effective brokering between policy makers and researchers, offers examples of policy-relevant research from Canada, Finland, Singapore, and the United Kingdom, and presents politicians' perspectives. This book provides a fresh angle on key issues facing policy makers, researchers, and school leaders today.

Part One
Setting the Stage: The Evidence Agenda and Methodological Issues

Chapter 1
The Evidence Agenda

Tracey Burns and Tom Schuller,
OECD Centre for Educational Research and Innovation

In this chapter, we examine the resurgence on interest in "evidence", outline the roles and relationships between major stakeholders, and provide an overview of this publication.

Part One: Setting the Stage: The Evidence Agenda and Methodological Issues

In recent years a number of public crises have seized the attention of the world and required rapid responses from governments to ensure the health and safety of the public and maintain their confidence in policy makers. The 2001 UK foot and mouth crisis and the emergence of SARS in Asia and Canada highlight the difficulties of decision-making for policy makers and also the necessity for time-sensitive information on which to base those decisions. In each of the above examples, dramatic action needed to be taken urgently. These decisions resulted in substantial economic and societal losses, as well as worldwide reaction to piles of scorched carcasses, delayed elections, passengers wearing masks while taking public transport, and restricted movement. The dangers were contained and the emergencies passed, but *post factum* evaluations revealed that perhaps the policy decisions taken were not, in fact, the most effective or efficient ones available (House of Commons Science and Technology Committee, 2006).

An even bigger issue is that of climate change. Here the debate has been very public, including much on the nature and reliability of the evidence presented, notably in respect of the report by Sir Nicholas Stern (Stern, 2006). Here the contestability of research and the difficulty of matching timeframe to analysis and action were very much in evidence.

Although extreme examples, these real-life instances illustrate the dilemma facing policy makers of all stripes, including education, who must make swift, time-sensitive decisions based on the information they have available. Yet often the information that is readily available is not "perfect" research on the subject. This could be either because the rigorous research relevant to policy needs has not been conducted; or because there is a disjoint between policy and research communities such that the relevant information is not widely disseminated and so overlooked by the policy maker; or simply that the research that is available is contradictory and so does not suggest a single course of action that could be reflected in policy. Yet clearly it is crucial that policy decisions be made

with the best available evidence, as the decisions made can, as the above examples demonstrate, have far-reaching impacts on all members of society.

Interest in and discussion on how educational policy is aided by research, and specifically on what kinds of evidence from research count or should count in policy and practice, have grown dramatically in the last several years. Discourse on the nature of scientific evidence, challenges for raising awareness of policy makers, and pleas to bring research into classrooms all trigger intense and sometimes heated debate on what constitutes "evidence-based" or "evidence-informed policy", the terms which have come to denote this field, and which we define as "the conscientious and explicit use of current best evidence in making decisions and choosing between policy options".

The activity from which this book (see below) comes was originally titled "evidence-based policy research" (EbPR). However, as the debate progressed this term became increasingly unsatisfactory for many of those participating. This is principally because it is seen to imply too tidy and rational an image of policy-making, as some kind of clinical and objective operation. "EbPR", moreover, can be ambiguous: is it the research or the policy which is evidence-based? The balance swung more towards *the use of research to foster evidence-informed policy*, which leaves more open the actual extent of the evidence use. So this report settles for *evidence-informed policy research* (EIPR) as its focal term. In very broad terms the research that is used to produce evidence-informed policy can be distinguished from purely scientific research in that the former is oriented to informing action while the latter is oriented to developing theory and testing hypotheses (although these are not mutually exclusive categories). The distinction is important, as burdens and standards of proof of causality are very different, and in many cases those responsible for evidence-informed policy are obliged to use the best available evidence at a given moment in time, whatever its strict epistemological status.

In brief, our concern is with the most appropriate forms of evidence to assist public policy-making, and the effective mechanisms for developing and utilising that evidence.

The (re)emergence of "evidence"

Despite the recent resurgence of interest in evidence-informed policy research (EIPR), the issues underlying the discussion are not new. More than a decade ago, the 1995 CERI report *Educational Research and Development: Trends, Issues and Challenges* raised the question of why the role of governments in promoting and using educational research had emerged as a prominent issue, and suggested several possible answers. These included:

- A belief that education and knowledge are increasingly important factors in innovation and economic growth in OECD countries.
- A growing concern with accountability in respect of educational expenditures.
- A concern about the quality and effectiveness of current educational research.

CERI's work on knowledge management pointed to the key role of knowledge-based innovation in education (OECD, 2004). However the country reviews of educational R&D have confirmed the following features as commonly (though not universally) characterising OECD systems (OECD, 2003):

- Low levels of investment in educational research.
- Generally low levels of research capacity, especially in quantitative research.
- Weak links between research, policy and innovation.

In most if not all countries, therefore, the issues of effective relationships between research and policy makers, capacity-building within those domains, and importance of allocating scarce resources in the most efficacious manner remain as important as they were ten years ago. So what has changed? Why revisit a discussion which was already well delineated ten years ago?

Key factors underlying this change are a greater concern with student achievement outcomes; a related explosion of available evidence due to a greater emphasis on testing and assessment; more explicit and vocal dissatisfaction with education systems, nationally and locally; increased access to information via the Internet and other technologies; and resulting changes in policy decision-making. These are accentuated by broader issues to do with the perceived legitimacy of policy-making in general.

A rising concern with the outcomes of education (*e.g.*, student achievement) is one of the most significant overall shifts in educational policy orientation. Much previous work, including the OECD's, concentrated on inputs (financial or otherwise) and participation rates. Today there is a mounting preoccupation with what happens as a result of these investments and activities. Outcomes are interpreted not only in terms of course completion or qualifications, but also in terms of skills and competences (as with the PISA study), access to and success in the labour market, and wider social outcomes such as health and citizenship attributable to education. In other words, policy makers are increasingly interested in what education actually delivers – and therefore with what educational research can tell us about that. A consequence of this has been the explosion of evidence of different kinds resulting from the enormous increase in testing and assessment.

A significant force behind this orientation to outcomes has been the greater interest shown by treasuries and finance ministries in the effectiveness of educational expenditure, as a major component of public expenditure generally. Where annual spending rounds have a strong grip on policy, requests from treasuries for evidence of results present challenges to their counterparts in educational ministries and to policy makers at other levels in the education system. The challenge is to gather evidence which is both appropriate and convincing. This is especially the case where the request is that impacts and effectiveness be given monetary values.

Increased access to information via the Internet and other technologies (including the easily digestible and publicity friendly information arising out of testing and assessment) is potentially the great equaliser, allowing as it does a greater number and wider breadth of individuals with access to all available information. At the same time, however, this process has effectively removed many of the established gate-keepers or quality controls for this information – a process sometimes called "disintermediation". More information is available, yes, but is it good information? And is it presented accurately and in an understandable fashion? Can the reader use it in a comprehensible and useful manner?

These twin effects (greater access with less quality control) have had great impact on policy-decision making, requiring as they do the policy maker to weed through immense amounts of information of unclear quality in order to make decisions on behalf of a more informed constituency. At the same time, there has been a shift across most OECD countries to de-centralise decision-making in education (unlike many of the other areas covered by the social sciences), giving more responsibility and mandating to local authorities. Given greater information, less quality control, a more informed public, and a greater diversity of policy makers, the need for clear, reliable, and easily available evidence on which to base policy-decisions has become more important than ever before,

as has the need to find mechanisms to obtain reliable answers to pressing policy questions. The role of research for evidence-informed policy, then, becomes newly important.

How do OECD countries understand and engage in EIPR? Our perception of the debate thus far suggests that the approach to this issue can be structured around the following dimensions:

- who are the players and the quality of communication/interactions between the different sets of agents involved in commissioning, executing and implementing EIPR;
- the kinds of methodologies and epistemological paradigms which dominate within policy and research communities; and
- the kinds of mechanism available to resolve the tensions and difficulties involved in the process.

Naturally, these interact with each other and need to be related to the structure of governance, notably the levels at which policy goals are set and policies fashioned and implemented, the availability of the evidence, and the substantive policy issues which happen to be prominent at any given time.

Further dimensions could certainly be added, but we suggest, as a basic organising framework, that it is the interplay between these dimensions which defines the different approaches to EIPR that are found across OECD member countries (and most likely beyond); and it is this interplay which needs to be explored in order for good practice to be understood and a useful agenda to be set.

This publication explores these dimensions and then pushes the discussion into the drivers and facilitators underlying the interaction between them, with the aim of investigating new challenges and new opportunities for EIPR in education. It arises out of a 2004-2006 CERI seminar series that brought together researchers and policy makers from all OECD countries to review the main aspects of evidence-based policy research – methods, costs, and capacity – and discuss what constitutes evidence for research in education, how that evidence can best be utilised, and possible solutions to challenges observed by participating countries.

The four seminars in the series were hosted by the United States (April 2004 in Washington DC), Sweden (January 2005 in Stockholm), the Netherlands (September 2005, at The Hague), and the United Kingdom (July 2006 in London). The contributions to this publication come from a selection of these attendees, and represent the diversity of roles, perspectives, and experiences that contributed to the seminar series.

In addition to the Secretariat discussion on the evidence agenda and the roles and relationships among the major stakeholders, this publication includes contributions on:

- methodological issues and what counts as evidence;
- the policy/research interaction: the role of brokerage agencies;
- research for evidence-informed policy in practice: examples from the field;
- the politicians' perspective.

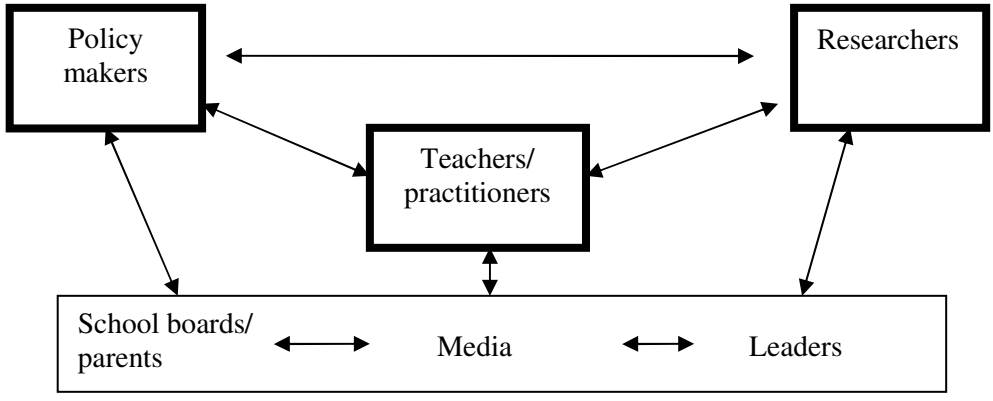
Policy-research interaction: who are the players and how do they interact?

In order to understand the policy-research interaction we need to focus on:

- the quality of the relationships between the primary stakeholders (*e.g.*, policy makers, researchers, practitioners, etc);
- how well they function in terms of information flows and trust levels;
- what could be done to improve the quality in each case.

As a starting point, Figure 1.1 summarises the potential lines of communication and interaction between these different agents. Given the focus of this discussion it privileges policy makers, researchers, and practitioners. Its function is primarily heuristic, *i.e.*, it is designed to prompt reflection and analysis in respect of specific country or other situations (see Levin [2004] for a more complete discussion of the various stakeholders and relationships between them).

Figure 1.1. Primary lines of communication/interaction



Source: OECD.

Ideas for education reform are often perceived to come from within the system of policy makers and, to a lesser extent, researchers working on education. These are the primary, but not the only, agents in the game. In the literature on evidence-based policy making there has been a good deal of discussion on how to bridge the gap between the researchers and policy makers, both in terms of communication of needs and priorities and in terms of coming to an understanding of the different timescales required by each community (*e.g.*, the mandate-limited scope of the policy maker versus the length of time it takes to do good research and pilot implementation as part of that research). This discussion implies that there is a unified concept and community both of researcher and policy makers and that the challenge lies in bridging the gap between these two disparate, but internally homogenous, communities. In real life, of course, nothing is ever that simple.

The researchers

Educational researchers are to be found in many different locations and roles, from university-based academics to hired contractors working for government or private sector. The latter group in particular has grown dramatically in recent years, particularly in countries such as the United States and Great Britain, as a response to increased need both for answers to particular questions and for researchers capable of conducting sophisticated quantitative analyses. Private sector research firms market themselves as capable of producing relatively swift analyses to focused educational research questions, and demand for their services has grown with increased desire to direct funding towards programmes that have proven effectiveness in education.

Educational researchers may be educationalists, other social scientists, computer scientists or architects, among other things. While comfortable with this diversity of identification, education researchers are less comfortable with the diversity of perspectives and backgrounds. The EIPR debate has been characterised by differences of opinion on these matters that have turned the discussion into a battleground between groups of researchers that seem incapable of communicating with each other. The current state of affairs could reasonably be described as mutual antipathy between those researchers deemed to favour "quantitative" approaches versus the "qualitative" researchers (although these labels do not satisfactorily capture the differences that exist between the different groups, see Gorard and Taylor, 2004). This antipathy is generally evidenced by mutual ignoring, with only occasional periods of interaction between the groups. The increased attention paid to education in research has resulted in a paradoxical situation where people seem more willing to talk and somewhat less willing to listen than before. For this and other reasons, aspirations to combine multiple methods are more often voiced than achieved (Chatterji, 2004).

The policy makers

Who are the policy makers? In any context and in every country, policy makers can be found at national, state, or local levels. Education, as a central component of citizenship building and cultural expression, can be a hotly contested and jealously guarded jurisdiction. In an international context, countries have dramatically different education systems and ministerial models, from the decentralised structures of the federated countries (*e.g.*, Canada, Germany), to the highly centralised (*e.g.*, France, Korea). These different systems have implications for decision making and locus of control and uniformity of policy across geographic areas. They also have an impact on the amount and uniformity of funding available for various initiatives for change, including the support of research directed at providing evidence-based policy in education. Unified support on policy lines translates into very real capacity-building of research programmes and implementation of reforms in the classroom.

In recent years there has been a shift across most OECD countries to de-centralise decision-making in education, giving more responsibility and mandating to local authorities. Non-centralised leadership of course results in a number of different policy contexts and priorities within each country, and in some countries a concurrent fragmentation of the solicitation and funding of educational research. Nations that do not have strong central planning may see evidence-based policy issues as unaffordable, both financially and conceptually. Regional policy makers might have less incentive to contribute to the EIPR dialogue and less confidence that their knowledge needs might be met. For regions with very specific short-term knowledge needs (*e.g.*, improving student

performance on standardised tests and/or working to reduce inequities within the system without damaging the performance of the most privileged), the pooling of resources required to engage the national research community might be perceived as a time-consuming and fruitless procedure. Even if particular regions or individual policy makers were convinced of the importance of evidence-based policy, the lack of generalised agreement on policy priorities coupled with long-standing national rivalries and greater or lesser willingness to share information could very well lead to a situation where policy makers decide the time and energy needed to invest in an evidence-based policy approach make it simply not practical.

There is a very real limit to what policy makers (national, regional, and local) can influence in their respective systems, and a very real concern among policy makers with choosing their battles so as to maximise effectiveness. Ideas for education reform that originate from policy makers will be as varied as the system itself. Whether or not those ideas are based on evidence and research will depend, in large part, on the priority given to evidence-based policy research in education in that particular country and/or region, as well as pressures from and connections to the stakeholders (researchers, community members, parents, etc). It will also depend on the ability of policy makers to hear and absorb the information that is being presented to them for use in evidence-based policy, and in their ability to understand the fundamentals of evaluation.

Institutional leaders, practitioners, and the community context

In addition to professional researchers and policy makers, a number of other players appear who are involved at different stages in educational reform, from producing ideas to gathering evidence to assessing the results. The role of practitioners – teachers, other educational staff and their unions – in the production, and interpretation of research evidence has been attracting increasing attention. This is a recognition both of the potential contribution which they can make, and of the need to maintain their confidence in the reform process if it is to be successfully implemented. School and college leaders wield considerable power, and can support a culture which favours the production and the use of research evidence, or inhibit it. One of the best ways to gain the support of school leaders is to involve them in the research process and thus give them a sense of ownership over the initiative (Slavin, 2004).

Until recently, there has not been a great deal of support for the production and use of research evidence in the classroom. The teaching profession by and large do not see themselves as practitioner-researchers, learning on the basis of research into their own activities by their peers. This is in contrast to doctors, who are trained to use and contribute to the medical research agenda. Encouraging the understanding and use of research by school leaders themselves could be particularly important for pedagogical research, both in terms of the validity and generalisability of field-based studies, and also as a way to encourage implementing research-based reform.[1] If the research itself is valued and used (and conducted) by school leaders, the shared ownership will encourage its implementation in a way that something imposed externally by researchers on teachers will not. This requires a great deal of capacity building.

[1] A current activity in the Education and Training Policy Division of the OECD's Directorate for Education deals with school leadership in depth.

Bottom-up campaigning for change on the part of parents, students, and local communities, stemming from observations and effective practices in the classroom and at home, also has the potential to drive educational reform. School boards in particular have been highly active in campaigns aimed at changing educational policy and thus educational practice (*e.g.*, reducing class size, use of formative assessment, etc). Many if not all of these ideas are based on personal observation and experience and have not been tested empirically, a fact which does not reduce the conviction of the observer regarding the efficacy of these ideas.

Ideas which are generally perceived as "intuitively reasonable" gain power and support of public opinion. This is especially the case where they are promoted by the media, who often play a major role in shaping, or stunting, the policy agenda. They can then be used as a basis for policy change and educational reform regardless of whether there has been any empirical testing. How then, to go from the power of intuitive knowledge that resonates with established observations and practices towards a strategic attempt to capitalise on the ideas generated in the field, test them, and then, if proven to be effective, implement them in policy? How to encourage practitioners, the community, and media to look for the evidence supporting a reasonable idea (and its corollary: how to encourage researchers to communicate results in a broad and accessible fashion)? Given the scarce resources for education reform, the high level of public interest in education and the importance it holds for national and regional policy makers and the diversity of opinions within the education research community, the ability to assess what works in education is critical. The rise in the use and priority given to evidence-based policy research is a direct result of this understanding.

Methodologies and epistemological paradigms: what counts and should count as evidence

There are various practical reasons why countries choose to involve themselves or not in work on evidence-based policy. These include the length of time required to obtain evidence, the funding required to support the research, and the possibility of jurisdictional wrangling. In addition to these practical considerations, the dominant research tradition within education and the social sciences may prevent serious engagement in the debate. There is little sense in operating with a model which assumes a linear and rational process that translates good evidence into effective policy, as the reality is much more complicated (see Nutley and Webb [2000], for a discussion). We can even ask whether good research is more likely to be the product of an effective system than vice-versa.

Practical issues aside, the debate on EIPR opens up some quite basic philosophical issues to do with the nature of knowledge, and how different methodologies are suited to different knowledge claims. Causation is a particularly problematic concept, but one that demands attention from policy makers who are responsible for allocating resources and accountable for the effects of these allocations. The debate reaches into OECD's own work: OECD, and the Education Directorate within it, would certainly claim to base policy recommendations on evidence, but the nature of the evidence varies considerably. It includes large-scale primary research, notably the PISA study; national and thematic country reviews which draw on quantitative evidence but draw their originality from the expertise of the examiners and the quality of the examination process; comparative case study work which necessarily adopts a quite flexible framework; and secondary analysis or synthetic research at different levels of scale and ambition. It would be fair to

acknowledge that there is no unanimity within OECD on where exactly to draw the lines around what counts as evidence, nor on how it might be best used.

To countries with research traditions that are less used to empirical and quantitative methodologies in the social sciences, the discussion may seem completely divorced from the national reality. Evaluation as a component to policy (in the form of both pilot trials and ongoing assessment after the implementation of reform) may not seem necessary or helpful. OECD countries vary in the degree to which they attempt to objectively assess policy effectiveness. Even the idea of objective verification can be seen as unnecessary in the traditional context of a lack of understanding of research but clear intuitive agreement among the population (which of course includes teachers and policy makers).

In addition, some countries have a strong and rich tradition of qualitative methodologies such as action research and case studies, and education research has traditionally been thought of as their domain (St. Pierre, 2002). Case studies have been conducted as pilot projects for a variety of initiatives with the intention of scaling up the project if the experience was evaluated as favourable. Experience with the project in some of these cases is evaluated through interviews, questionnaires, and a variety of other qualitative research methodologies, which are then synthesised into a general evaluation of the project, including recommendations for change (Bogdan and Biklen, 2002). One of the strengths of qualitative research is precisely that it can give the depth of information required for, for example, recommendations for change or possible explanations of *why* something does or does not work. The policy research debate, dominated as it has been by disagreements on methodology and whether qualitative research is permissible at all, might be incomprehensible given national standards of research and education reform.

The debate regarding appropriate methodology stems from a very real concern about the overall quality of education research. It is unfortunately not difficult to find low quality education research with poorly derived research questions, inappropriate methodology and analyses, and misleading interpretation of the data. A low threshold for research quality has led to rifts within the community as well as damaging the reputation of education researchers in the eyes of other researchers (for example, social science researchers) (Feuer, Towne, and Shavelson, 2002). This in turn has led to difficulties for funders and confusion on the part of the policy maker.

In methodological terms one clearly identified area of weakness is in quantitative skills and the use of large data sets. This applies to assessment and evaluation data as well as experimental research design. Randomised controlled trials (RCTs) have received a great amount of attention and are clearly useful for causal questions and provide the rigour appreciated by funders. However, it seems clear that, as with all other methodologies, RCTs have both strengths and weaknesses and can be proposed as one of a set of appropriate methodologies for education research.

The urge to improve the overall quality of education research is one of the fundamental drivers of the EIPR discussion. The need for quality research and the appropriate methodology for questions of a causal nature (*e.g.*, does a programme have the intended effect?) has been thoroughly spelled out (*e.g.*, Angrist, 2004; Boruch, DeMoya and Synder, 2002; Cook, 2003). Hierarchies have been proposed to help guide evaluation of research aimed at addressing effectiveness issues, with preference being given to robust experimental and quantitative designs (Cook, 2005). For non-causal questions (*e.g.*, *how* does a programme work and why?), another set of methodologies is required, again with emphasis on the rigour of the investigation. The importance of deciding *first* on a research question and *then* choosing the appropriate methodology with

which to investigate the question is clear (Shavelson and Towne, 2002), yet often this point is overlooked by researchers and policy makers alike (Berliner, 2002). All this supports our basic proposition that there is no single best method for or type of evidence-based policy research. A variety of proposals have been advanced regarding how best to combine methodologies in education and other social sciences, with the view to providing concrete proposals and explicit strategies (Gorard and Taylor, 2004).

Part One of this publication takes a look at the issue in the form of a dialogue between two well-respected methodologists: Tom Cook from Northwestern University (United States) and Stephen Gorard from York University (United Kingdom). Each author first states his position in terms of the key concerns for education research methodology, and what counts and should count as evidence. For policy-making, is it a question of using the best available evidence, or should only the best evidence be used (which would imply that evidence that does not meet the gold standard could not be used)?

After positioning their arguments, the authors engage in a dialogue around areas of disagreement and unresolved issues. Much of the unresolved issues can be conceptualised under the heading of capacity-building in educational research and methodologies. The discussion can be boiled down to two main unresolved issues:

- Which forms of capacity are most in need of expansion/strengthening;
- How and by whom should this be done.

Capacity-building

As Cook and Gorard point out in Chapter 2, capacity building is required to encourage more participation in the evidence-based policy research discussion. This is true in a range of different contexts: in national and international discourse but also by educational researchers, policy makers of all levels, and teachers. Dyson and Desforges (2002, quoted in Sebba, 2004) distinguish between strategies to *broaden* and *deepen* capacity. The former is concerned with doing more of the same, but better; the latter with enabling the system to do new things. They identify three themes around which capacity can be built:

- Development opportunities for researchers (refresh, retrain, update).
- Infrastructure development (dedicated centres, IT, networks).
- Practitioner and policy makers' capacity to use and produce evidence.

Educational researchers, or the people doing research on education (see Cook, 2003, for a discussion on who is actually doing the research in education), need to be given enough training, financial support, and access to expertise to be able to conduct quality research. Additionally, there needs to be support from within the academic community for research that is policy relevant so that researchers who engage in this kind of research are not penalised in promotion ranking relative to colleagues who are pursuing a more traditional set of topics and publishing options. For those researchers aiming to provide policy-relevant work, specialised training is needed. At the moment, one area of clear weakness is that far too often researchers neglect to assess resource implications – when, for good reasons, this will be the second question the policy maker asks (after effectiveness). This requires knowledge of a range of techniques in addition to methodologies (*e.g.*, understanding opportunity costs).

Researchers need to be able to disseminate the results in a manner that can be understood by the general public, including the policy maker. Research results that remain within the realm of academia will not be able to be understood or accessed when needed, greatly limiting their impact. This is a lesson that has already been learned (at least to some extent) by researchers in other fields such as medicine and agriculture. In the multi-lingual international context, the dissemination of research results in a readily accessible and easily understood manner increases the likelihood that researchers can build off of each other's work, instead of squandering limited funds repeating the same basic research in various countries (we acknowledge that research must also be sensitive to cultural context to be relevant).

Policy makers may wish to reflect on why evidence-based policy is important and to understand the research process required to produce that evidence, allowing for the creation of realistic projects and deadlines for reporting on the part of the researchers. They might also be encouraged to think carefully about the importance of quality evidence and the need to trust the source of information (*i.e.*, use a "gate-keeper" to filter all available information such that only the best-available evidence is used for decision-making). On regional and national levels, this would facilitate the funding and solicitation of research that seeks to address questions pertinent to policy. It would also allow for more interface between the policy and research communities and (at least theoretically) more reasonable expectations from both sides. On local levels, support for evidence-based policy research from school boards and teachers is an important step in allowing funding to be allocated towards research with the intention of improving practice. It also allows for the possibility that funding might be made available to build capacity of the providers to help ensure faithful implementation of reforms. Alternatively, without accepting any obligation to fund research, boards might demand to see the external evidence on which a proposed innovation is based before approving a change in policy or practice.

Teachers need to be supported so that they have the time and energy to implement education reform in the manner it was intended as well as understand the research that fuels such reforms. This requires giving them access to research that is written for the non-scientist, as well as some background in research and how to interpret results. Otherwise how realistic is it to implement top-down policy change based on research in a system of professionals without giving them the capacity to understand and evaluate the research for themselves? The strategy of exposure to research methods and using research as part of the teacher training programme is one very concrete way to build capacity. In every national context and every kind of system, education reform is, ultimately, only as good as what actually takes place in the classrooms. Reforms unpopular with teachers or perceived as unnecessary (or worse, ill-conceived) will be at best half-heartedly implemented, at worst actively resisted. Cordingley (2000; 2004) offers a summary of factors influencing the use of research by teachers, including perceived relevance, evidence of learning outcomes, and clear links from the research to classroom practice. Garnering support from the people on the ground is one of the most strategic approaches to encouraging active implementation of evidence-based policy.

As mentioned above, a very good way to achieve this is to expect teachers not only to understand research, but take the lead in initiating it at local level to give them a sense of ownership. The capacity-building required for this (in research methods, in creating networks for experience sharing, in building support from management for the time and energy required, providing in-service training for ongoing development) is extensive and implies a re-thinking of existing structures and expectations. It also, rather fundamentally,

implies a degree of teacher and school autonomy such that teachers could act on the basis of research findings. The United Kingdom's Teaching and Learning Research Programme described in Part Three of this book specifically includes a significant capacity-building component.

Part Two: Mediating the Research/Policy Interface: The Role of Brokerage Agencies

Bridging the gap between internally and externally heterogeneous groups of researchers, policy makers, and educators is no easy task. One of the most common, and possibly most dangerous, ways to do this is to by-pass the communities themselves and turn instead to the Internet as a source of "research" on any given policy concern. The problem is, of course, that there is no system of quality control, and, if contradictory information is presented, there is no formalised process for consolidating or challenging evidence presented as research. The ability to assess the quality of evidence available – whether pulled from the Internet, presented by researchers, or offered by parents and teachers – is a fundamental prerequisite for informed policy-making.

One strategy to bridge this gap has been through think tanks and other intermediate agencies. These are most common in the Anglo-Saxon countries, but are becoming more common in continental Europe. They occupy a particular space in the divide in that they are not quite researchers, but not quite policy makers either. They have traditionally been quite aware of the power of public opinion and seek to harness it to lobby for particular causes and arguments. They are an important "bridging" institution in that they can be extremely effective and highly professional, but of course are not neutral, instead usually marshalling research evidence that would reinforce their particular policy priority.

A number of different initiatives aimed at bridging the divide between policy makers and researchers as well as assessing the quality of evidence available have been developed as general examples of "brokering". Brokering can be informal, *e.g.*, the exchange amongst colleagues of research evidence and information related to a policy issue at hand. It can also be more formal, *e.g.*, the creation of ties between national research institutions and their closest policy counterparts. In the past ten years this process has developed to the extent that formal brokerage agencies have been developed to officially facilitate both the process of information sharing and ensuring a certain level of quality control. Brokerage agencies vary in type and can be designed to be in-house and aid a particular Ministry to increase effective communication regarding the research and policy interface, evaluate proposed changes and policy recommendations, and assess the implementation of these programmes (*e.g.*, Norway). However most brokerage agencies have a broader agenda and seek to collaborate with as wide a community of researchers and policy makers as possible, so as to broaden the relevance of their work and findings.

Part Two of this publication provides a series of studies of existing and newly formed brokerage agencies. In lifespan, they range from EPPI-Centre (*eppi.ioe.ac.uk*) in the United Kingdom (formed in 1993) to the Knowledge Chamber of the Netherlands, which was created in 2006 as a result of the OECD/CERI workshop series that is also the basis for this publication. The Danish contribution describes the process behind the creation of a brokerage agency and the political and administrative aims guiding this process.

The brokerage agencies are also distinguished in their goals and means, with New Zealand providing an example of a brokerage programme embedded within the Ministry

that provides hands-on guidance to those wishing to conduct a synthesis of available evidence (*www.minedu.govt.nz/goto/bestevidencesynthesis*). In contrast the Canadian Council on Learning (*www.ccl-cca.ca/*), although also federally funded, is not embedded with the various provincial Ministries and, in addition to various other roles, is prepared to conduct reviews and syntheses based on policy priorities and within a very short timespan. Similarly the What Works Clearinghouse (*www.whatworks.ed.gov*) (United States) works in collaboration with a number of other institutes and subcontractors, to provide information and databases of research syntheses of replicable high-quality interventions, with a particular focus on the methodology of randomised controlled trials. Interestingly, it also conducts consumer surveys and questionnaires in order to ensure that the service it provides is meeting the demands of the users (researchers, practitioners, policy makers, etc).

As part of the continuing effort to draw parallels between education and other subject areas and learn from relevant experience, Part Two also provides an example from the field of social care in the form of the Social Care Institute for Excellence (*www.scie.org.uk*) (United Kingdom). From this contribution we can observe both the similarities of experience, goals, and efforts of a brokerage agency in this domain, as well as some shared challenges.

All of these centres have the goal of encouraging dialogue between policy makers, researchers and educators with the aim of providing tools and capacity-building within these communities to evaluate what works and what does not work in education. An important first step in this process is the creation of a database of quality education research on particular topics that are of interest to policy makers, as well as providing clear goals and criteria for conducting and evaluating educational research. These criteria serve as a baseline for conducting reviews of research, reviews which can then be used to provide systematic evidence as to the effectiveness of particular policy objectives or classroom practices. A key component to these brokerage agencies is the transparent exchange of findings: all reviews are available on the various brokerage agency websites, and all methodologies used by the review process are defined in detail. Many of the centres require reviewers to commit to updating their work on a regular and pre-defined basis, so as to include new evidence and maintain a state-of-the-art synthesis on each particular topic. And all of the centres have a commitment to disseminating research results to as wide an audience as possible, in order to work on affecting top-down and bottom-up change to the system.

The agencies aim to address one of the key issues identified in the OECD's 2003 report on *New Challenges for Educational Research*, and elsewhere (*e.g.*, Raudenbusch 2005): the accumulation of knowledge. Educational research is conspicuously weak in its ability to continuously develop and refine a body of knowledge which is quasi-universally acknowledged as well-founded. The research community, through its induction and training procedures, has the crucial role in this, but brokerage agencies can also have a major part in designating the most recent authoritative additions to the knowledge pile.

Such brokerage agencies can and do play a key role in bringing together the disparate communities and bridging the gaps in the EIPR process. They have provided resources and tools for researchers, policy makers, and educators to openly engage in the discussion of what works in education, and allowed for capacity-building in each of those domains. They are potentially a vital mechanism in aligning supply and demand and there are valuable formative lessons to be learnt from their experiences to date. As Part Two of this

publication demonstrates, standard challenges that have yet to be resolved by these agencies include how best to:

- incorporate all stakeholders into the process;
- address the tension between the time required for solid research and the necessity of quick results for policy-making;
- disseminate findings to all stakeholders, including media, parents, and students;
- ensure sustainability and stability of funding.[2]

Part Three: Evidence-based Policy Research in Practice: Examples from the Field

The examples from the field are drawn from a number of different countries (Canada, Finland, Singapore, United Kingdom) as well as a variety of different models of how best to put the use of evidence into practice. Canada's contribution looks at the launch of the National Children's Agenda, which focused on the long-term vision of fostering good Canadians by aiming for lifelong learners, productive workers, nurturing parents, and engaged citizens. In order to best achieve this, a number of decisions were made regarding key developmental outcomes for children and the need for a body of evidence from multiple data sources and analytic methods. This national data system has flourished in part because it was intended to provide reliable and stable flows of data and as such was assured over time and protected from short-term budget cuts and thinking.

In contrast Life as Learning (Finland) was set up as a discrete project of the Academy of Finland and ran from 2002-06. As a national research programme, it aimed to encourage the development of a research culture and support interdisciplinary and international research projects. A number of interesting developments have already come out of this process and have been well-received in Finland – however the time-bound nature of the exercise and funding difficulties discussed in the paper speak to a challenge shared by many research communities.

The contribution from Singapore is unique in that it is a non-OECD country with an unparalleled set of funding and research opportunities, including a strong database with data on the entire student population. This reflects its major commitment to a profile as a knowledge-based high-skill economy. In addition, this breadth and depth of information allows for comprehensive investigation into a variety of education issues and also permits a particular emphasis on arguments presented to policy makers. In this sense this example is unique not only in the methods, funding, and scope of the research, but also in the relationships involved between the major stakeholders.

Lastly the United Kingdom's Teaching and Learning Research Programme (TLRP) coordinates over 500 researchers in 60 project teams. An immensely complex project with assured medium-term funding (2000-2011), it aspires to improve the quality of education research in the United Kingdom as well as contributing new knowledge. An explicit emphasis is how best to disseminate activities, results, and events of the

[2] OECD/CERI proposes to investigate brokerage agencies more thoroughly in its 2007-08 Programme of Work, addressing the following key questions: What are the different roles played by brokerage agencies in different countries? And how effective are they at this role, and how have they attempted to address the challenges outlined above?

programme, and as such there are active ties to policy makers, practitioners, community members and media in addition to the researchers.

These examples touch on the key issues of the preceding sections, allowing us to look at how the various interactions play out in a real world situation. For example, all of the contributions allow us a close look at the quality of the relationships between the primary stakeholders (*e.g.*, policy makers, researchers, practitioners, etc), and particularly the two points as raised in the preceding section on roles and relationships:

- How well they function in terms of information flows, trust levels.
- What could be done to improve the quality in each case.

In addition, the examples all provide the answer that that particular programme has chosen to the question posed in Part One: What Counts as Evidence? Both Canada and Singapore provide cases of extensive and thorough data collection, using all means available (in the case of Canada, the focus is on socio-demographics rather than education, whereas Singapore uses its unique funding and database opportunities to compile an unparalleled evidence-base). Finland and the United Kingdom, on the other hand, mix models of methods and types of service delivery to best serve the question and mandate of the organisation.

All of the programmes must also concern themselves with the dissemination of their findings and the best way to bridge the gap between the evidence they can provide and the needs of the policy maker, and in this respect all of the programmes benefit from their close ties to government. The UK's Teaching and Learning Research Project in some senses acts as its own brokerage agency, with dissemination and communication of results playing a central role in project planning and development. Life as Learning (Finland) faces perhaps the most traditional challenge of a research project, in that it must strive to maintain a high profile and disseminate its results. In contrast both the Canadian and Singaporean examples are much more closely tied to government and their respective Ministries and so the brokering in some sense has already been done (*i.e.*, the research is high priority and centrally incorporated in policy-making).

This part thus serves as both a set of concrete examples interesting in their own right and as a way to observe the various theoretical issues discussed in this volume in practice. Similarly to Part Two, it includes an example from outside the education field, in this case a programme governed by the Department of Human Resources and Social Development (Canada).

Part Four: The Politicians' Perspective

The evidence-based policy research debate has to large extent been focused on the major players of policy maker and researcher. In Part Four of this publication we hear the voice of the politicians, charged with making decisions and evaluating the best way forward for the education system in their jurisdiction. In this part we hear from Johnny Nilsson, the former Secretary of State for Education in Sweden, who speaks of the experience of the past and some of the limitations and concerns of the politician in a period without the current emphasis on evidence-based policy. We hear also of the serious weight that has been placed on this process in the Netherlands, and how and why this work was so important to the Minister of Education, Science, and Culture, Maria van der Hoeven. The Netherlands has a number of initiatives that have followed from the OECD/CERI work and are looking to the future.

In addition, Jane Davidson of Wales uses her experience as the longest serving Minister in the United Kingdom to discuss the role of evidence in policy-making and the struggle she has engaged in to bring it to the forefront of the policy-making process. Finally we have a unique case study from Poland, which has recently brought together current and former Ministers and decision-makers in an on-going effort to raise the priority and profile of education research in policy-making in that country. Jerzy Wisniewski, former senior civil servant and CERI Governing Board member, contributes the challenges and aspirations of a country that is just beginning to focus on these issues.

Concluding note

The debate on what counts as evidence will certainly recur. In another dozen years we may be noting the same weaknesses in educational research, and the same flaws in the communication between research and policy on education. Or fresh obstacles may have emerged, from some of the trends observed in this chapter: the massive diffusion of information of varying quality, without established intermediaries; the pressure on policy makers to reach decisions, whether or not good evidence is to hand; or further modifications in the relationship between different parts of the decision-making process such as a tightening grip of finance ministries on education spending. But some progress will also have been made, in all probability; we can guess that rigorous research techniques will become more widely understood and applied; that practitioners and perhaps also policy makers will broaden their evidence base; and that the potential of brokering will have been explored in many countries. But that is to write the history of the future (Attali, 2007); for now it is only to be hoped that lessons from the past can be learned.

References

Angrist, J. (2004), "American Education Research Changes Tack", *Oxford Review of Economic Policy*, Vol. 20(2), pp. 198-212.

Attali, J. (2007), *Une brève histoire de l'avenir*, Fayard, Paris.

Berliner, D.C. (2002), "Education Research: The Hardest Science of All", *Educational Researcher*, Vol. 31(8), pp. 18-20.

Bogdan, R.C. and S.K. Biklen (2002), *Qualitative Research for Education: An Introduction to Theories and Methods, 4/E*, Allyn and Bacon, Needham Heights, MA.

Boruch, R., D. DeMoya and B. Synder (2002), "The Importance of Randomized Field Trials in Education and Related Areas", in F. Mosteller and R. Boruch (eds.), *Evidence Matters: Randomized Trials in Education Research*, The Brookings Institution, Washington DC.

Chatterji, M. (2004), "Evidence on 'What Works': An Argument for Extended-Term Mixed-Method (ETMM) Evaluation Designs", *Educational Researcher*, Vol. 33:9, pp. 3-13.

Cook, T. (2003), "Why Have Educational Evaluators Chosen not to Do Randomized Experiments?", *Annals of American Academy of Political and Social Science,* Vol. 589, pp. 114-149.

Cook, T. (2005), Consultant's Report, Meeting on Evidence-based Policy, Stockholm Sweden, OECD, Paris.

Cordingley, P. (2000), *Teacher Perspectives on the Accessibility and Usability of Research Outputs*, paper presented at the British Educational Research Association annual meeting.

Cordingley, P. (2004), "Teachers Using Evidence: Using What we Know about Teaching and Learning to Reconceptualize Evidence-based Practice", in G. Thomas and R. Pring (eds.), *Evidence-based Practice in Education*, Open University Press, New York, pp. 77-90.

Dyson, A. and C. Desforges (2002), "Building Research Capacity", National Education Research Forum, see *www.nerf-uk.org*

Feuer, M., L. Towne and R. Shavelson (2002), "Scientific Culture and Educational Research", *Educational Researcher*, Vol. 31(8), pp. 4-14.

Gorard, S., and C. Taylor (2004), *Combining Methods in Educational and Social Research*, Open University Press, England.

House of Commons Science and Technology Committee (2006), *Scientific Advice, Risk and Evidence Based Policy Making*, Document HC 900-1, The Stationary Office Limited, London.

Levin, B. (2004), "Making Research Matter More", *Educational Policy Analysis Archives*, Vol. 12(56), pp. 1-20.

Nutley, S. and J. Webb (2000), "Evidence and the Policy Process", in H.T.O. Davies, S.M. Nutley and P.C. Smith (eds.), *What Works? Evidence-based Policy and Practice in Public Services*, The Policy Press, England, pp. 13-42.

OECD/CERI (2003), *New Challenges for Educational Research*, OECD, Paris.

OECD/CERI (2004), *Innovation in the Knowledge Economy: Implications for Education and Learning*, OECD, Paris.

Raudenbusch, S.W. (2005), "Learning from Attempts to Improve Schooling: The Contribution of Methodological Diversity", *Educational Researcher*, June/July, pp. 25-31.

Sebba, J. (2004), *Developing an Evidence-based Approach to Policy and Practice in Education*, UK Higher Education Academy, Discussion Paper 5.

Shavelson, R.J. and L. Towne (eds.) (2002), *Scientific Research in Education*, National Academy Press, Washington DC.

Slavin, R. (2004), "Translating Research into Practice on a Large-scale: Lessons from Success for All (US)", Keynote presentation at DfES research conference, "Research in Education, What Works?", 19 November, London, England.

St. Pierre, E.A. (2002), "Science Rejects Postmodernism", *Educational Researcher*, Vol. 31(8), pp. 25-27.

Stern, N. (2006), *Review Report on the Economics of Climate Change*, HM Treasury, London.

What Works Clearinghouse.org (2005), *www.ed.gov/nclb/methods/whatworks/ww/13*.

Chapter 2
What Counts and What Should Count as Evidence

Thomas Cook, Northwestern University
Stephen Gorard, University of York

In this chapter, we present our views on the place of experimental design in evidence-based policy-making and practice. Learning where we agree should help readers identify where they can be relatively confident about method choice. Learning where we disagree may help them identify which method choice decisions remain problematic, and where maximal caution is required in evaluating claims about new knowledge for improving the outcomes from education.

Introduction

The main issue we address in this brief chapter is the place of experimental design in evidence-based policy-making and practice. We were asked to write this chapter because one of us, Thomas Cook, has written widely on randomised control experiments in education (Cook and Campbell, 1979; Cook, 2002), detailing their merits, outlining the assumptions and threats to validity associated with them, and listing and refuting most objections usually raised to doing them while accepting some. The other of us, Stephen Gorard, is probably better known for writing about the value of mixing methods of different kinds (Gorard, 2001; Gorard with Taylor, 2004). Our difference in research emphasis speaks to a vexing issue of method choice that currently bedevils the educational research community as it seeks to ground education policy decisions in better evidence. As usually framed, the issue is: Should policy-centered education research be predominantly experimental or based on mixed-method studies? As will become clear, our positions are quite similar when we address such general propositions. But we differ in some particulars of great importance for deciding which kinds of education research to commission in order to improve the policy yield of education research. Learning where we agree should help readers identify where they can be relatively confident about method choice. Learning where we disagree may help them identify which method choice decisions remain problematic and where maximal caution is required in evaluating claims about new knowledge for improving the outcomes from education. Each of us will first individually present some research principles and propositions, and we then later draw together our areas of agreement and difference.

Thomas Cook's propositions

1. Educational policy speaks to many different kinds of issue and question, most associated with different method preferences. So, comprehensive "evidence-based research must be multi-method".

Among other issues, educational policy has to be concerned with "who gets what?"; "what does a given educational service cost?"; "what is classroom life like?"; "how well are students performing?"; "how are teachers trained?"; and "what works to improve student performance?". The majority of these questions are descriptive; only the last is explicitly causal. Theorists of method in the social sciences broadly agree that the best methods for dealing with non-casual issues require theory, ethnography, interviews and surveys, among other methods. Experiments hardly help. If educational research is to speak to the comprehensive knowledge needs of the education policy community, it can, should and must involve multiple methods. Framing the issue as a choice between experimental or mixed methods is silly. Even questions that seem purely causal at first glance are embedded within contexts where we also need to know: "who gets the new educational practice under evaluation?"; "what does the program cost?"; "which social values does the intervention speak to?"; and so on. Even the major institutional advocate of experiments today, the Institute for Educational Sciences of the United States Department of Education, routinely commissions experimental evaluations that also include theoretical analysis of the programme under review and observational measures of programme implementation. It also funds many, and some very large, non-experimental surveys of educational resources and performance, like the National Assessment of Educational Progress (NAEP). Arguing for mixed method research is anodyne, given the heterogeneity of knowledge needs in education and the research design practices of even the most passionate advocates of experiments. The debate needs to be framed differently: about (1) the priority to give to causal versus non-causal issues in educational research today; and (2) when causal questions are central – and only then – the priority that should be given to randomised control experiments versus other causal methods. I address these two basic themes in the points below.

2. Causal questions have a special importance in educational policy research.

My rationale for this assertion is that policy makers are selected or elected to make decisions. These decisions often touch on how to change schools and colleges to raise the performance of teachers and students. This is always a pressing concern, but especially in nations where comparative studies like PISA indicate disappointing levels of average performance. But even in nations currently doing well, novel ideas are needed if they are to maintain their relatively high standing. Where are these ideas to come from, and how should they be tested before being implemented on a broad scale? I believe many descriptive issues are important in education; but identifying "what works" deserves a special status among the concerns of those accountable for the quality of educational performance, as does learning about "what works" in the most secure ways. Moreover, I also believe that the need to learn what works is especially acute right now, raising even more the priority of gaining accurate causal knowledge in education. The main reason for believing this is immediately below.

3. The causal knowledge now being generated in education is inadequate for providing a secure stock of knowledge about effective educational practice.

Empirically-based causal assertions are rampant in today's educational research, very few of them the product of experiments. How valid are they in general? No definitive answer is possible, given that an answer depends on the very standards of evidence that are in contention among educational researchers today. But in the countries I know best – the United States, the United Kingdom, France and Germany – no secure body of literature exists that policy makers can rely upon to learn what should be changed in schools in order to improve student achievement and social behavior. Cacophonous claims about effective practices abound. But we will later see that their technical warrant is generally weak when evaluated by the most widely accepted causal methods in statistics and across the social sciences as a whole, as opposed to the standards currently operating in large parts of the educational research community. When the fundamental values buttressing policy choices are at issue, all educational policy makers should welcome active dispute since contention about values is the mother's milk of democracy. But to welcome dispute about the effects of discrete educational practices is another matter. Evidence-based policy depends on a reasonably clear *research-based* consensus about effective practices as one central input into decision-making, though all decision makers realise that total consensus is impossible. Yet typically decision makers do not get even an approximation to consensus. Some decisions have been endorsed by education researchers in the past and were widely disseminated without much quality research evidence to back them up. Some of these turned out to be quite disastrous – *e.g.*, new math and whole language reading instruction. I believe that causal issues are central to educational policy and that the causal knowledge generated by educational researchers to date has generally not been trustworthy. So the key is to learn more about what works in education. One proposal to do this involves radically increasing the incidence of random assignment experiments, since in Cook's (2002) review of the relevant literature they constitute from 1% to 5% of all the educational research studies that claimed a causal finding. Why stress experiments?

4. For answering causal questions, the randomised experiment is well warranted theoretically and empirically.

The theoretical warrant for experiments comes from a minor variant on the same statistical theory that undergirds the highly successful survey research industry. This minor variant uses statistical theory to create, not a single sample that formally represents the population from which it was drawn, but two or more samples that represent the same population, whatever it might be. Since the groups so created are initially identical on expectations, any final difference between them must be due to whatever intervention one group had that the other (or others) did not. However, this is not the only warrant for experiments. Over the last decades we have had considerable experience implementing them in sectors other than education and even some experience in education, albeit primarily in the United States. We are fast learning how to improve their implementation in order to regularly meet all the assumptions the method requires. The survey research industry could not exist without both a statistical theory and decades of wisdom (much from small-scale experiments) about how to implement surveys so as to reduce bias. The needed statistical theory already exists for experiments, and knowledge is being quickly accumulated about how to implement them more often and better (Cook, 2002). I do not

want to argue that experiments are perfect, only that they are superior to their current alternatives. Their imperfections are of several kinds.

5. The valid causal interpretation of experiments depends on assumptions being met.

To produce unbiased causal results experiments require several assumptions that are routinely described in method texts. The major ones are that a correct random assignment procedure is chosen; that it is correctly implemented; that no differential attrition occurs across the groups being compared; and that contamination of the intervention details from one group to another is minimal. Also, the analysis of experiments depends on standard statistical assumptions being met, as do other causal studies too. Each of these assumptions can be violated, but methodologists know about them and about how to avoid or limit their influence in complex settings like schools and colleges. However, while many educational researchers know about the necessary statistical theory, far fewer of them are experienced in implementing experiments so that their assumptions are demonstrably met in school-based research, and on a quasi-automatic basis (Cook and Foray, in press). Experiments are only sufficient for unbiased causal knowledge when the above assumptions are demonstrably met, and meeting them is not difficult for those with experience conducting experiments.

6. Being limited in their capacity to generalise causal findings, experiments do not always answer the question of greatest policy relevance.

Many experiments are limited to those schools, teachers or students that agree to whatever treatment they are assigned by chance (Cook, 1991). The causal findings so generated will be bias-free, but only apply to those who volunteer for a random assignment study. Other types of causal study will also depend on volunteers, but not necessarily volunteers of the same kind. Experiments have other restrictions to their generality. They do not guarantee that any obtained effects will hold in the future; and the effects of an intervention may change if it is implemented on a much broader scale that leads to different causal processes being involved in the smaller experiment than in the extrapolation to, say, an entire nation. Once again, though, these restrictions apply to varying degrees to other kinds of causal study too. The limited generalisation of findings from single experiments helps explain why advocates of experiments prefer policy to depend on multiple experimental studies, each with a different population of persons, settings and times as well as on different ways of instantiating the intervention and measuring the outcome. Alternative causal methodologies are also limited in their capacity to generalise, although not all in the same ways as experiments. What are these alternatives? And how good are they? We must answer this to support the claim that experiments are marginally superior to their alternatives, albeit not perfect.

7. In human history, valid causal knowledge has often come from non-experimental and non-quantitative sources.

It would be preposterous to maintain that experiments are necessary for causal knowledge. Our ancestors learned about the causal effects of making fires millennia before there was formal experimentation. And scholars knew that out-group threats usually cause in-group cohesion long before R.A. Fisher created the first formalisation of experimental design. The case for experiments is that they are needed for detecting

effects that are smaller than many of the others humans have learned about in the past. We have learned from studies of educational performance net of various student background characteristics that, within the limits of the models used, schooling effects are indeed very small and swamped by individual differences, particularly familial and psychological ones, not to speak of the genetic ones still to be examined in detail (Coleman, 1966; Jencks, 1972). This may be why the Institute for Educational Sciences designs its evaluations to detect achievement gains of $1/5^{th}$ of a standard deviation – typically over several years and thus equivalent to a total of about one year's change in growth over these years. As important as such effect sizes are, they are not obviously "large" and are manifestly far from transformational. Experiments are also needed because many educational practices that might be effective are enmeshed in real-world school or college life within complex systems involving many other variables. This makes it difficult to identify the unique causal role of any one educational practice, or set of practices, unless these practices have first been isolated and then systematically varied.

8. In social science, experiments are not the only method known from theory to be capable of generating unbiased causal knowledge.

Four alternatives to the experiment are known to generate unbiased casual inferences under certain conditions. (1) From statistical theory and comparative empirical research reviewed in Cook (in press) and in Cook and Wong (in press), we know that regression-discontinuity studies can produce the same causal estimates as experiments. These studies depend on an educational resource being distributed according to an eligibility score along some quantitative continuum, often a specific level of need or merit but sometimes a specific date of birth or order of applying for the service under review. The key is that everyone on one side of the eligibility score receives the service and those on the other side do not. (2) We also know from theory that instrumental variable approaches can result in unbiased causal inference when an instrument is found that is correlated with the treatment but not with errors in the outcome (Angrist, Imbens and Rubin, 1996). We also know that casual inferences are unbiased if (3) the process of assignment to treatment is perfectly known or (4) the outcome is perfectly predicted (Cronbach, 1982).

9. These theoretically unbiased alternatives have assumptions that cannot be as clearly met in actual research practice, making them technically inferior to the experiment.

Regression-discontinuity has less statistical power to detect effects than the experiment (Trochim, 1984), and it depends on strong assumptions about the functional form of the relationship between the assignment variable and outcome (Rubin, 1977). As for instrumental variables, it has proven very difficult to find many of them that meet the requirement of being uncorrelated with the outcome – the ironic exceptions being random assignment (Angrist, Imbens and Rubin, 1996) and regression-discontinuity (Hahn, Todd and VanderKlauuwe, 2001). Most causal claims to date using such instruments, particularly in economics, have been hotly contested and thus limit our confidence that an instrumental variable approach can be widely used to promote causal inference. Both random assignment and regression-discontinuity derive their intellectual warrant from the fact that the process of assignment into the different treatment conditions is completely known and hence easily modeled. This is not the case with quasi-experiments or non-experiments where attempts are made to model the treatment assignment process. Empirical research on attempts to do this via selection models (Heckman, 1979) and

propensity scores (Rosenbaum and Rubin, 1984) shows that these statistical tools nearly always fail to recreate the results of experiments that share the same intervention group and so vary only in how their control group is formed – at random or not (LaLonde, 1986; Glazerman, Levy and Myers, 2003; Cook, Shadish and Wong, 2007). So full knowledge of the treatment assignment process has not yet turned out to be a viable and practical causal tool. And it is almost always impossible in actual research practice to totally predict any educational outcome, even when schools are the unit of study. The foregoing implies that the main case for preferring experiments is that they are practically superior to the other causal methods known from theory to be unbiased.

10. Many other methods are also currently used for supporting claims about what works in education, but they are generally inferior because they do not enjoy an independent theoretical or empirical warrant as unbiased.

A great array of other methods is used to justify causal claims in education. They range from site visits to countries that are performing well in PISA through to highly statistical difference-in-differences or causal modeling studies. Also included are ethnographic accounts, secondary analysis of survey data, and quasi-experiments. None of them enjoys an independent and theoretically infallible warrant sufficient to justify the causal knowledge gained. The shortfalls are many and vary by method. Suffice it to note here that Campbell and Stanley (1963) and its successors (Cook and Campbell, 1979; Shadish, Cook and Campbell, 2002) have detailed many threats to the validity of causal conclusions that are associated with even the better of these study types. Moreover, Glazerman, Levy and Myers (2003) have documented how practice among economists, including some who work in education, regularly fails to produce the same results as experiments that share the same treatment group. The absence of both a theoretical and empirical warrant for the many types of study from which causal conclusions are regularly drawn in education today could well be a major reason why so many causal claims have failed to stand up to hard scrutiny and have not led to clear cumulative learning about what works.

11. In many sectors where policy is currently made, experiments enjoy more credibility than other kinds of causal study.

This is the case in health, public health, agriculture, the prevention sciences, criminal justice, and legal studies of compliance with gender- and race-based hiring laws. And even in survey research, improvements to practice have often depended on experiments. They are also common in research on early childhood education in the United States where Congress requested that its largest national programme, Head Start, be evaluated experimentally. Also, the pre-school studies regularly cited to promote the "universal preschool" policy in the United States are held in such high regard because they are experimental and involve decade-long effects on children's lives (Schweinhart, Barnes and Weikart, 1993; Reynolds *et al.*, 2001; Ramey and Campbell, 1991). To advocate against randomised experiments requires a compelling argument that schools are systematically different from institutions in other sectors in ways that either make experimentation infeasible or bias the results obtained. Such advocates also have to explain why experiments are common both in pre-schools and in school-based research with prevention rather than academic achievement outcomes. It is important to note that experimentation does not exist in a vacuum.

12. Any single experiment assumes prior knowledge that need not itself be the product of experiments.

Experiments require prior substantive theory and the experience of persons knowledgeable about what is feasible in school life. They also require the availability of good measures of the preferred outcomes, or the ability to construct such measures. Further, they require at least local political and administrative support for the study. And finally, they depend on prior causal studies. These can be experiments, but need not be so in order to confer marginal advantages for constructing future experiments. For instance, statistical power calculations depend on variance estimates from other studies, as do bigger picture issues like how an intervention is conceptualised, chosen and implemented. All experiments build on the shoulders of prior scholars in theoretical and applied fields. They do not exist in a methodological vacuum, and experimenters are not a new priesthood that can afford to declare itself independent of educational research' past.

13. Having information from experiments does not guarantee that this information will be used in policy debates, and certainly not used to form a decision.

Although experiments give a marginally superior causal answer compared to other methods, this does not guarantee that these results will be more often used in debates about educational change. And when evidence from experiments is used, it certainly does not mean that they will alone shape policy decisions. The history of educational research is replete with examples of study results not apparently used; and in democracies decision-making does, and should, depend on many factors other than scientific knowledge alone.

14. But having scientific information from experiments probably increases the odds of the information being used in policy debates.

It is difficult to argue this point for education today, given the recent history of school-based experiments with random assignment. However, in other fields of study, causal results from experiments are routinely preferred over the results from other kinds of study. This is especially true in medical, public health and prevention contexts, and also when the results from multiple studies are synthesised in search of an effective policy option. Indeed, it is standard practice in meta-analyses to analyse the results from experiments separately and to add non-experimental results to the review only if their average effect size does not differ from that from experiments (Cook *et al.*, 1992). This is even the case in those rare educational instances where a very large number of studies of an intervention exist, creating enough experiments to analyse separately even if they are but a tiny fraction of the whole corpus of studies – for two instances in early childhood reading, see Ehri (2001a and b). In more qualitative review contexts, at least in the United States, expert panels commissioned to review the literature for a governmental agency often pay special attention to the experiments in formulating conclusions for policy consideration within a government agency, deliberately giving them more weight than the non-experimental evidence.

In conclusion, the argument is that learning "what works" is crucial in educational policy-making, and that it is especially a problem today. This is because we have failed over the last 30 years to accumulate a secure body of knowledge about effective educational practices. So I believe that the case for more causal research is clear – that is,

relative to other kinds of study with a claim on educational research funds. To do more experiments does not mean that only experiments are valuable and that only they should be funded. But it does mean that they deserve, at least temporarily, a higher profile than they received over the last 30 years or so. But only if the causal studies provide more secure causal knowledge of what works, and the best method for achieving this involves doing experiments, given their independent warrant in statistical theory and also in past practice in sectors outside of school-based education. Experiments are not perfect. But no other method currently exists that does as well, and this is broadly acknowledged in sectors other than education. But it is also acknowledged in two sectors with close links to traditional education – in research on cognitive outcomes in pre-schools and on prevention outcomes in research in schools. Experimentation is not a novelty in school-based research; merely something whose sphere of application needs to be extended to meet a commitment to learn more about what works in a context of international crisis about educational performance levels in many larger countries.

Stephen Gorard's propositions

Like Tom Cook, I shall set out a number of summary propositions. Interested readers can trace the further basis for these propositions in my research writings – examples of which are provided. In my own writing I am concerned with education as an area of public policy, including pre-school, post-compulsory, and adult, provision, whereas Tom Cook writes for the context of schools. I see no reason why this difference should affect our methods approach.

1. A key ethical concern for those conducting or using publicly-funded education research ought to be the quality of the research, and so the robustness of the findings, and the security of the conclusions drawn.

Until recently, very little of the writing on the ethics of education research has been concerned with quality. The concern has been largely for the participants in the research process, which is perfectly proper, but this emphasis may have blinded researchers to their responsibility to those not participating in the research process. The tax-payers and charity-givers who fund the research, and the general public who use the resulting education service, have the right to expect that the research is conducted in such a way that it is possible for the researcher to test and answer the questions asked. Generating secure findings for widespread use in public policy could involve a variety of factors including care and attention, sceptical consideration of plausible alternatives, independent replication, transparent prior criteria for success and failure, use of multiple complementary methods, and explicit testing of theoretical explanations through randomised controlled trials or similar experimental designs (Gorard, 2002a).

2. It is helpful to consider the research enterprise as a cycle of complementary phases and activities, because this illustrates how all methods can have an appropriate place in the full cycle of research.

Experimental designs, like in-depth work or secondary analysis, have an appropriate place in the cycle of research from initial idea to development of the results. The main reason to emphasise experiments at this point in time is not because they are more important than other phases in the cycle, but because they represent a stage of work that is

largely absent in education research. If nearly all of education research were currently conducted as laboratory experiments then I would be one of the commentators pleading for more and better in-depth work or secondary analysis, for example. Other weak points in the cycle are currently the systematic synthesis of what we already know in an area of work, the design or engineering of what we already know into usable products for policy and practice, and the longer-term monitoring of the real-world utility of these products (Gorard with Taylor, 2004; Gorard, Rushforth and Taylor, 2004).

3. Working towards an experimental design can be an important part of any research enterprise, even where an experiment is not envisaged or even possible.

Sometimes a true experiment, such as a large randomised controlled trial, is not necessary, and sometimes it is not possible. An experiment is not necessary in a variety of research situations, including where the research question does not demand it, and where a proposed intervention presents no *prime facie* case for extended trialling. An experiment may also not be possible in a variety of research situations, including where the intervention has complete coverage, or has already been implemented for a long time, and where it would be impossible to allocate cases at random. However, a "thought experiment" is always possible, in which the researcher considers no practical or ethical constraints except answering the research question as clearly as possible. In then having to compromise from this "ideal" to conduct the actual research, the researcher may come to realise how much more they could be doing. There might then be more natural experimental designs, more practitioner experiments, and surely more studies with appropriate comparison groups rather than no explicit comparison at all (a situation which reviews show is the norm for UK academic research in education). There might also be more humility about the quality of the findings emanating from the compromise design (Gorard, 2002b, 2003a).

4. Part of the problem of research quality lies in traditional research methods training and "experts".

In the United Kingdom, traditional methods training for new researchers in university departments of education generally starts by introducing students to differences between types of research, and emphasising the purportedly incommensurable values underlying the variety of approaches to discovery. Most obviously, researchers are introduced to a supposed paradigmatic division between "qualitative" and "quantitative" studies in a way that encourages methods identities based on a choice of only one of these "paradigms". This leads many of us to indulge in paradigmatic strife, or write off entire fields of endeavor – as being "positivist", for example. Some commentators try to heal these schisms after they have been created, but there is a shortage of texts and training resources that take the far superior approach of assuming that there is a universal underlying logic to all research. Such an approach leads from the outset of training to a focus on the craft of research, thus bringing design, data collection, analysis, and warranting results to the fore, leaving little or no place for paradigms (Gorard, 2003b, 2004a).

5. Part of the problem of research quality lies in a lack of appropriate use of numbers.

One of the main reasons why there is not more mixed methods education research is clearly that there are few researchers willing and able to work with numbers. Since experimental designs are seen by many, incorrectly, to be "quantitative" in nature, this could also be part of the reason for the lack of experimental work. There may be a number of influences at play here, including poor maths teaching in schools, lower ability of social science students in comparison to other disciplines both in terms of maths and perhaps also overall, the selection of methods courses by students in terms of perceived ease, and the widespread misunderstanding that being a "qualitative" researcher means never having to deal with numbers. However, I am coming increasingly to the view that a major share of the blame lies with "quantitative" researchers. They seem to prefer devising more and more complex methods of analysis rather than devoting their energy to creating higher quality datasets that are easier to analyse. They often present their research in exclusive and unnecessarily technical ways. They generally assume, incorrectly, that numbering is the same as measuring, that reliability is the same as validity, that probabilistic statistics can be used with purposive samples or even with population figures, and that any use of numbers must be based on sampling theory. This is not the way forward (Gorard, 2006a, 2006b).

6. Part of the problem of research quality lies in an unwillingness to test our cherished theories.

Another element of the methods crisis stems from our love of specific theories, and our consequent unwillingness to test them for failure. A typical piece of evaluation in UK education is either commissioned by, or conducted by, those responsible for the programme being evaluated. There may then be pressure from funders to "finesse" the results. I have certainly been contacted by evaluators seeking some new kind of analysis that will gainsay the surface findings, and which will support instead their underlying belief that the programme must be being effective. This is no different, in principle, to the dredging of data that goes on shamelessly *post hoc* in other forms of research as well. I have also experienced far too many cases in which researchers simply make up or distort data in order to help preserve their prior beliefs. Some methods experts actually advise researchers to "take sides" before conducting research, and not to publish negative or otherwise unhelpful results. Of course, it remains true that the evidence-based approach to policy-making and practice is itself untested in education, and still far from fully satisfactory in fields such as health sciences. But this is a reason to test it, not to reject it out of hand (Gorard, 2004b; Gorard and Fitz, 2006).

7. Much of the solution lies in greater scepticism, because the problem is not really one of methods at all.

Some of the criticism of education research during the 1990s was concerned with relevance. But education is a very applied field of research. I do not find much published research that has no relevance to some important or useful component of education. The criticism is more properly about the poor quality of much research, so that even though the findings may have relevance they still cannot be used safely. In response, capacity-building activities have tended to focus on solutions in terms of methods, such as having more complex quantitative work, more systematic reviews, or more experiments. These,

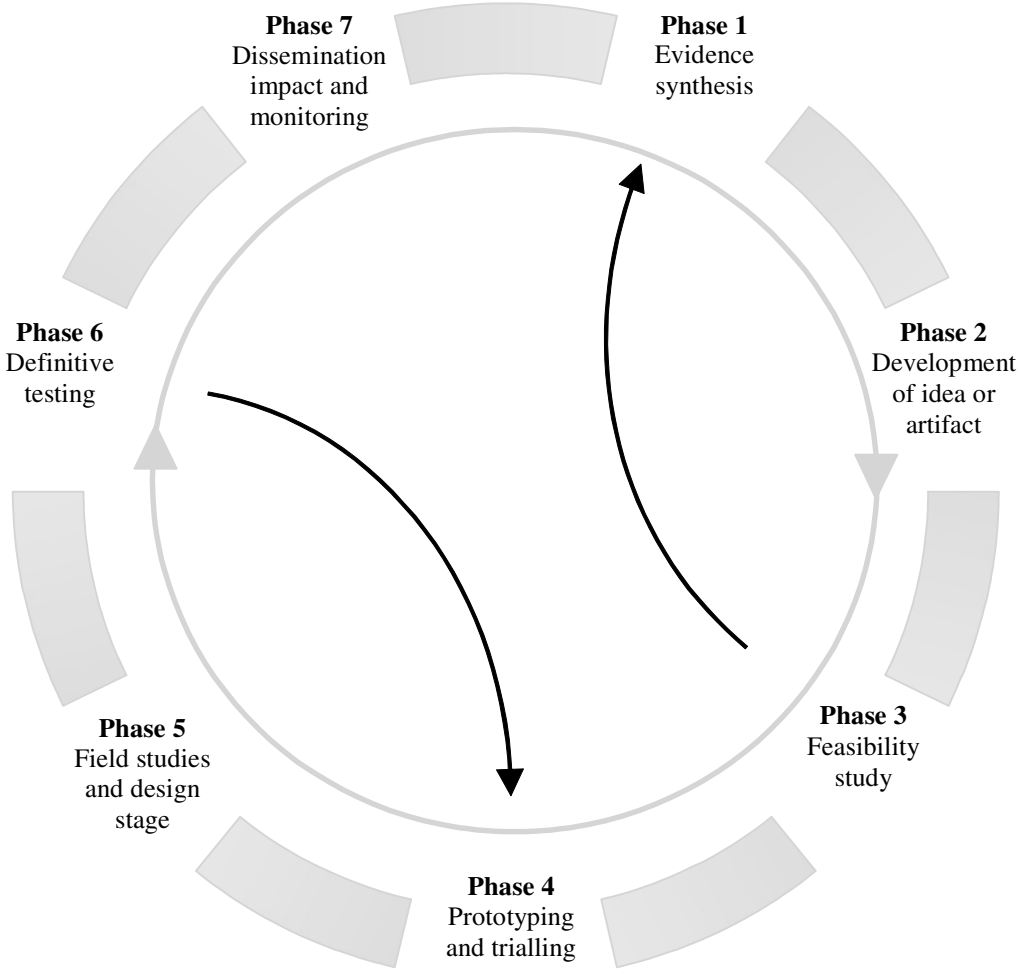
to my mind, are not the answer in themselves. A more general change is needed in the culture of research. The answer for me lies in genuine curiosity, coupled with outright scepticism. These characteristics lead a researcher to suit methods to purpose, try different approaches, replicate and triangulate, and attempt to falsify their findings. It leads them to consider carefully the logic and hidden assumptions on the path from evidence to conclusions, automatically generating caveats and multiple plausible interpretations from the standard query – "if my conclusions are actually incorrect, then how else could I explain what I have found?". Some improvement may come from researcher development, but, somewhat pessimistically for an educator, I have come to believe that the role of capacity-building is limited here. Some people appear genuinely curious and sceptical anyway. Some, on the other hand, tend to be devoted "believers" of things, and their development may involve simply a change of the subject of those beliefs as when a committed religious person becomes an enthusiastic Marxist, or when a "qualitative" researcher turns heavily "quantitative" (Gorard, 2002c, 2005). In a sense, what we need for evidence-based policy making and practice is more real research, where the researcher is genuinely trying to find something out. From this, all else will likely follow – including more and better experiments for many of the reasons advanced by both authors in this chapter so far.

Agreements and disagreements

Intriguingly, having written out our opening positions independently, it seems that we are mostly in agreement, though there are differences of emphasis we will mention. We agree that all commonly used methods have a valid purpose and a place in the larger cycle of education research. Our capacity-building should, therefore, focus on filling in the existing gaps within the cycle so as to create the needed expertise and practices, on trying to overcome mono-method identities where researchers reject the use of all but one type of evidence, and on teaching respect for all methods in their place, as difficult as it is to identify these places.

We also agree that the full research cycle represented in Figure 2.1 presents a simplified and stylised, but useful, model of the research cycle. In this cycle, reviews and secondary analyses might appear in Phase 1, theory-building and small-scale fieldwork in Phase 2, *et cetera*, with smaller experiments being part of Phase 5 and a full randomised controlled trial only appearing once, in Phase 6. We agree that experimental designs are not privileged for all of these phases and that other means are preferable, especially for the first four phases. We also agree that experiments are currently lacking in education research practice writ large, and that most education research gets stuck in Phases 1 to 4. In other words, it is stuck working towards a randomised trial that hardly ever gets done.

Figure 2.1. An outline of the full cycle of education research



Source: Gorard and Taylor (2004).

We further agree that it is important to answer descriptive questions such as "Who gets what?" or "How are teachers trained?". But these questions are no sooner broached than we usually also want to learn how to improve things in these domains and causal questions then arise, like: "How can we train better teachers?" or "How can we better share out resources?" Thus, a complete programme of education research will generally lead to a need to make causal claims, and so to an ethical need for researchers to use something like a randomised controlled trial to make these claims responsibly.

Important consequences follow from our agreement that most education research gets stuck in Phases 1 to 4 and that experiments have a special role to play in the underrepresented Phases 5 through 7. For a fixed research budget, doing more experiments in the later phases will entail fewer resources for those researchers working on Phases 1 through 4, this being the vast majority of education researchers. So these individuals will not, and do not, like increasing the priority accorded to causal questions and methods. This priority is deeply threatening to them intellectually and instrumentally, hence their lack of support for the call to conduct more school-based experiments.

Drawing attention to the neglected later phases of the research cycle indirectly serves to raise the priority accorded to them. After all, there is little point to a model that rarely meets its ultimate goals! Without explicit or implicit priorities, Figure 2.1 is conservative in its implications. It is a recipe for more of the same since so few education researchers want to work on the later phases, or even know how to do so if experiments are required. They might want to argue that Phases 1 through 4 are necessary for the later phases, thus justifying much more work on the earlier than the later phases, especially since the figure presumes a winnowing process – only some modest fraction of the ideas initially generated ever get to have a randomised experiment devoted to them later. However, we both agree that the early phases are not necessary conditions for the later ones, as advantageous as it is to have them. Indeed, many educational practices that are currently widespread have never been through even the first four Phases of Figure 2.1. They are widely implemented despite theory that is weak or even non-existent and, if any studies support these practices at all, they are not strong in terms of internal or external validity, having mostly been conducted in contrived settings or tested in a few schools and with few classrooms or children. In the past, we have been accepting of educational reforms that have hardly benefited from Phases 1 through 4, let alone 5 and 6. Even in logic, there is no need for potential school reforms to have gone through a multi-year testing process before being implemented in schools.

Also pushing towards conservatism is that an un-prioritised Figure 2.1 leaves the funders of education research with total freedom of action. They never need take stands about priorities, and so they need not fear alienating their constituencies in universities and ministries. In many policy environments, setting priorities is a political headache one would like to avoid if possible. Figure 2.1 may be a good normative description of some Platonic research cycle, but it will only change education research practice if it is linked to acknowledging two things we both agree on concerning its last phases – that they are: (1) indispensable to evidence-based policy research since much of policy is about improving educational performance; and (2) they are neglected in current education research practice, making secure knowledge about what works in education a current gap of some significance.

Where we may differ more is on the urgency of the need to fill this gap and hence on the extent to which experiments are needed. Tom Cook is more worried that current education research rarely gets to a point where it reliably tests its ideas in the hurly-burly of school life, and that so few organisations responsible for education and research on education are fazed by this. He believes that those commissioning education research have a responsibility for hurrying along the research cycle and for short-circuiting it on a regular basis by jumping quickly to Phases 5 and 6. He argues that the last phases in Figure 2.1 are the *sine qua non* of evidence-based education research. Without them, policy makers do not have secure causal evidence, arguably the most relevant of all kinds of evidence for forming policy. Consequently, policy makers cannot truly meet their accountability obligation to tax-payers. Of course, there are always researchers willing to offer policy makers causal knowledge; but without experiments they cannot offer causal knowledge that is known to be secure because it results from a valid statistical theory based on random assignment and from the wisdom about implementing experiments that has accumulated from doing them in complex settings in the past, including even from randomised experiments on doing randomised experiments (*e.g.*, Shadish, Luellen and Clark, 2007).

Stephen Gorard sees the need for more causal studies at the end of the research cycle in Figure 2.1, and also the need for more experiments in Phase 6. Indeed, he has

supported both as Director of the ESRC-funded Research Capacity-building Network in the United Kingdom. This helped convince him of the difficulty of shifting the culture in UK higher education research, though he nevertheless continues to take on the task and is currently leading an ESRC-funded Researcher Development Initiative designed to promote the use and understanding of randomised controlled trials (*http://trials-pp.co.uk/*). However, he is less worried about the shortage of knowledge about effective educational practices than Tom Cook is; and he is also less sure of the size of the premium that experiments deserve when causal knowledge is needed. So he does not use the rhetoric of crisis and, if we were to re-assign some hypothetical education research budget, he might not assign as much money to experiments as Tom Cook would. However, this is a difference of degree rather than a fundamental difference about the relative importance of causal questions and experimental methods.

However, we do disagree on whether calling for more genuinely mixed methods is "anodyne", as Tom Cook terms it. Stephen sees the dominance of qualitative studies in UK education journals and regrets the number of researchers who fail to accept the principle that different kinds of questions (phases) require different (multiple) approaches. Tom Cook sees different kinds of questions requiring different methods, but not each kind of question requiring multiple methods. For a given kind of question, one method is often superior to another. It is only across all of education research with its many different kinds of question that multiple methods are needed. And we both agree on this last proposition. However, Tom Cook sees it as so obvious that it is not worth claiming as a great intellectual principle. In this sense, it is anodyne for him, however gripping the need for mixed methods may be as part of a political battle between research factions that struggle to be at the table for prestige, funds and self-vindication. But the main point is that we both agree that randomised controlled trials are the best available primary method for answering causal questions. We both want to know, therefore: How can we get more of them done?

References

Angrist, J.D., G.W. Imbens and D.B. Rubin (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association,* Vol. 91, pp. 444-472.

Campbell, D.T. and J.C. Stanley (1963), *Experimental and Quasi-experimental Designs for Research*, Rand-McNally, Chicago.

Coleman, J.S. (1966), *Equality of Educational Opportunity*.

Cook, T.D. (1991), "Clarifying the Warrant for Generalized Causal Inferences in Quasi-experimentation", in M.W. McLaughlin and D.C. Phillips (eds.), *Evaluation and Education: At Quarter-century*, National Society for the Study of Education, Chicago, pp. 115-144.

Cook, T.D. (2002), "Randomized Experiments in Educational Policy Research: A Critical Examination of the Reasons the Educational Evaluation Community has Offered for Not Doing Them", *Educational Evaluation and Policy Analysis*, Vol. 24(3), pp. 175-199.

Cook, T.D. (in press), "Waiting for Life to Arrive": History of the Regression Discontinuity Design in Psychology, Statistics and Economics, *Journal of Econometrics.*

Cook, T.D. and D.T. Campbell (1979), *Quasi-Experimentation: Design and Analysis Issues for Field Settings*, Rand-McNally, Chicago.

Cook, T.D., H. Cooper, D.S. Cordray, H. Hartmann, L.V. Hedges, R.J. Light, T.A. Louis and F. Mosteller (eds.). (1992), *Meta-analysis for Explanation: A Casebook*, Russell Sage Foundation, New York.

Cook, T.D. and D. Foray (in press), "Building the Capacity to Experiment in Schools: A Case Study of the Institute of Educational Sciences in the U.S. Department of Education", *Economics of Innovation and New Technology*.

Cook, T.D., W.J. Shadish and V.C. Wong (2007), "When Non-experimental and Experimental Effect Size Estimates do and do not differ: A Review of the Within-Study Comparisons Literature", Institute for Policy Research, Northwestern University, Evanston, Ill.

Cook, T.D. and V.C. Wong (in press), "Empirical Tests of the Validity of the Regression Discontinuity Design", *Annales d'économie et de statistique*.

Cronbach, L.J. (1982), *Designing Evaluations of Educational and Social Programs*, Jossey-Bass, San Francisco.

Ehri, L., S. Nunes, S. Stahl and D. Willows (2001a), "Systematic Phonics Instruction Helps Children Learn to Read: Evidence from the National Reading Panel's Meta-analysis", *Review of Educational Research*, Vol. 3, pp. 393–447.

Ehri, L., S. Nunes, D. Willows, B. Schuster, Z. Yaghoub-Zadeh and T. Shanahan (2001b), "Phonemic Awareness Instruction Helps Children Learn to Read: Evidence from the National Reading Panel's Meta-analysis", *Reading Research Quarterly*, Vol. 36, pp. 250–287.

Glazerman, S., D.M. Levy and D. Myers (2003), "Nonexperimental versus Experimental Estimates of Earnings Impacts", *The Annals of the American Academy,* Vol. 589, pp. 63-93.

Gorard, S. (2001), *Quantitative Methods in Educational Research: The Role of Numbers Made Easy*, Continuum, London.

Gorard, S. (2002a), "Ethics and Equity: Pursuing the Perspective of Non-participants", *Social Research Update*, Vol. 39, pp. 1-4.

Gorard, S. (2002b), "The Role of Causal Models in Education as a Social Science", *Evaluation and Research in Education*, Vol. 16(1), pp. 51-65.

Gorard, S. (2002c), "Fostering Scepticism: The Importance of Warranting Claims", *Evaluation and Research in Education*, Vol. 16(3), pp. 136-149.

Gorard, S. (2003a), *Quantitative Methods in Social Science: The Role of Numbers Made Easy*, Continuum, London.

Gorard, S. (2003b), "Understanding Probabilities and Re-considering Traditional Research Methods Training", *Sociological Research Online*, Vol. 8(1), p. 12.

Gorard, S. (2004a), "Scepticism or Clericalism? Theory as a Barrier to Combining Methods", *Journal of Educational Enquiry*, Vol. 5(1), pp. 1-21.

Gorard, S. (2004b), "Three Abuses of 'Theory': An Engagement with Nash", *Journal of Educational Enquiry*, Vol. 5(2), pp. 19-29.

Gorard, S. (2005), "Current Contexts for Research in Educational Leadership and Management", *Educational Management Administration and Leadership*, Vol. 33(2), pp. 155-164.

Gorard, S. (2006a), *Using Everyday Numbers Effectively in Research*, Continuum, London.

Gorard, S. (2006b), "Towards a Judgement-based Statistical Analysis", *British Journal of Sociology of Education*, Vol. 27(1), pp. 67-80.

Gorard, S. and J. Fitz (2006), "What Counts as Evidence in the School Choice Debate?", *British Educational Research Journal*, Vol. 32(6), pp. 797-816.

Gorard, S., K. Rushforth and C. Taylor (2004), "Is there a Shortage of Quantitative Work in Education Research?", *Oxford Review of Education*, Vol. 30(3), pp. 371-395.

Gorard, S., with C. Taylor (2004), *Combining Methods in Educational and Social Research*, Open University Press, London.

Hahn, J., P. Todd and W. VanderKlaauw (2001), "Identification and Estimation of Treatment Effects with a Regression-discontinuity Design", *Econometrica*, Vol. 69(1), pp. 201-209.

Heckman, J.J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica*, Vol. 47, pp. 153-161.

Jencks, C. (1972), in F. Mosteller and D.P. Moynihan (eds.), *On Equality of Educational Opportunity*, Random House, New York.

LaLonde, R. (1986), "Evaluating the Econometric Evaluations of Training with Experimental Data", *The American Economic Review*, Vol. 76(4), pp. 604-620.

Ramey, C.T. and F.A. Campbell (1991), "Poverty, Early Childhood Education, and Academic Competence: The Abecedarian Experiment", in A.C. Huston (ed.), *Children in Poverty: Child Development and Public Policy*, Cambridge University Press, New York, pp. 190-221.

Reynolds, A.J. and J.A. Temple (1995), "Quasi-experimental Estimates of the Effects of a Preschool Intervention: Psychometric and Econometric Comparisons", *Evaluation Review*, Vol. 19, pp. 347-373.

Reynolds, A.J., J.A. Temple, D.L. Tobertson and E.A. Mann (2001), "Long-term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile Arrest", *Journal of the American Medical Association*, Vol. 285(18), 2339.

Rosenbaum, P. and D.B. Rubin (1984), "Reducing Bias is Observational Studies Using Subclassification on the Propensity Score", *Journal of the American Statistical Association*, Vol. 79, pp. 516-524.

Rubin, D.B. (1977), "Assignment to Treatment Group on the Basis of a Covariate", *Journal of Educational Statistics*, Vol. 2(1), pp. 1-26.

Schweinhart, L.J., H.V. Barnes and D.P. Weikart (1993), *Significant Benefits: The High/Scope Perry Preschool Study through Age 27*, High/Scope Foundation, High/Scope Press, Ypsilanti, MI.

Shadish, W.R., T.D. Cook and D.T. Campbell (2002), *Experimental Quasi-Experimental Designs for Generalized Causal Inference*, Houghton Mifflin Company, Boston.

Shadish, W.R., J.K. Luellen and M.H. Clark (2007), "Propensity Scores and Quasi-experiments: A Testimony to the Practical Side of Less Sechrest", in R. Boootzin (ed.), *Measurement, Methods and Evaluation*, American Psychological Association Press, Washington DC.

Trochim, W.M.K. (1984), *Research Design for Evaluation*, Sage Publications, Beverly Hills, Ca.

Part Two
Mediating the Research/policy Interface: The Role of Brokerage Agencies

Chapter 3
What Works Clearinghouse, United States[1]

Robert Boruch, University of Pennsylvania
and Rebecca Herman, American Institute for Research

In this chapter, we outline the main features of the What Works Clearinghouse (WWC). The WWC was designed by the Institute of Education Sciences (United States) to provide educators, policy makers, researchers, and the public with a central and trusted source of scientific evidence on what works in education.

In the United States whenever a science has made remarkable advances, the government has formed new organisations to recognise, foster, and support the science. The creation of the Department of Agriculture, National Institutes of Health, the National Science Foundation, and the National Aeronautics and Space Administration are cases in point. The Institute of Education Sciences (IES), created under the Education Sciences Reform Act of 2002, is a new case in point. Its promise is as substantial as that of its older siblings.

The What Works Clearinghouse (WWC) is an IES initiative. The WWC was designed by IES to provide educators, policy makers, researchers, and the public with a central and trusted source of scientific evidence on what works in education.

WWC is not designed to endorse particular interventions. Rather its focus is on reviewing and summarising the evidence pertaining to the effects of educational interventions, notably evidence that permits causal inferences. Nor does the WWC conduct randomised field trials or quasi-experiments to estimate the effects of interventions. Rather, part of the mission is to assure that all reports on such studies in a WWC topic area are identified and screened for dependability of the evidence.

In what follows, we outline the main features of the What Works Clearinghouse as of 2006. The effort is evolving. Readers are encouraged to consult the WWC website – *http://whatworks.ed.gov* – for up-to-date information.

[1] The What Works Clearinghouse is funded (2001-2006), through a contract from the US Department of Education Institute of Education Sciences to the Campbell Collaboration and the American Institutes for Research, a Joint Venture. This report is about the *facts* on the Clearinghouse. The personal views expressed in this paper do not necessarily agree with the views of the US Department of Education, nor do they necessarily disagree.

The What Works Clearinghouse and embodiments of science

The WWC embodies science in at least three ways. The first concerns the WWC's attention to unbiased estimation of an intervention's effect. As a practical matter, this means the WWC puts randomised controlled trials at a high priority, a status these studies have had in medicine since the 1950s, and in employment, training, and welfare research since the 1970s.

Randomised trials produce fair comparisons because, at the outset, the children, or families, or schools, etc., who are involved in one intervention do not differ systematically from those that are involved in another intervention that is purported to be more effective.

The WWC's focus on unbiased estimates based on randomised trials does not preclude estimates based on quasi-experiments. But the WWC recognises that the results of quasi-experiments are frequently more equivocal than those based on randomised trials because sources of bias in the latter cannot always be identified, much less estimated (Duncan, Magnuson and Ludwig, 2004; Boruch, 1997). The WWC's emphasis on randomised trials accords with the IES emphasis on higher quality evidence about what works, especially randomised trials (US Department of Education, 2003a, 2003b).

The second embodiment lies in science's emphasis on cumulation of knowledge. As a practical matter, the WWC depends on state-of-the-art methods developed over the past 20 years in the science of systematic reviews. The WWC's aims are to search literatures so as to produce an unbiased assembly of studies, screen them on the basis of the trustworthiness of the scientific evidence they have produced, and analyse and synthesise the information so as to properly understand and communicate the results (see, for example, Lipsey and Wilson, 2001).

For education research in the United States, the WWC's approach to instantiating reliance on scientific evidence is unique. There is similar interest in other countries, of course. For instance, OECD reviews of education research in Mexico and the United Kingdom point out the value of scientific research as a basis for informing policy and practice. The World Bank's Operations Evaluation Department Biennial Conference in 2003 focused substantially and for the first time on randomised trials in education and other sectors.

Both the Cochrane Collaboration in health care (*http://www.cochrane.org*) and the Campbell Collaboration in the social, criminological, and education sectors are international (*http://www.campbellcollaboration.org*). Their cross-discipline efforts aim to advance higher standards of evidence in the review – and ultimately the production – of studies. The WWC has built on these international initiatives, and expects that these other initiatives will capitalise on the WWC's work. The WWC also builds on earlier efforts in the United States that transcended political squabbles and that depended on the interest of teachers, administrators, and researchers in learning what works, notably Herman *et al.* (1999).

The third way that the IES's What Works Clearinghouse embodies scientific standards is through the use of transparent decision rules and protocols, developed under the guidance of substantive and methodological experts. The What Works Clearinghouse's Technical Advisory Group (TAG) contributed to the early development of WWC study review standards, and individual TAG members help resolve technical issues as they arise. The WWC's reliance on independent peer review is basic to vetting

the quality of the reviews that the WWC products. The review production system relies on explicit, consistent protocols, coding guides, and technical guidance, and the work of expert teams, led by principal investigators who are themselves experts in the areas under review.

Assumptions and prospects

The success of the Institute for Education Sciences' What Works Clearinghouse depends on some things that are in the WWC's control and some that are not. The prospects, for instance, depend partly on the public appetite for good evidence on what works. The No Child Left Behind Act attaches high value to scientific evidence. But if public interest in good evidence diminishes, governmental support for producing good evidence might then also decline.

The WWC reviews reports on field studies rather than executing such studies. Consequently, the WWC has no direct control over the production of high-quality research on the effects of interventions, especially randomised trials. If the supply of such studies is cut short, the WWC mission might have to change. The WWC can and does, of course, encourage production of high quality field tests indirectly, partly by recognising the value of randomised trials and what appear to be good quasi-experiments, and by enhancing their visibility in its standards for reviewing the research. Further, the WWC operates a Help Desk to help researchers understand and apply WWC review standards in their own work.

The prospects for success depend heavily on resources, especially people, for the production of reviews of evidence. The intellectual resources include published work on standards of evidence and reporting on individual studies in the health sector, such as the CONSORT statement (Altman *et al.*, 2001), and advances in the social, behavioral, and education sciences that direct special attention to producing fair estimates of an intervention's effect (Boruch, 1997; Mosteller and Boruch, 2002; Sherman, 2003).

The intellectual resources include procedures, methodological advances in conducting meta-analyses and systematic reviews of impact evaluations, and standards that have been developed for assessing assemblies of studies and reporting systematic reviews of studies in health care (Moher *et al.*, 1999) and in the social, behavioral, and educational sectors (Cooper, 1998; Halvorsen, 1994). They also build on precedents such as Herman *et al.* (1999) in education and Chalmers (2003) in health care, among others.

Operating principles

Assuring the quality of evidence is the first of the WWC's operating principles, represented partly in the WWC's focus on scientific excellence. The first principle is embodied in the standards developed for assessing evidence that are posted on the WWC's website. A second operating principle requires the WWC to be procedurally and organisationally efficient. Identifying dependable studies from the morass is demanding and complicated; the task requires efficiency to serve the public interest. Because the WWC is exploring new terrain, a willingness and capacity to improve is a third operating principle. Technical issues, for instance, emerge often, and technical guidance documents are developed on a "case law" basis to facilitate reviews in particular domains of education research. Emphasising accessibility and transparency in organisation and

procedures, in identifying and explaining the evidential standards, and in efforts to improve constitutes a fourth operating principle under the contract.

The WWC's credibility depends on these basic operating principles, of course. But as an ancient Latin aphorism puts it, being virginal is not sufficient. One must also appear virginal. Independence in the sense of anonymous and independent peer review, for example, is a theme that is instantiated in the WWC operations. Science asks to be surpassed and outdated. Consequently, the WWC is attentive to the need for course correction as the knowledge base changes. Course corrections depend on everyone who contributes to WWC, include people in the IES, sibling organisations such as the Cochrane Collaboration in health and the Campbell Collaboration in the social sectors, and others who contribute to the effort. Some corrections depend on the critics of WWC's products, and critics are an important resource.

Contemporary history

The WWC's aims and operating principles, described above, were made explicit in a competitive contract that the IES awarded in 2001 to a joint venture of Campbell Collaboration (C2) and the American Institutes for Research (AIR).

During 2001-2003 in a process of incremental and demanding improvement, the WWC developed tools and standards for assessing quality of evidence. During 2002, the WWC's Technical Advisory Group (TAG) was assembled. The prospective members' knowledgeability about scientific evidence, including randomised trials and measurement, and the production of systematic reviews of evidence were crucial to their invitation to serve.

During 2004, the WWC undertook a pilot phase to test the application of WWC standards in the review area of Middle-School Math Curricula. While the pilot test affirmed the use of WWC standards in reviews, it did reveal major challenges in designing detailed reporting formats that would give WWC users, including practitioners and researchers, what they need to know about each study. The WWC website and WWC reports underwent at least three major changes and many smaller modifications to shape the WWC's presentation of review results.

During 2004-2006, the volume of production of reviews increased from one to seven topics: early childhood education, beginning reading, elementary school mathematics, dropout prevention, English language learners, character education, and updated reviews on middle school mathematics. All of these focused on named interventions – including programmes and practices – and were based on reviews of randomised trials and quasi-experimental designs that met WWC standards.

The WWC'S products

The WWC's reviews of evidence on education interventions, at two levels of reporting, are the WWC's most important products. The WWC's standards of evidence are a deeper level of product. They underpin all the WWC work. The WWC's Evaluator Register, another product, was designed to assure that capacity for generating higher quality evidence can be fostered and exploited well. The use of the reviews by policy makers, researchers, and practitioners is itself an important ultimate product of the effort.

Standards of evidence as a WWC product

A major theme underlying all standards enunciated by the What Works Clearinghouse is that one must be able to make causal inferences about what works and what does not work based on dependable evidence.

Operationally, this means that randomised trials get top priority. They are more dependable in making a causal inference about what works than quasi-experiments. This also means that quasi-experiments have a lower priority, and are designated as meeting a lower standard of scientific evidence in any reports produced by the WWC. Randomised trials with no serious problems in their design or execution are rated as by WWC "Meets Evidence Standards". Quasi-experiments that (1) match on a pretest (or a good proxy) and other appropriate matching variables or (2) covary on these measures are rated as "Meets Evidence Standards with Reservations". The phrase "with reservations" is intended to remind readers that a quasi-experiment cannot provide the assurance of unbiased estimates of difference that a randomised trial can other things being equal. The WWC is also exploring standards for dependability of regression discontinuity studies (which is a quasi-experimental design with especially strong causal validity) and single subject designs.

Beyond the broad rating, WWC reviewers also examine and describe certain features of studies to assure that the studies can be interpreted properly and reviewed accurately and uniformly. These features include descriptions of the intervention, outcome measures, study settings, subgroups tested, and analysis statistics. WWC, for example, encounters reports at times that do not contain basic statistical information such as variance within groups being compared. A study that does not provide enough information to compute – and verify – study authors' reported findings would be screened out. The WWC uses a uniform query to request the missing information from study authors in such cases so as to assure reviewers have all pertinent information.

The WWC's efforts to develop standards must confront the fact that we do not know the answers to some questions, and that we must be attentive to the accretion of empirical evidence that could help address such questions. Consider, for instance, a randomised trial in which children or families attrite from one arm of the trial at a 5% rate and in the second arm at a 20% rate. Is this potentially serious difference important enough to incorporate into a standard that directs attention to internal validity of a trial? Does it depend on a recruitment process and context? How do we take into account the continuously accumulating evidence on attrition rates from well-conducted trials, and then make judgments about the dependability of the evidence at hand? And how do we incorporate this into a standard? WWC is working on such issues and how to take new evidence into account.

The WWC standards underwent repeated scrutiny and modification during 2002-2005, based on the Technical Advisory Group, public comments, and comparisons to related standards in the medical arena. The earliest versions were eventually put aside because of complexity in presentation; many seasoned researchers could not understand them. The more transparent and up-to-date standards are given on the WWC website. The WWC also develops technical guidance to provide more detailed decision rules for operationalising the standards. For example, the WWC standards indicate that severe attrition is problematic. The related technical guidance explains what should be considered "attrition" and the levels at which attrition is problematic. The standards and technical guidance are periodically updated on the WWC website. The WWC is developing an archive of technical issues confronted in WWC reviews, their resolution,

and application of the resolution in WWC review standards. Readers are encouraged to see the site for the most recent version.

WWC Evaluator Register

In 2005, the WWC launched an Evaluator Registry that provides information about organisations and individuals that have the capacity to produce high-quality evidence on the effects of educational interventions. Entries to the register are based on registration by evaluators who provide information on their performance – for instance, in designing and executing trials and in having the products of their research and evaluations published in peer-reviewed scientific venues.

The intended consumers and their use of WWC products

The WWC aims to assure that its products are used by policy makers, practitioners, researchers, and others. The WWC understands that getting research used is no easy task. In the medical research arena, for instance, it takes 5 to 10 years for a tested innovation to be incorporated into practice. In the education arena, the results of Tennessee's class size trials were not recognised, much less used, by many policy people for over 5 years. The WWC would like to foster a brisker pace.

Because WWC depends on advances in the state of the art in conducting studies, and advances the state of the art in reviewing them, researchers are part of the target for WWC reviews. The WWC aims to vet ideas and products in peer-reviewed scientific forums. Consequently, papers covering some WWC activities have been developed for peer-reviewed journals such as the *Annals of the American Academy of Political and Social Sciences* (Turner *et al.*, 2003) and edited books.

The public and professional media are important, given the WWC's interest in assuring that teachers, parents, and policy makers can learn about and use the WWC's products. Media related information has been put up on the WWC's website. Such information and a broader communications strategy has led to new WWC reviews being covered frequently in the popular press and in trade journals such as *Education Week*.

Attracting attention to websites and assuring repeat visits can be a fiercely competitive enterprise. The WWC's website has undergone at least three major changes in the years since its creation, and WWC continues in its effort to improve. Nonetheless, one must confront the fact that there are hundreds, if not thousands, of websites that purport to tell "what works" on topics ranging from astrology to zoo keeping, and that the phrase is also common in sites that purport to provide evidence about education practice and policy. Despite the competition, the WWC website has substantial usage, with an average of over 1 300 unique visitors per day.

The WWC topics and workflow

The WWC aims to be as attentive to quality and as transparent as possible. Most important, the workflow includes quality control at repeated definable points.

At the first stage of the WWC's workflow, people submit their opinions about what topics, interventions, or studies ought to be reviewed by the WWC. The people who make submissions can include anybody – parents, teachers, executives in publishing houses,

researchers, or other individual or organisations who have an interest in discerning what works or who might benefit or suffer from a WWC review on what works. Candidate topics also are nominated in professorial forums to which WWC contributes. Certainly they also include advisors to the IES, including substantive area specialists.

The WWC's choice of a particular topic for review depends on (a) the relevance of the topic to current education policy and practice, (b) the topic's probable importance in decisions about what interventions can be adopted, and (c) the level of evidence available. These are complex interrelated criteria. Reaching decisions has involved assuring that different prospective users of information weigh in on the information they want: policy makers, practitioners, and researchers. As of 2006, the topics for review include Elementary-School Math, Middle-School Math, Dropout Prevention, Character Education, Beginning Reading, English Language Learning, and Early Childhood Education. Each topic has a review team consisting of a PI, project coordinator, and coders.

A WWC review in a topic area begins with detailed protocol, developed by the PI, that defines the intervention and inclusionary criteria, the target population including high-risk subpopulations, the outcome variables that are pertinent, and the study designs that are eligible for a WWC review under WWC standards.

The WWC's process for generating a review in a particular intervention area continues with comprehensive literature searches and full-text readings of published and unpublished reports. Outcome studies that depended solely on testimonials or simple correlations are eliminated at the outset, for example. Randomised trials and high-end quasi-experiments on relevant interventions were admitted to candidacy for WWC review.

When eligible studies are identified, the coding process begins with basic categorical distinction between randomised trials and quasi-experimental designs. For each category of study, characteristics that influence internal validity are identified. For instance, a randomised trial that has large difference in the attrition rate between intervention arms must be recognised. As a result, it might subsequently be downgraded to quasi-experimental status "Meets Standards with Reservations", in the absence of other information that speaks to the biases that such attrition could engender.

Characteristics of studies are double coded by two independent coders to assure that coding reliability can be estimated. Differences of opinion in coding are adjudicated by a principal investigator and a project coordinator. Principal investigators provide substantive expertise to professional review teams and weigh in on topic-specific decisions. Some people might expect that adjudication issues are few and take little time. That has not been the case. Adjudicating ambiguities in a report from a peer-reviewed journal can easily take hours. Because standards of reporting research in journals have changed, and because the WWC may cover up to 20 years of preceding research in a review area, the number of adjudicated cases can be large.

Draft Intervention Reports and Topic Level Reports are reviewed by members of the WWC Technical Review Team, anonymous peer reviewers who are engaged by the IES directly, and by senior IES staff. The aims of these external peer examinations are to assure accuracy in the WWC reports, to minimise ambiguity, and to verify uniform adherence to WWC standards.

Concluding remarks

The Institute of Education Sciences' What Works Clearinghouse (WWC) is unprecedented in its focus on the quality of evidence that is generated about the effects of education interventions and its focus on scientific standards in making judgments about evidence quality. It is also unprecedented, in education, for operationalising standards that are as public and transparent as possible, across a wide variety of topics. The WWC is unprecedented in creating an organisation, processes and procedures, and teams of people that are essential in developing reviews at this scale and with this level of transparency.

Despite lack of these precedents, the IES's Clearinghouse has depended heavily on experience and advances in understanding how to build scientific knowledge. This includes work over the last three decades on randomised trials so as to produce unbiased estimates of the relative effects of interventions. It includes scientific work over roughly the same period – in health care, criminology, and welfare, as well as education – to understand how to summarise the results of studies uniformly and against clear standards.

The aims are high and the products important. In identifying what works, the Clearinghouse will help us, as a fine aphorism suggests, to "Test all things and hold fast to that which is good."

References

Altman, D.G., K.F. Schulz, D. Moher, M. Egger, F. Davidoff, O. Elbourne *et al.* (2001), "The Revised CONSORT Statement for Reporting Randomized Trials", *Annals of Internal Medicine*, Vol. 134(8), pp. 663-694.

Boruch, R.F. (1997), *Randomized Experiments for Planning and Evaluation: A Practical Guide*, Sage, Thousand Oaks, CA.

Chalmers, I. (2003), "Trying to Do More Good than Harm: The Role of Rigorous, Transparent, Up-to-date Evaluations", *Annals of the American Academy of Political and Social Sciences*, Vol. 589, pp. 22-40.

Cooper, H. (1998), *Synthesizing Research: A Guide for Literature Reviews*, Sage, Thousand Oaks, CA.

Duncan, G.J., K.A. Magnuson and J. Ludwig (2004), "The Endogeneity Problem in Developmental Studies", *Research in Human Development*, Vol. 1(I&2), pp. 59-80.

Halvorsen, K. (1994), "The Reporting Format", in H. Cooper and L. Hedges (Eds.), *Handbook of research synthesis*, Sage, New York, pp. 425-438.

Herman, R., D. Aladjem, P. McMahon, F. Masem, I. Mulligan, A. O'Malley *et al.* (1999), *An Educators' Guide to Schoolwide Reform*, American Institutes for Research, Washington DC.

Lipsey, M.W. and D.R. Wilson (2001), *Practical Meta-analysis*, Sage, Thousand Oaks, CA.

Moher, D., D.J. Cook, S. Eastwood, I. Olkin, D. Rennie and D.T. Sunup (1999), "Improving the Quality of Reports of Meta-analyses of Randomized Controlled Trials: The QUORUM Statement", *Lancet,* Vol. 354(27), pp. 1896-1900.

Mosteller, F. and R. Boruch (2002), *Evidence Matters: Randomized Trials in Educational Research*, Brookings Institution Press, Washington DC.

Sherman, L.W. (ed.) (2003), "Misleading Evidence and Evidence-led Policy: Making Social Science More Experimental", *Annals of the American Academy of Political and Social Science*, Vol. 589(l), pp. 6-19.

Turner, H., R. Boruch, A. Petrosino, J. Lavenberg, D. de Moya and H. Rothstein (2003), "Populating an International Web-based Randomized Trials Register in the Social, Behavioral, Criminological, and Education Sciences", *Annals of the American Academy of Political and Social Sciences*, Vol. 589(l), pp. 203-223.

US Department of Education, Institute of Education Sciences (2003a), *Random Assignment in Program Valuation and Intervention Research: Questions and Answers*, Washington DC.

US Department of Education, Institute of Education Sciences (2003b), *Identify in and Implementing Educational Practices Supported by Rigorous Evidence: A User Friendly Guide*, Washington DC.

Chapter 4
The Evidence for Policy and Practice Information and Co-ordinating (EPPI) Centre, United Kingdom

David Gough, Institute of Education, University of London

In this chapter, we present the Evidence for Policy and Practice Information and Co-ordinating (EPPI) Centre of the University of London, United Kingdom. The Centre aims to develop and promote participatory and user-friendly systematic reviews that address important questions in policy, practice and research in the public interest.

Aims and function

The Evidence for Policy and Practice Information and Co-ordinating (EPPI) Centre is part of the Social Science Research Unit at the Institute of Education, University of London (*http://eppi.ioe.ac.uk/*). Its work on systematic research synthesis for evidence informed policy and practice started in 1993 with the aim of developing and promoting participatory and user-friendly systematic reviews addressing important questions in policy, practice and research in the public interest (Oakley *et al.*, 2005). It attempts to achieve these aims through a number of interrelated strategies.

The first strategy is to develop a broad conceptual framework for understanding, undertaking and using question-driven reviews. This includes examining the origin and nature of the questions being asked, the answers produced, and the relationship with policy, practice and individual decision-making. Second, by developing methods and tools for systematic reviews answering all types of research questions and including all types of research evidence. Third, undertaking reviews in-house and supporting others to undertake reviews and adding to the evidence base in different discipline areas. Fourth, providing support and training to develop capacity in evidence-informed policy and practice. Fifth, working with others nationally and internationally to achieve these aims and products.

Methods

A broad conceptual framework

To be systematic

Any individual research study is fallible and more reliance can usually be put on the full corpus of research relevant to a research question. Traditionally reviews of literature have not been explicit about their methods and so it is not evident why a review was taken in a particular way, why it included some but not all of the potentially relevant research literature and why it came to the conclusions that it did. In contrast, systematic reviews aim to meet the standards of primary research of being explicit about their methods so that the results are accountable; so that one can assess the appropriateness of the methods used and be convinced that the findings were not subject to some form of hidden bias. Reviews are also important to ensure that research is available to non researchers so that decisions can be informed by research as well as values, resources, and practice knowledge (Hargreaves, 1996; Hillage *et al.*, 1998).

Users question driven and interpreted and implemented reviews

The EPPI-Centre's interest in evidence-informed policy and practice is based on the use of systematic reviews to make a difference by answering the questions of policy makers, practitioners, users of services and other members of society. Systematic reviews ask what do we know from research in relation to different questions. Different individuals and groups will have different concerns and different questions, and this should lead to a range of different user, question-driven reviews. A framework for different types of review thus needs to take into account that there will be a plurality of reviews being used in different ways by different individuals and groups.

Similarly, the findings of research usually have little meaning without separate processes of interpretation and implementation, both of which involve users of research engaging with other types of knowledge. In the same way that reviews have formal systematic procedures to ensure accountability, different but equally important sets of procedures are needed for the two processes of interpretation and implementation of reviews. Such developments in the use of evidence for policy and practice are manifested in brokerage agencies such as the Canadian Council on Learning, and the English National Institute for Health and Clinical Evidence and Social Care Institute of Excellence (see other chapters in this volume).

A framework for all systematic reviews

The systematic approach to reviewing literature has become routine in the review of quantitative experimental studies through statistical meta-analysis in health research, but is less common in other disciplines or in addressing other sorts of questions. In parallel and in cooperation with many colleagues across the world, the EPPI-Centre develops methods of systematic review that apply to all research questions and thus can include all types of research data including both quantitative and qualitative data and synthesis.

The range of current systematic reviews is increasing rapidly and reviews now vary on many dimensions such as (Gough, 2006):

- the question being asked. For example, exploring or generating theories of cause or testing the efficacy of interventions;
- whether the method of review is specified a priori or develops iteratively during the progress of the review. For example, the iterative approaches used in some forms of meta-ethnography (Noblitt and Hare, 1988), critical interpretative synthesis (Dixon-Woods *et al.*, 2006), realist synthesis (Pawson, 2006) and meta-narrative reviews (Greenhalgh *et al.*, 2005);
- whether the literature is searched exhaustively, is sampled in a purposive way and/or until a sufficient amount of references has been found (saturation);
- whether empirical or conceptual data is being considered;
- whether numerical or narrative data is used as evidence for the review and in the analysis and process of synthesis of the review;
- whether the synthesis is predominantly meta-empirical as an integration of "facts" within an accepted theory or world view or meta-conceptual as an integration of different conceptual views.

As the Methods for Research Synthesis Node of the ESRC National Centre for Research Methods, the EPPI-Centre is creating a matrix of all the research questions used in the social sciences and the actual or potential methods of research used to answer these questions. It is too soon to provide details of the matrix but it is likely to include two dimensions of the research activity (such as describe, measure, compare, relate and evaluate) and the use of theory (such as generate, explore or test theory). We are then applying this matrix to examine: (i) all the actual and potential questions that could be asked by systematic reviews; (ii) the actual or potential methods of review; and (iii) the conceptual and practical challenges that these methods involve. In doing so, we aim to provide an overall framework to understand the range and nature of systematic reviews, to chart their development and to assist further methods development in the future.

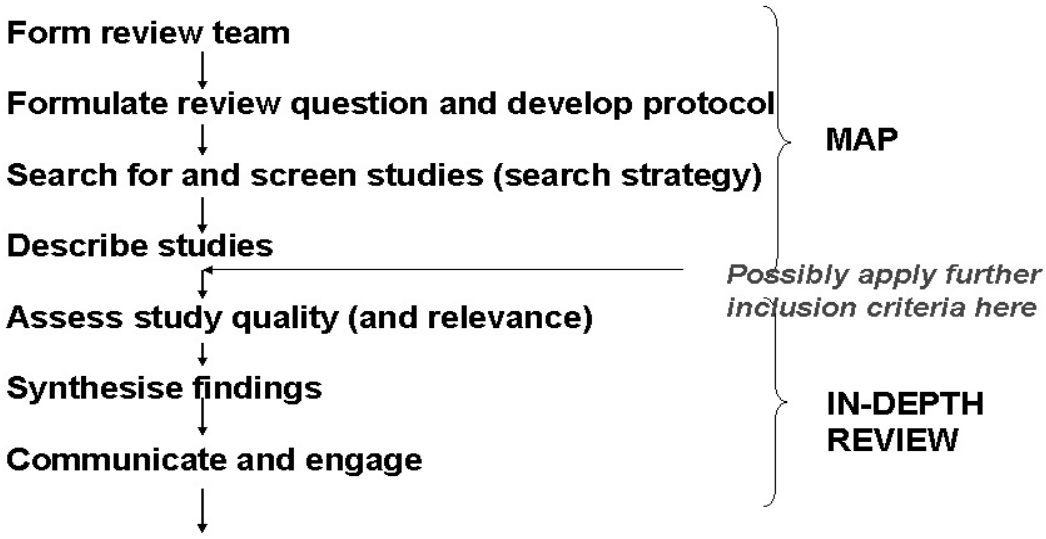
In addition, the framework has to take account of the fact that not all reviews are of the same size and scope. They vary in terms of the breadth of the issues considered, the depth to which they are examined, and the time and financial resources invested to achieve these aims (Gough, 2006).

Methods and tools

As review questions and methods can vary so extensively (as shown by the matrix of all types of review) the EPPI-Centre develops procedures that can be used for many types of reviews.

One example of such a procedure is systematic mapping of research that describes the nature of the research that has been undertaken (Peersman, 1996). This is a description of all the research identified by the systematic review as relevant to answering the review question (see Figure 4.1). The map is a useful product in its own right in providing an analysis of research that has been undertaken and also helps inform the nature of the synthesis that could be of all of the map or just part of the map. This ability to narrow down from the map to the in-depth review and synthesis means that the original question can be broader than it might otherwise have been (Gough 2005, in press a).

Figure 4.1. Mapping and synthesis



Another example of flexibility of methods is the development of a process for quality and relevance appraisal of studies. There are many quality appraisal tools available but these typically assess the quality of a study in its own right rather than in terms of what value it brings to answering the review question. The EPPI-Centre's Weight of Evidence system provides a process for distinguishing the generic judgement of quality of execution of a study, from the review specific judgements of appropriateness of the research design for answering the review question and the focus of the study (Gough, in press a). The Weight of Evidence approach does not provide detailed criteria for making these judgements but a system for such judgements to be made and described by the authors of reviews.

A further example is the use of mixed methods reviews, where a review question is addressed by asking subquestions which are addressed by different methods and then compared with each other. In a review on barriers to healthy eating in children and young people, a systematic review of experimental studies of the efficacy of health promotion interventions to increase healthy eating was undertaken in parallel with a conceptual synthesis of research on children and young people's views about health and eating (Harden and Thomas, 2005). The studies of efficacy showed that the health promotion interventions were effective to some degree but may have been much more effective if they had been devised taking into account user views. For example, the synthesis of views studies showed that children considered fruit and vegetables as very different but most health promotion interventions combined messages to eat more fruit and more vegetables. Also, children thought that health was an issue for parents rather than them and so were probably less likely to be convinced by exhortations to eat fruit and vegetables in order to be healthy. Fashion and image might be much more effective health promotion strategies. It would be wrong to assume that health promotion interventions are

not very effective simply because they have been devised without much consideration of the research on the target audience.

The Centre also develops review tools such as EPPI-Reviewer, a web based software system to manage all stages of a review. This includes bibliographic capture of references from electronic bibliographic databases, management of those references and associated electronic and hard copies, screening against review inclusion criteria, data coding for mapping or data extraction, quality assurance, both quantitative statistical meta-analysis and qualitative thematic analysis, and data organisation for reporting of the review. EPPI-Reviewer can be used with different screening, coding and analytic schemes and several in-house guidelines have been developed for EPPI-Centre reviews with generic, discipline specific and review specific coding frameworks (or guidelines). This is accompanied by a review companion to help review authors undertake a review as well as a tool to help in the assessment of the quality of completed reviews. Just as with the development of broad methods for undertaking reviews, the software does not dictate detailed decisions about how a review should be undertaken but provides tools to enable the review process for all types of reviews.

The same enabling approach is taken with the structure of the Centre's technical reports that in being transparent detail all aspects of the methods of a review. The structured approach makes it easy to check how each part of the review has been undertaken. Such detailed reports are not, however, suitable for all audiences so the Centre has developed a four-level communication strategy of a one page summary, a fifteen- to twenty-page main report, a full technical report, and web access to all the data codings (from EPPI-Reviewer) on which the report was based.

Capacity-building

Although the need to review what we know before undertaking new research, policy or practice has been known for a long time, the widespread use of systematic methods of review is quite recent (Chalmers, Hedges and Cooper, 2002). The increased use of systematic review evidence requires a culture change in the use of research and balancing the investment in new primary work and in consolidating what we know. We still need much new creative research but this has to be balanced against the wastage of many under resourced and ineffective studies, the duplication of work already done, and research not focused on the issues of most relevance to decision-making.

One part of the culture change is the capacity in understanding and undertaking reviews. The Centre attempts to assist with capacity-building by developing methods and tools for review, supporting external groups in undertaking reviews and in providing a range of training resources in reviewing. This includes tailored workshops for other organisations, stand alone workshops for individuals, and a full MSc in Evidence for Public Policy and Practice. The workshops contain didactic sessions with discussion and small group work, but e-learning will soon become an important mode for training and review support.

Work with others

None of this work would be possible if the Centre was not supported by many partners and collaborators. The Centre is funded by the university and by many grants from a range of government research councils, charitable foundations and government agencies and departments.

The Centre works with many others who are also facing the same challenges of developing methods and resources for reviews including the external EPPI-Centre groups. For example, the Centre has formal links with the Cochrane Collaboration (*www.cochrane.org*) (the body that coordinates reviews on the efficacy of health interventions) as co-directors of the Cochrane Health Promotion and Public Health Field.

As part of this the Centre promotes the production and use of reviews in health promotion and maintains a web-based register of experimental trials and of systematic reviews. The Centre is also a formal partner of the Campbell Collaboration (*www.campbellcollaboration.org*) which is an international umbrella group to support reviews on social interventions.

Issues

Despite all the international activity to develop methods of reviews there are a number of major challenges faced by those committed to evidence-informed policy and practice.

Firstly, the culture change is still in its infancy and there are many who are unaware of the importance of such an approach. Reviews are not cheap and need resources in order to be carried out just as with primary research. Any major change in funding and support for reviews might quickly reverse the culture change in support for evidence-informed policy and practice of the last few years.

Second, we need to acknowledge that there are those who are sceptical about the value of systematic reviews. Some of these concerns are simply critiques of poor reviews or processes that need to be developed, and this needs to be taken seriously and seen as a resource to drive improvements in reviews by the systematic review community. Some other concerns are due to misunderstandings such as the belief that reviews are only of randomised controlled trials rather than all types of research questions and research data. Others critiques are more fundamental and arise from those with different values or views of science or who have an interest in maintaining the status quo without explicit methods of synthesis of empirical or conceptual knowledge (Oakley, 2006).

Third, in order to achieve such a culture change, reviews need to be shown to be useful, but this is easier to demonstrate with a critical mass of evidence reviews rather than relatively few single reviews. In health, the Cochrane Collaboration has such a body of review evidence which has made a difference to policy and practice but it is still early days for reviews in education and other social sciences.

Fourth, more reviews need to be demand-led so that they are more likely to be of use. Academics are users of research and are well placed to determine the focus of primary research and of reviews. But they may not be so well placed to determine the focus of all research that is relevant to other users of research such as policy makers, practitioners, and members of the public. Involving these others users in driving demand for reviews and thus also for primary research (Gough, in press b) should make research more democratic, more fit for purpose and more demand led.

Fifth, whoever determines the focus of reviews, we need to develop better formal processes for the interpretation and implementation of review findings. Undertaking reviews, however sophisticated, is not going to be sufficient if we do not also have sophistication in other parts of the evidence to decision-making cycle.

References

Chalmers, I, L.V. Hedges and H. Cooper (2002), "A Brief History of Research Synthesis", *Evaluation and the Health Professions*, Vol. 25(1), pp. 12-37.

Dixon-Woods, M., D. Cavers, S. Agarwal, E. Annandale, A. Arthur, J. Harvey, R. Hsu, S. Katbamna, R. Olsen, L.K. Smith, R. Riley and A.J. Sutton (2006), "Conducting a Critical Interpretive Review of the Literature on Access to Healthcare by Vulnerable Groups", *BMC Medical Research Methodology*, Vol. 6, p. 35.

Gough, D. (2004), "Systematic Research Synthesis", in G. Thomas and R. Pring (eds.), *Evidence-based Practice in Education*, Open University Press, Buckingham, pp. 44-64.

Gough, D (2006), *Typology of Evidence Reviews*: *I Main Dimensions; II Breadth, Depth, Quality and Use*, 19th April, EPPI-Centre, London.

Gough, D.A. (in press a), "Weight of Evidence as a Process to Appraise Studies for Systematic Reviews", in *Research Papers in Education*, special edition edited by J. Furlong and A. Oancea on quality in applied and practice-based research.

Gough, D.A. (in press b), "Giving Voice: Evidence Informed Policy and Practice as a Democratizing Process", in M. Reiss *et al* (eds.), *Discourses of Difference within and Beyond Education*, Trentham Books, Stoke.

Greenhalgh, T, G. Robert, F. Macfarlane, P. Bate, O. Kyriakidou and R. Peacock (2005), "Storylines of Research in Diffusion of Innovation: A Meta-narrative Approach to Systematic Review", *Social Science & Medicine,* Vol. 61(2), pp. 417-430.

Harden, A. and J. Thomas (2005), "Methodological Issues in Combining Diverse Study Types in Systematic Reviews", *International Journal of Social Research Methodology,* Vol. 8, No. 3, pp. 257-271.

Hargreaves, D. (1996), "Teaching as a Research-based Profession: Possibilities and Prospects", *Teacher Training Agency Annual Lecture*, TTA, London.

Hillage, J., R. Pearson, A. Anderson and P. Tamkin (1998), *Excellence in Research in Schools*, Department for Education and Employment/Institute of Employment Studies, London.

Noblitt, G. and R. Hare (1988), *Meta-Ethnography: Synthesizing Qualitative Studies*, Sage, London.

Oakley, A. (2006), "Resistances to 'New' Technologies of Evaluation: Education Research in the UK as a Case Study", *Evidence & Policy: A Journal of Research, Debate and Practice*, Vol. 2(1), January, pp. 63-87(25).

Oakley, A., D. Gough, S. Oliver and J. Thomas (2005), "The Politics of Evidence and Methodology: Lessons from the EEI-Centre", *Evidence & Policy: A Journal of Research, Debate and Practice*, Vol. 1, No. 1, pp. 5-31.

Pawson, R. (2006), *Evidence-based Policy: A Realist Perspective*, Sage, London.

Peersman, G. (1996), *A Descriptive Mapping of Health Promotion in Young People,* EPPI-Centre, Social Science Research Unit, Institute of Education, University of London, London.

Chapter 5
The Iterative Best Evidence Synthesis Programme, New Zealand[1]

Adrienne Alton-Lee,[2] Chief Education Adviser, Ministry of Education,
Wellington, New Zealand

In this chapter, we look at New Zealand's Iterative Best Evidence Synthesis Programme, which seeks to develop and use bodies of evidence to explain what works and why in education, with special attention on context.

New Zealand's Iterative Best Evidence Synthesis (BES) Programme is a collaborative knowledge building strategy to develop a series of inter-linked syntheses that explain influences on diverse learner outcomes. Information about the programme can be found at *http://educationcounts.edcentre.govt.nz/goto/BES*. A series of syntheses focused on the major influences on student outcomes (family, teaching, professional development and leadership influences) has been progressively developed as part of medium term strategic policy work. The initial BESs were published in 2003. These informed *Guidelines* (*http://educationcounts.edcentre.govt.nz/goto/BES*) for subsequent BES development. Four more BESs have been developed via collaboration across policy, research and practice with the guidelines as a foundation.

The primary purpose of the programme is to support sustainable educational development whereby a whole education system and its communities strengthen a range of desired outcomes for all learners through iterative processes of shared knowledge building and use. The iterative approach is designed to be a collaborative tool and catalyst to intensify and embed the interplay of research and development (R&D) as a systemic lever for sustainable development in education.

BES has been valued by the New Zealand secondary teachers' union[3] for its challenge to what they call the "snake oil" myths and fads that have beset teachers. The work has been valued for the insights that explain what can make a bigger positive difference and lessen teacher stress. Some examples are: enhanced academic and social outcomes through strengthening student self-regulation, problem solving and conflict resolution

[1] Thanks to Dr Penny Moore whose work on the *Evidence Based Policy Project* informed the latter part of this paper (*http://educationcounts.edcentre.govt.nz/goto/BES*).

[2] Feedback and critique of this paper is welcomed at adrienne.altonlee@minedu.govt.nz

[3] Presentation by the Post Primary Teachers Association to the Minister of Education, Trevor Mallard, at Brackenridge Retreat to discuss implications of the first cohort of BESs, July 16-17, 2003.

skills, intensifying reciprocal peer supports for learning, and optimising school-home-community linkages in ways that dramatically lift the achievement of at-risk students (Alton-Lee, 2003).[4]

The Iterative BES approach to knowledge brokerage

The BES development process requires of BES researcher-writers, iterative engagement with colleagues across educational policy, research and practice. The rationale is that use is embedded in development. A stronger evidence-base and wider ownership and use of synthesis findings are possible when a brokerage role is taken to BES development and use; not only between policy and research communities, but also with educational practice communities.

Fit-for-purpose synthesis methodology

The Iterative BES Programme synthesises bodies of educational research that provide credible evidence about influences on a range of desired outcomes for diverse learners (*what? what magnitude of impact? under what conditions? for whom? why? and how?*). The approach uses a fit-for-purpose methodology that attends to the New Zealand context including indigeneity and the historic pattern of wide disparities in New Zealand's educational outcomes. The approach requires rigorous eclectism, attention to theoretical coherence, and vignettes exemplifying findings in practice to be embedded throughout synthesis reports.

Part of the rationale for the focus on impacts on learners is the compelling evidence across studies that have linked educational goals, processes mediating learning and student outcomes, that well-intentioned, caring and experienced teachers and teacher educators can unknowingly teach in ways that have impacts counter to their own goals (Alton-Lee, 2006; Alton-Lee and Nuthall, 1995; Bossert, 1979; Doyle, 1983; Nuthall, 2004; Timperley *et al.*, 2007). The concern for impact on outcomes is similarly critical for well-intentioned policy settings and initiatives that can also have impacts counter to their goals and do harm, for example, policy initiatives related to drug education (Biddulph, 2003).

BES writers are required to draw upon systems thinking about the inter-dependencies and ecological relationships that influence effectiveness of any one part of the education system. For example, the BES focused on family and community influences (Biddulph, 2003) highlights the impact of poverty and health issues such as student hearing on educational outcomes, calling for a wider societal and inter-agency policy response to support educators in their work.

BES development guidelines

One of the key challenges in BES development is the contestation of what counts as rigorous evidence amongst researchers especially when so much of educational research has been traditionally siloed within different paradigms and methodological traditions (Alton-Lee, 2004). In order to gain the confidence of the educational research and

[4] One small experimental study of four parent workshops designed by a teacher adviser enabled students achieving at 18 months behind their chronological age on average to almost catch up the gap, showing an upward achievement trajectory 15 months later. Students who didn't receive the intervention dropped even further behind still reading at an 8-year level as 10 and 11 year olds.

practice communities and their engagement in iterative processes of BES development and use, the Ministry of Education drew upon research expertise across the country. The process included not only research but also policy and teacher union representation to strengthen the approach and to get a high level of agreement about the methodology. The approach taken was to gain agreement about the purposes which then informed a fit-for-purpose methodology described in *Guidelines for Generating a Best Evidence Synthesis Iteration* (*http://educationcounts.edcentre.govt.nz/goto/BES*). This allowed for the beginning of a national, structured and transparent process of dialogue to inform BES development.

The *Guidelines* provide a critical resource to support the collaborative process and are themselves subject to iterative review. While international formative quality assurers have provided valued criticism and substantial suggestions for improving the *Guidelines*, they have been a useful and transparent tool to mediate the iterative process across different stakeholders. Professor Paul Cobb, formative quality assurer for the Effective Pedagogy in Mathematics/Pangarau BES Iteration (Anthony and Walshaw, 2007) commented: "The BES Guidelines are outstanding and are clearly grounded in the hard-won experience of synthesising research findings to inform both policy and practice" (Cobb, 2006).

Rationale for a collaborative approach across policy, research and practice

The decision to take such a collaborative approach meant more time would be needed for BES development but laid the foundations for more impact. While such dialogue is challenging, Ginsburg and Gorostiaga (2003) explain the costs of not taking such a collaborative approach:

> *Dialogue isn't necessarily more efficient, but it's more democratic and, therefore, more effective.…Our preference is also based on the belief that in the long run dialogue and participation by a wide range of stakeholders produce better and more relevant educational research, policy and practice. …Certainly, it may be easier – and, in that sense, more efficient – for researchers, policy makers, and practitioners in education to engage in action (or even in praxis) in isolation of members of the other groups. However, the decisions that are made and the actions that are pursued are likely to be less effective. This is the case not only because the quality of judgements may be lower but also because the activities of one group may detract from or cancel out those of other groups. (p. x)*

There is a mandate within the New Zealand public service for the kind of intensive engagement with stakeholders used in BES development. Eleven case studies of innovation in the New Zealand public service (Wright and de Joux, 2003) identified the following implications for effective and innovative policy development and implementation:

- Develop diverse and diffuse invisible colleges, partnerships, and collaborations across agencies, individuals and organisations.
- Exploit opportunities by consistent forward planning and engagement with stakeholders.

A recent review of evidence about the links between research and practice (Walter, Nutley and Davies, 2005) found that interactive approaches such as the development of partnerships and collaborations between researchers, policy advisers and practitioners facilitate the adaptation of research findings to local contexts. The reviewers note that

success is constrained by "the time and energy required to establish effective working relationships, differences in culture, goals, information needs, timescales, power, regard, systems and language, issues of project control and direction (p. 344)". The Iterative Best Evidence Synthesis Programme is seeking to negotiate these kinds of constraints through agreed national *Guidelines*, strategic partnerships, power sharing and iterative processes that enable policy workers, researchers and educators to learn not only from emerging BES findings but also from each other.

Iterative processes of stakeholder engagement in BES development

Educational leaders, educators and policy colleagues are able to influence the scoping and the search strategy for a BES development by raising issues from their experience that they consider significant.

This collaborative knowledge building process has forged sector and policy ownership and greater rigour, trustworthiness and usefulness in BES development but is not without its tribulations. For example, when BES writers share early and emergent work in progress then sector stakeholders have the chance to proactively engage with and provide feedback. If early work in progress is used as a political weapon then risk management is heightened in the policy context, researcher writers become vulnerable, and the iterative process may be threatened. The process needs trust to work.

Strategy for use

BES has been instrumental in enabling teachers to recognise and reclaim the research on educational practice as their own. Because New Zealand has a highly devolved school-based management model, a partnership with educational leaders, particularly principals, will be critical to the potential of BES being realised. The Secondary Principals' Association of New Zealand and the Principals' Council have been proactive in supporting and contributing to the work of the Iterative BES Programme. However, New Zealand primary principals' conferences have featured some concerns and reservations about BES (Flockton, 2005)[5]. The New Zealand Principals' Federation supports the Leadership BES in principle but is concerned about the paucity of outcomes linked to research on New Zealand educational leadership and whether the BES will reflect the reality of school leadership.

The iterative process has allowed one venue for grappling with and addressing the fears of some stakeholders. But it has been when BES findings have fulfilled their promise for principals,[6] and word has spread amongst the networks, that remarkable shifts in student achievement, enjoyment of learning or other valued outcomes are occurring as a result of teachers and leaders using BES, that concerns diminish and the work is valued. Early findings from the Educational Leadership BES (Robinson *et al.*, 2007) emphasise how important pedagogical knowledge is for effective school leadership, particularly when integrated with a

[5] Flockton, L. (2005, July), Closing address to New Zealand Principals' Federation Annual Conference, Otago, New Zealand.

[6] Dr Lorna Earl is being contracted to develop a protocol for evaluating the sector-led developments in which principals have tracked marked improvements in student academic outcomes and enjoyment of learning linked to use of BES or particular approaches sourced through access to BES.

transformational approach to leadership that involves staff in decision-making. This is a challenge in New Zealand where the school-based management reforms of 1989 favoured a more generic management model (Task Force to Review Educational Administration, 1988).

The challenge is to communicate the synthesis findings in ways that facilitate their effective use by leaders, teachers and teacher educators (as opposed to a death-by-bullet-point approach). BES writers Graeme Aitken and Claire Sinnema (forthcoming) have been pulling together evidence about the ways in which research information can be more effectively presented to teachers.

Early findings from the *Teacher Professional Learning and Development BES* (Timperley *et al.*, forthcoming) are compelling. That BES includes an analysis not only of what facilitates the kind of teacher learning that made marked improvements in student outcomes, but also, analyses of interventions that led to student achievement deteriorating from what it had been before intervention. Such findings will be critical in policy development.

The findings highlight the importance of external and challenging expertise with strong pedagogical content knowledge to facilitate and support changes in practice; although poor expertise even from the research community can result in negative impacts on student outcomes. The findings indicate the importance of engaging teachers' theories and challenging discourses that are a barrier to improvements for some students. The findings highlight the importance of sufficient time for extended opportunities for teachers to learn and of the importance of using time effectively – particularly using diagnostic information about students' understandings in a teacher's own context.

Brokerage from a policy agency: constraints and opportunities where there is an evidence gap

The Iterative Best Evidence Synthesis Programme carries out its brokerage role from a government agency, the New Zealand Ministry of Education. The Ministry of Education has a commitment to strengthening the evidence-base informing policy.[7] This commitment is critical within a policy context not only for the use of BESs but also for the integrity of BES development to ensure that the outcomes-linked findings produced cannot be altered for immediate political exigencies but are a trustworthy product transparently generated through an open process.

Perhaps the most substantial gap in the available evidence-base is that which explains the links between policy decisions, activity and outcomes for diverse learners, or explains the communication, organisational learning and other processes that mediate policy decisions and activities. Reid (Reid, 2003) could find no significant international or national body of academic research on the actual process of research integration with policy as seen from the policy advisers' viewpoint.

Court and Young (2003), in their study of fifty case studies in developing countries, found two critical factors influencing policy uptake of research to be:

- the nature of the *evidence* and whether the research was credible and relevant in terms of operational usefulness and problem solution; and
- the *social context* linking researchers and policy makers.

[7] The Ministry of Education has an explicit commitment to effectiveness and "Evidence-based Policy and Practice (p. 23)" within its *Statement of Intent* 2006-2011, Ministry of Education, Wellington.

BES brings strengths consistent with both of these findings. However, Court and Young (2003) found that political context was the most important factor affecting the degree to which research had an impact on policy.

A recent study of effective innovation within the New Zealand public sector (Wright and de Joux, 2003) found the following to have to been critical to success: sufficient resources; tireless risk management; senior management support, mandate, commitment, faith and trust; and management of diverse stakeholder interests, concerns and their tolerance for risk. Risk is a big issue in a democracy where evidence of what does and doesn't work can be a gift to the political opposition particularly if current government policy is inconsistent with the findings (Levin, 2005). The risks would be heightened if a government was not briefed early and its policy agencies were not proactive in integrating the implications of new findings into its work. Cranefield's (2005) study of knowledge transfer in the New Zealand State Sector found organisational factors (such as CEO support), knowledge-related factors (such as representation of knowledge and the strategy for staff engagement with the new knowledge), and gatekeeper-related factors to be critical to a shift towards outcomes-focused policy.

Court and Young (2003) found that policy uptakes were greatest where influencing and communication strategies were in place from the beginning of research programme. Kirst (2000) noted a discrepancy between the pervasive view that policy research either does not reach or is not used by educational policy advisers and the frequent citation or acknowledgement of policy research in the United States. Kirst noted that decades of research on issues in research dissemination help to explain this gap. Nutley, Walter and Davies' (2003) *Framework for Understanding the Evidence-into-Practice Agenda* helpfully suggests six research fields that may advance knowledge about "research utilisation". These are research on: diffusion of innovations, institutional theory, managing change in institutions, knowledge management, individual learning and organisational learning. Drawing upon this framework, adding in a consideration of information literacy, and conducting an interview study about the use of BES within the Ministry of Education, a small pilot study has been carried out to help inform our developing theories of action, communication strategy and strategic planning about policy influence (Moore, 2006). A strength of the BES approach in the policy context at this time is the use of relevant policy partners to collaborate throughout each BES development so that the iterative process and emerging findings feed progressively into policy thinking from the outset.

The single most compelling finding across the BESs is that effective R&D has enabled educational practice to make a much bigger positive difference for diverse learners. In the light of Coburn's (2003) analysis of the evidence of a history of failed educational reform, the magnitude of positive impact for, the responsiveness of, the sector ownership gained and the futures orientation of the most effective R&D are compelling. Often such R&D has gone through many iterations to create the kind of educational development that can work powerfully for diverse learners. As an initial step, through funding educational researchers and the collaborative and iterative processes necessary to undertake first iteration BES developments, BES is seeking to build the capability of the national research community to transform relevant but fragmented research knowledge into a more useful tool for both policy makers and practitioners. BES is also seeking to steer the research community towards a greater focus on informing educational development through R&D.

Each completed BES iteration is an invitation to researchers and educators to engage with the gaps in our knowledge base, the areas of need and the areas of most potential to contribute more deliberatively to a cumulative agenda to strengthen educational practice. The vision is that the Iterative BES Programme will act as a catalyst for policy makers to fund, and researchers and practitioners to build, an integrated outcomes-focused research-and-development culture in education that enables systemic capability building, transformation and sustainable renewal.

There are significant challenges for building national capability in effective educational R&D and cumulative innovation. Such challenges arise particularly in the tertiary and initial teacher education sector where non-research linked market competition has been a recent New Zealand policy model. Education has a low profile in R&D in New Zealand (MORST, 2006). The forthcoming *Teacher Professional Learning and Development BES* demonstrates that New Zealand's most effective research-and-development in education compares relatively well internationally. But education jurisdictions are under-investing in R&D internationally (OECD, 2003) and where there is investment it may not be helpful, or can even be a waste of investment. In the OECD comparison cited above R&D has a much wider meaning to denote research in general so even these estimations are conservative for productive R&D. Despite recent initiatives New Zealand is under-investing in R&D even compared to relatively small investment in other OECD countries.[8]

While there are world-class and even internationally leading researchers in New Zealand education, tertiary academics in education overall, particularly in teacher education, are predominantly research inactive or the quality of their research is not judged to be high by their colleagues (Alcorn *et al.*, 2004). Because of the influential role of the tertiary sector in credentialising both professionals and knowledge, the follow-on effects for initial teacher education cannot be under-estimated.

Where educational research is of high quality, much research may be of interest to academics for its own sake, but not concerned with, or useful for, improving practice. Reward structures and hierarchies for academics can mitigate against rather than value as high status, productive R&D. If the BES strategy is to be more than a set of remarkable books frozen in time the challenge posed in the *2006 World Yearbook of Education* (Ozga *et al.*, 2006) "steering the knowledge-based economy …research steering in national contexts" needs to be seriously and strategically addressed in New Zealand educational policy and research. Because of the critical role of education in society our future as a knowledge society will depend on it.

In conclusion, BES is a collaborative knowledge building strategy. The approach draws upon the expertise and engagement of policy, research and practice communities in education to develop and use bodies of evidence that explain what works and why in education with careful attention to context. The strategy is to use BES as a catalyst for inquiry, cumulative R&D and systemic change. The question of whether such a programme will be sustainable is an open one.

[8] "At the same time New Zealand invests far less in research and development of any kind than other developed countries, and has far lower R&D personnel per million population than Australia or Western European countries. New Zealand is successful educationally, but is, by R&D standards, not becoming a knowledge economy" (p. 89, OECD, 2003).

References

Aitken, G. and C. Sinnema (forthcoming), "Effective Pedagogy in Social Studies/Social Sciences/Tikanga-a-iwi: Best Evidence Synthesis Iteration", Ministry of Education, Wellington.

Alcorn, A., R. Bishop, C. Cardno, T. Crooks, P. Fairbairn-Dunlop, J. Hattie, A. Jones, R. Kane, P. O'Brien and J. Stevenson, J. (2004), "Enhancing Education Research in New Zealand: Experiences and Recommendations from the PBRF Education Peer Review Panel", *New Zealand Journal of Educational Studies*, Vol. 39(2), pp. 275-301.

Alton-lee, A. (2003), *Quality Teaching for Diverse Students in Schooling*, Ministry of Education, Wellington, *http://educationcounts.edcentre.govt.nz/goto/BES*

Alton-Lee, A. (2004), "Improving Educational Policy and Practice through an Iterative Best Evidence Synthesis Programme", Paper prepared for the April 2004 Joint OECD US Evidence Based Policy Research in Education Conference, Washington DC, *http://educationcounts.edcentre.govt.nz/goto/BES*

Alton-Lee, A. (2006), "How Teaching Influences Learning: Implications for Educational Researchers, Teachers, Teacher Educators and Policy Makers", *Teaching and Teacher Education*, Vol. (22), pp. 612-626.

Alton-Lee, A.G. and G.A. Nuthall, with J. Patrick (1995), "Reframing Classroom Research: A Lesson from the Private World of Children", in G. Capella Noya, K. Geismar and G. Nicoleau (eds.), *Shifting Histories: Transforming Education for Social Change*, Harvard Educational Review, Reprint series No. 26, Cambridge MA.

Anthony, G. and M. Walshaw (2007), "Effective Pedagogy in Mathematics/Pangarau: Best Evidence Synthesis Iteration", Ministry of Education, Wellington, *http://educationcounts.edcentre.govt.nz/goto/BES*.

Biddulph, F., J. Biddulph and C. Biddulph (2003), "The Complexity of Community and Family Influences on Children's Achievement in New Zealand: Best Evidence Synthesis", Ministry of Education, Wellington, *http://educationcounts.edcentre.govt.nz/goto/BES*

Bossert, S. (1979), Tasks and Social Relationships in Classrooms: A Study of Instructional Organisation and its Consequences, Cambridge University Press, London.

Cobb, P. (2006), "Formative Quality Assurance Report", Vanderbilt University, Tennessee, 13 February.

Coburn, C. (2003), "Rethinking Scale: Moving beyond Numbers to Deep and Lasting Change", *Educational Researcher*, Vol. 32(6), pp. 3-12.

Court, J. and J. Young (2003), "Bridging Research and Policy: Insights from 50 Case Studies", Working paper 213, Overseas Development Institute, London.

Cranefield, J. (2005), "Inter-organisational Knowledge Transfer in the New Zealand State Sector: An Investigation into Enablers and Barriers from the Gatekeeper's Perspective", unpublished research project submitted in partial fulfilment of a Master of Information Management, Victoria University, Wellington.

Doyle, W. (1983), "Academic Work", *Review of Educational Research*, Vol. 53, pp. 159-199.

Flockton, L. (2005), *Insights and Oversights 1. Education Policy and Regulation: Evidence Based or Babble and Quackery?*, Column on the New Zealand Principals' Federation website, *www.nzpf.ac.nz/main.htm*

Ginsburg, M. and J. Gorostiaga (2003), "Dialogue about Educational Research, Policy, and Practice: To What Extent is it Possible and Who Should be Involved?", *Limitations and Possibilities of Dialogue among Researchers, Policy Makers and Practitioners: International Perspectives on the Field of Education*, Falmer, New York, pp. 1-36.

Kirst, M.W. (2000), "Bridging Education Research and Education Policymaking", *Oxford Review of Education*, Vol. 26(3&4), pp. 379-391.

Levin, B. (2005), "Improving Research-policy Relationships: The Case of Literacy", in N. Bascia, A. Cumming, A. Datnow, K. Leithwood and D. Livingstone, *International Handbook of Educational Policy*, Springer, Dordrecht.

Moore, P. (2006), "Evidence Based Policy Project. Iterative Best Evidence Synthesis Programme", Ministry of Education, Wellington, *http://educationcounts.edcentre.govt.nz/goto/BES*

MORST (2006), *Research and Development in New Zealand: A Decade in Review*, MORST, Wellington, June.

Nuthall, G. (2004), "Relating Classroom Teaching to Student Learning: A Critical Analysis of why Research Has Failed to Bridge the Theory-practice Gap", *Harvard Educational Review,* Vol. 74(3), pp. 273-306.

Nutley, S., I. Walter and H. Davies (2003), From Knowing to Doing: A Framework for Understanding the Evidence-into-practice Agenda, Vol. 9(2), pp. 125-148, Sage, London.

OECD (2003), *Knowledge Management: New Challenges for Educational Research*, OECD, Paris.

Ozga, J., T. Seddon and T. Popkewitz (eds.) (2006), *World Yearbook of Education: Educational Research and Policy: Steering the Knowledge-based Economy*, Routledge, London, p. 170.

Reid, F. (2003), "Evidence-based Policy: Where is the Evidence for it?", Masters Thesis presented in completion of MSc at Bristol University, *www.bristol.ac.uk/sps/downloads/working_papers/sps03_fr.pdf*

Robinson, V. (2007), "The Impact of Leadership on Students", *Educational Leadership: Best Evidence Synthesis Iteration*", Chapter 2, Ministry of Education, Wellington.

Task Force to Review Educational Administration (1988), *Administering for Excellence* (Picot Report), Government Printer, Wellington.

Timperley, H., A. Wilson, H. Barrar and I. Fung (forthcoming), "Teacher Professional Learning and Development: Best Evidence Synthesis Iteration", Ministry of Education, Wellington, *http://educationcounts.edcentre.govt.nz/goto/BES*

Walter, I., S. Nutley and H. Davies (2005), "What Works to Promote Evidence-based Practice? A Cross-sector Review", *Evidence and Policy*, Vol. 1(3), pp. 335-363.

Wright, A. and V. de Joux (2003), *Case Studies of Innovation in the Public Service*, New Zealand Treasury, Wellington, *www.treasury.govt.nz*

Chapter 6
The Canadian Council on Learning, Canada

Charles Ungerleider, Canadian Council on Learning[1]

In this chapter, we are describing the Canadian Council on Learning, which was created in 2004 to: promote knowledge and information exchange among learning partners; inform Canadians regularly of Canada's progress in learning; and address knowledge gaps and provide evidence-based information to improve investments in lifelong learning.

The establishment of the Canadian Council on Learning

The Canadian Council on Learning (CCL) was created as a consequence of a grant from the Government of Canada in March 2004 to provide Canadians with the most current information about effective approaches to learning.

Among nations, Canada is not alone in seeking ways to improve the learning available to its citizens. Member countries of the OECD have set the goal of making lifelong learning a reality for all of their citizens. Many countries have established national-level leadership for developing and sharing learning information. For example, the United States recently created the Institute of Education Sciences within its National Department of Education. The Institute provides national leadership in expanding knowledge and understanding of education through its National Centres dedicated to education research, statistics and evaluation. In the United Kingdom, the Department of Education and Skills has identified National Learning Targets, which focus on increasing participation in and attainment of learning in schools, in the workplace, and throughout life.

The vision behind CCL was originally announced at the national Summit on Innovation in Toronto in November 2002 as the "Canadian Learning Institute". Leaders from all walks of Canadian life – education, business, labour, government, aboriginal organisations and non-governmental organisations of many kinds – agreed that Canada must move beyond rhetoric about lifelong and life-broad learning. They wanted to see links among the various parts of our learning systems – a national roadmap for a culture of learning throughout an individual's lifespan.

Unlike other nations, Canada does not have a federal department devoted to learning. As a consequence, the Government of Canada supported the creation of an independent, not-for-profit organisation with a mandate to: promote knowledge and information

[1] Charles Ungerleider is Director of Research and Knowledge Mobilization for the Canadian Council on Learning.

exchange among learning partners; inform Canadians regularly of Canada's progress in learning; and address knowledge gaps and provide evidence-based information to improve investments along the full span of lifelong learning, on a pan-Canadian basis.

CCL's mandate is extremely broad, encompassing the continuum of lifelong learning (childhood, youth and adulthood), diverse settings (home, classroom, community and workplace), and different forms of learning (informal and formal). CCL's mandate is also inclusive, recognising the linguistic duality of Canada as well as the diverse cultures and circumstances of Canada's population.

Organisation and illustrative activities

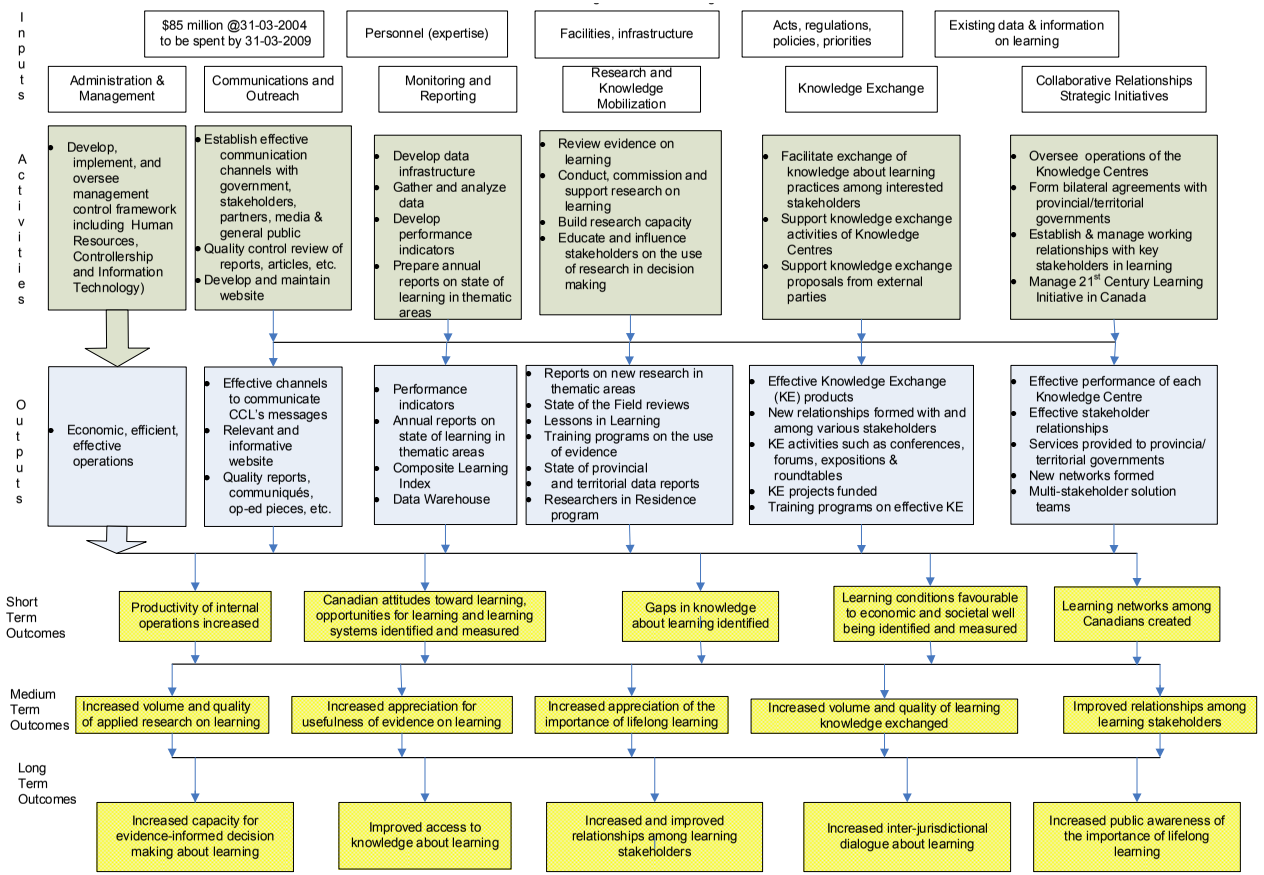
CCL carries out its work through four operational units – Monitoring and Reporting; Research and Knowledge Mobilisation; Knowledge Exchange; Strategic Initiatives; two support units: Communications and Outreach; and Administration and Management (see Figure 6.1); and five knowledge centres, consortia of decision makers and researchers engaged in knowledge exchange in the areas of early childhood learning, Aboriginal learning, work and learning, health and learning, and adult learning. Each of the operational units is responsible for numerous knowledge brokering activities and projects, some of which include:

- The *Composite Learning Index* (CLI): For more than forty years, social scientists have talked about the creation of an index of the learning conditions favourable to Canada's social and economic well-being. In May 2006, CCL developed the Composite Learning Index (CLI)
- *16 indicators* representing the four dimensions of learning identified by French economist and two-time president of the European Commission, Jacques Delors: learning to know; learning to do; learning to be; and learning to live together (*www.ccl-cca.ca/CCL/Reports/CompositeLearningIndex2?Language=EN*) in the Report to UNESCO of the International Commission on Education for the Twenty-first Century (*www.unesco.org/delors/delors_e.pdf*).

Rather than relying on weights assigned by an expert panel or equal weighting, the index was developed using a multi-stage statistical weighting that follows the structure of the four pillars of learning and is based on the relationship between the indicators and outcomes such as adult literacy, civic engagement, employment, income, and health status. CCL is able to generate a CLI score for any of the 4 760 census sub-divisions of Canada. The CLI data is refreshed yearly, providing the opportunity to appraise change over time in the learning conditions favourable to Canada's economic and social well-being.

- *Survey of Canadian Attitudes Toward Learning (SCAL)*: CCL conducts an annual survey of more than 5 000 Canadians to assess their attitudes toward various dimensions of learning, including early childhood learning; elementary, secondary and post-secondary learning; adult work-related learning, and health and learning (*www.ccl-cca.ca/CCL/Reports/SCAL?Language=EN*)
- *Lessons in Learning*: CCL produces a bi-weekly series of electronic publications presenting data, evidence, and conclusions across a broad range of learning issues. Each article addresses a specific learning issue by answering the questions: What is the issue?; Why is the issue important?; and What can we do to address the issue?. Publication of Lessons in Learning began in September 2005 on the CCL website and will continue indefinitely (*www.ccl-cca.ca/CCL/Reports/LessonsInLearning? Language=EN*).

Figure 6.1. Canadian Council on Learning: results-based logic model



- *Question Scans and Reviews of Evidence*: CCL uses question scans, an exploratory tool that gauges the volume of literature in a field and summarises the predominant observations contained in the body of literature, to determine the feasibility of conducting a further, more comprehensive, systematic review of the literature devoted to a given question. The material identified through the question scanning process is not critically appraised for its quality. At this point in its development, CCL has produced more than 50 question scans. CCL also undertakes systematic analyses of the literature devoted to a particular topic, systematic reviews of evidence.
- *Researcher in Residence*: CCL's Researcher in Residence programme provides co-funding to support a "researcher in residence" at various learning-related organisations across the country. The programme is designed to encourage new research initiatives and increase Canada's capacity to perform important research on learning. For example, CCL and local school board have launched a three-year partnership to sponsor a Researcher in Residence to work with the district's community and inner city schools. The main focus of the research is to identify policies and practices that offer support to students at risk, as well as effective ways to track the progress of at-risk learners. The initiative will help build the school board's research capacity at both the staff and district programme level.
- *CCL Suite of Tools to Support Learning*: The Canadian Council on Learning is developing a suite of tools to support learning that is to be available from CCL on its Internet site in 2007. The suite of tools were developed to support the educators who do the challenging work that makes successful learning possible, and consists in an assessment tool, a self-assessment tool, and a tool to collect data from students about their experiences as learners.
    - The CCL Assessment Tool:
        - Automates test construction from an item bank.
        - Scores test data using user-supplied answer keys.
        - Uses plain language to explain how items need to be improved in order to make the final test more efficient and effective.
        - Illustrates statistical results graphically.
        - Provides easy-to-follow instructions.
        - Gauges the extent to which items are measuring what they ought to be.
        - Estimates optimal cut-points for reporting of results by proficiency levels (for norm or criterion-referenced assessments).
        - Can equate test items across exams.
    - The Student Self-Assessment Tool:
        - Enables students to test their declarative and procedural knowledge in a variety of curricular areas at a variety of grade levels.
        - Automates the construction of self-assessments from an item bank.
        - Scores each self-assessment.

  - Uses plain language to explain results, providing easy-to-follow instructions and feedback.
  - Enables any given test to be linked to year-end, grade-wide, district-wide or provincial assessments so that students can learn how they might perform.
- The Student (Parent) Monitoring Tool:
  - Enables data collection from standard or specifically designed instruments.
  - Can be used on a sample or census basis.
  - Permits information about the school, school board, community, neighbourhood, and other sources (such as the Composite Learning Index, CLI) to be linked.
  - Represents data graphically on a geographic basis.
  - Automates analysis and will answer questions (depending upon the data collected) such as: Do the responses of urban and rural learners differ? Do responses vary by the respondents' neighbourhood socio-economic conditions? Do relationships vary across administrative units, and if so, why?
  - Data are revised dynamically and presented in real-time as new data is entered.

Opportunities and challenges

As an organisation in its infancy and with such a broad mandate, the Canadian Council on Learning faces many opportunities and challenges. Its broad mandate – fostering the use of evidence and research to support learning across the life course – means that it must be strategic in identifying from among the universe of opportunities available those that are likely to have the most significant long-term impact.

A related challenge is identifying the relevant audiences for its work. CCL has deliberately used the term *decision makers* rather than *policy makers* as the focus for its efforts, believing that the former term encompasses everyone (individual citizens who must make decisions about their learning; frontline practitioners who must each day make decisions about the learning of those for who they are responsible; and policy makers who set the broad directions that learning takes within their areas of jurisdiction) where the latter term draws attention only to policy makers.

Getting and maintaining the attention of the audiences for its work is another challenge for the Council. CCL exists in an information rich environment in which many organisations seek attention for the purposes of advocacy. CCL does not advocate for particular policies or practices, preferring instead to encourage the careful consideration of the evidence available to inform decisions. It is a challenge for CCL to get and maintain attention in the face of competition from organisations less concerned about evidence.

The approach CCL has taken thus far has helped it to meet another of its challenges: establishing its credibility. By emphasising the use of evidence and research to inform

decisions, CCL is helping to develop the capacity of decision makers to distinguish between factors likely to influence policy and practice and approaches based upon current fashion or ideology.

The breadth of CCL's mandate poses the challenge of working across jurisdictional and perceptual boundaries. The knowledge centres, knowledge exchange, and knowledge mobilisation works of the council have met with some success in encouraging collaboration among individuals and agencies without prior experience working together. Even though it is a pan-Canadian organisation, CCL has also met with success in working with provincial and territorial authorities which have exclusive jurisdiction for education.

Though early in its mandate, CCL faces the challenge of meeting the expectations of the various audiences with which it works. The creation of CCL engendered high expectations for what it might do and how quickly it might accomplish the things that various audiences thought it might do. Managing those expectations, articulating priorities, and meeting the expectations in its priority areas are consistent challenges for the young organisation.

Only half way through its five year mandate, CCL works hard to maintain its focus, meet its commitments in a timely fashion, and – most important – ensure that its work meet the highest standards. The combination of a daunting agenda and a young organisation pose the ultimate challenge for CCL: Can it show sufficient impact in five years to merit its continuation as an agent to provide an evidence base to improve lifelong learning across Canada?

Chapter 7
The Knowledge Clearinghouse, Denmark

René Bugge Bertramsen, Danish University and Property Agency

In this chapter, we discuss how new demands and expectations from the decision makers and stakeholders in educational R&D are affecting and challenging the Danish institutional framework.

Introduction

In recent years the political interest in educational R&D in Denmark has exploded. OECD's PISA results in 2000 and 2003 have been an eye-opener for decision makers at all levels. The PISA examinations indicated that the competences of the Danish pupils at the age of 15 in general were close to the OECD average. To a small, open economy – which depends on the ability to integrate a large share of the population in knowledge-based professions – the results were disappointing and testify to the need to strengthen the quality of education. Educational R&D is perceived as an effective means of addressing the problem of weak performance – not only in the primary and secondary schools but at all levels of the educational system. The purpose of this article is to discuss how new demands and expectations from the decision makers and stakeholders on educational R&D are affecting and challenging the institutional framework within the field.

A central expression of these new demands and expectations is the Globalisation Strategy, "Strategy for Denmark in the Global Economy" (*www.globalisering.dk/page.dsp?area=52*), tabled by the Danish government in April 2006. Strategy was the outcome of a process featuring 14 meetings in the Globalisation Council – a council set up by the government comprising 26 key decision makers and chaired by the Prime Minister.

The Strategy contained 350 specific initiatives, launching a comprehensive reform within the field of education, research and innovation affecting actors from primary schools to higher education, from public research institutions to private businesses. It focused on improving the efficiency of public spending on education and research, in particular by allocating more public funds in open competition and by increasing competition and internationalisation in the Danish economy as a whole. Competitiveness was perceived as a means to ensure prosperity and cohesion in the society.

The initiatives will be financed by a Globalisation Fund. Ten billion Danish Kroner (DKK) will be set aside for future investments up to 2012. The pledge will get gradually

phased in, with another DKK 2 billion allocated per year until 2010 and DKK 1 billion allocated in 2011 and 2012.

As described below the initiatives of the Globalisation Strategy are presently transforming the map of educational R&D in Denmark. Starting from a brief description of the institutional framework of educational R&D, new demands and expectations towards sector are discussed. The discussion is focused on two specific incidences/cases/processes, which are central to the understanding of the actual development of the institutions of educational R&D in Denmark. One is the response to the OECD evaluation on educational R&D in Denmark in 2004, and the other the impact of the governments Globalisation Strategy.

The institutional framework of educational R&D in Denmark

Educational R&D is undertaken by multiple actors. The organisation of the activities is not formalised and coordinated in Denmark. However, the usual perception is that the 12 universities – and a few research institutions – are involved in the research activities funded by the Ministry of Science, Technology and Innovation, while 17-19 university colleges/Centres of Higher Education (CVUs) in partnership with local authorities and schools, are responsible for developmental activities financed by the Ministry of Education.

The institutional framework, as it appears at the moment, is a result of an institutional reform which was undertaken in 2000. One of the central elements in the reform was the establishment of the Danish University of Education (DPU) by a merger of the Royal Danish School of Educational Studies, the Danish School of Advanced Teaching and the Danish Educational Institute. The aim was to create a leading international research university in education. Learning Lab Denmark (LLD) was merged into DPU as well.

DPU appears as the largest environment for academic research in didactics and education in Denmark. In a European context the university is also unique. It employs approximately 200 researchers in permanent positions, more than a third of the researchers (calculated as full year's work for one person) conducting educational studies in Denmark. Approximately 5 000 students are enrolled in study programmes, making the university the largest in Europe within the field. An international alliance with similar universities in Europe has been initiated. It is expected that the Institute of Education, University of London, which is already a well-known collaborator, will be part of this alliance. Additionally, DPU is a leading partner in the ASEA-project: the ASIA-EUROPE Education and Research Hub for Lifelong Learning.

Another central element of the reform launched in 2000 was the merger of a wide range of institutions in the college sector into new Centres of Higher Education, CVUs. CVUs offer bachelor degrees for teachers at the primary school level, nurses and other professions in the health sector as well as degrees in engineering and business. In the legislation of the CVUs, it was envisaged that they be "research affiliated" and this was ensured through cooperation agreements between CVUs and relevant research institutions in Denmark (as well as other countries) on user-inspired research projects. The DPU has a specific duty to support CVUs in R&D projects. The purpose of the "research affiliation" of the CVUs was to ensure that the CVUs had ready and on-going access to recent research-based knowledge, to develop qualifications of CVU teaching staff and to build bridges between "scientific research" and "applied research and development". At the moment there are 18 CVUs.

Since 2004 a number of regional knowledge centres have been established. The purpose of the centres is to collect, process and transmit knowledge established by institutions offering short-cycle and medium-cycle study programmes to researchers, teachers, students, field workers, parents, etc. Many centres involve CVUs with special competences of didactics. This is for example the case in relation to the Centre for Reading Research, the Centre for Curriculum Education and Development, the Centre for Adult Learning and Education. The Ministry of Education is funding the centres, with an annual budget in 2007 of nearly DKK 50 million.

New expectations and demands

The Danish response to the OECD evaluation of educational R&D

On the request of the Minister of Science, Technology and Innovation, OECD examined national educational R&D in Denmark in 2004 (*www.oecd.org/dataoecd/56/21/33888206.pdf*). The purpose of the review was to assess the extent to which it serves its function of creating, collating, distributing and applying the knowledge on which practitioners and policy makers can draw. Thus the aim of the examination was broader and different from a traditional educational R&D review that would focus on the quality of the research. In particular, the examination focused on interactions between producers and users of educational R&D.

OECD concluded that there was no explicit national strategy for educational R&D. The R&D activities were organised in a large number of small-scale projects largely determined by the individual interests of the researchers. Coordination and dialogue between the various actors – especially between researchers at the universities on the one hand, and trainee teachers at the university colleges on the other – were modest. In general, there appeared to be a lack of basic research, a low capacity to apply a range of research methods (in particular quantitative methods) and few mechanisms for accumulating and sustaining the knowledge.

Based on these conclusions and on experiences from the United States and the United Kingdom, the OECD suggested the establishment of the following new institutions:

- A National Education Research Forum, which could hold regular meetings (annually, or two or three times a year) for researchers, teachers and policy makers to share ideas on education issues and needs.
- A clearinghouse on education, which could be a means to offer regular reporting on significant educational research gained nationally as well as internationally.
- An Educational Observatory in Denmark.

In the wake of the OECD report, the Ministry of Education and the Ministry of Science, Technology and Innovation initiated a process of round table discussions to clarify the organisation of the new institutions.

There is general acceptance of the need to build new institutions at the national level in order to develop a national research strategy and strengthen coordination between actors. However, there seems still to be widespread disagreement on strategies, priorities and needs. At the first conference in the National Education Research Forum held in March 2006 these disagreements were exposed.

Two questions have been focal points of the differences of opinions.

The *first question* concerns the concept of evidence. The disagreement is based on a fundamental academic dispute of the definition of the concept of evidence and consequently the appropriateness of the idea of establishing a new clearinghouse focusing on evidence-based research. Should the concept of evidence be restricted to formal, quantitative empirical research based on classical experimental design model? Or should the concept be more broadly defined enclosing a wider array of interpretative disciplines and implying a close interaction between student and practice in the search for evidence? The discussion has partly run parallel to the discussion between the university and university college sector on the dissemination of evidence, with the former sector tending towards the narrow definition.

The *second question* concerns the organisation of educational research in general and the new clearinghouse in particular. At the moment DPU is carrying out 40% of the research projects, while other universities and research institutions are responsible for 60% of the projects. Some have argued that the process of establishing a new clearinghouse indicates a political ambition to control the research agenda and concentrate activities. The dominant role of DPU in the clearinghouse process has been criticised by a group of universities, which is also opposing a further centralisation in the sector.

Due to the implementation of the government's Globalisation Strategy, and especially the ongoing committee work concerning a merger of the DPU with a university (see below), the process of building new institutions is put on hold at the moment.

University mergers and the Strategy for Denmark in the Global Economy

Among other measures in the Globalisation Strategy the government proposed a major institutional reform directed towards institutions offering higher education. In the university college sector the aim was to merge the 17-19 existing CVUs into 6-8 multi-disciplinary, regionally based university colleges. In the university sector the ambition was to create world class universities by integrating the 13 government research institutions in 12 universities.

In October 2006 the government made a formal decision in relation to the issue of university mergers. The decision implied that the current 25 universities and research institutions by January 2007 will be reduced to 11, and 97% of their activities will be concentrated in seven universities (*www.ubst.dk/uk/page_university_mergers.html*).

Denmark will have three large universities (University of Copenhagen, University of Aarhus and Technical University of Denmark), four medium-sized universities (Aalborg University, University of Southern Denmark, Roskilde University and Copenhagen Business School) and one small university (IT University).

The position of the DPU was not a part of the decision made by the government in October 2006. The reason was that the government in June 2006 decided to set up a committee to describe models of a merger of DPU with University of Copenhagen or another university. Below, considerations concerning the future of DPU are outlined.

New solutions

The recent development in the wake of the Globalisation Strategy is expected to change the initial set up fundamentally and in effect implement the solutions proposed by the OECD.

The position of DPU is not finalised, but the most likely result seems to be a merger of DPU in University of Aarhus. A number of reasons make the University of Aarhus an attractive partner. First, the new University of Aarhus will be a modern university with a unique research profile combining basic research, user-oriented research and research services for public authorities. Secondly, as the merger enters into force in 2007 the university will introduce a new model of organisation. Inspired by the Anglo-American tradition the faculty model will be supplemented by a new model of university school. The DPU is expected to merge into University of Aarhus as a school of education. This construction enables close contact and cooperation between the new university school and representatives of the users and the professions. The cooperation will be formalised in an advisory board. The Board will advise the university school about future educational needs, development of new study programmes, research strategies and strategies for cooperation between institutions in the field.

The specific conditions of the merger and the new advisory board are presently (December 2006) uncertain. However, the results of the actual negotiations between University of Aarhus and DPU are highly relevant to the process of building new, national institutions, to define national strategies on educational R&D and to secure cooperation among the actors within the field.

It is expected that the ideas of a National Education Research Forum and clearinghouse will be further developed as the institutional reforms of universities and university colleges are in place.

Chapter 8
The Knowledge Chamber, Netherlands

Hans Stegeman and Rien Rouw
Dutch Ministry of Education, Culture and Science

*In this chapter, we shed light on the Dutch "Knowledge Chamber" (*Kenniskamer*). This Chamber was created in 2006 to bring together stakeholders on education policy and on knowledge of education policy in an environment which takes into account both politics and knowledge.*

Introduction

Education policy is a sensitive phenomenon. Schools are jealous of their autonomy, but at the same time desire guidance from authorities. Authorities define the problems which they wish to solve and design education policies which should bring about solutions. At the same time, interested third parties (parents, employers) are often quite outspoken in formulating their wishes.

Within this complicated framework, which is further bedevilled by the dynamics of politics, the processes leading to concrete policy-measures are often not primarily shaped by rational and knowledge-oriented considerations. It is probably an illusion to think that a completely rational and knowledge-oriented method of policy-making is possible. There will always remain conflicts of interest and struggles for influence. But also in a highly-politicised environment evidence may play a role. In the Netherlands, an effort is currently being made to bring together stakeholders on education policy and on knowledge of education policy in an environment which takes into account both politics and the knowledge factor. This is the so-called "Knowledge Chamber" (*Kenniskamer*), which met for the first time in the summer of 2006, on the initiative of the Dutch Ministry of Education, Culture and Science.

In the Netherlands, the international debate on the importance of evidence-based policy-making has not gone unnoticed. Also the Netherlands have been internationally active, both in the debate generated by OECD as well as, on their own initiative, in putting evidence-based policy-making on the "European agenda". At the same time, the influential Education Council of the Netherlands, the government's chief advisory committee on educational matters, has emphatically demanded attention for the necessity to base educational policy-making on a more evidence-based footing (January 2006). And the Advisory Council for Science and Technology policy (AWT – *Adviesraad voor Wetenschaps- en Technologiebeleid*) published in May 2005 an advice on the knowledge

policy of the national government called *Knowledge for policy – policy for knowledge.* Both councils emphasised the necessity of a coherent knowledge policy. These developments took place against a background of increasing civil dissatisfaction with educational policy. Retroactively many of the policies and strategies which had helped shape education since the 1980s were called into question or even repudiated as downright counterproductive. Great strategies were partially dismantled, as in the case of *basisvorming* ("basic education"), the semi-comprehensive schooling system for 12-16 year-olds, or the *studiehuis* ("studying house") which was to enable 16-18 year-olds to develop independent learning capacities. Doubts were cast on the very capacity of the ministry of Education, Culture and Science to develop effective policies at all.

The Knowledge Chamber is partially an expedient political response to recently expressed doubts concerning educational policy-making. But in the Dutch context it also represents a time-honoured method of tackling problems in that it brings together government and stakeholders in a structuralised give-and-take of views, information and knowledge. As such, there are good conditions to produce, offer and obtaining knowledge, while at the same time there remains a certain measure of room for political manoeuvring.

In this paper we first shed some light on the background of the Knowledge Chamber and after that we will discuss the design of the Chamber.

The Ministry desires a new way to deal with knowledge

Before we get to the background of the Knowledge Chamber it is useful to specify the notion of "knowledge" we use in this paper. With regard to the work of the Knowledge Chamber we follow the AWT and define knowledge "as being empirical data, concepts, analyses and theories that are considered true and correct and enable people to take decisions" (*www.awt.nl/uploads/files/a63uk.pdf*). This means codified, stored and traceable knowledge that is publicly validated, mostly in an academic forum. This kind of knowledge is often the result of scientific research, which is performed mainly at universities, research institutes, planning offices and advisory councils and less frequent by commercial consultants.

What then is the background of the Knowledge Chamber on Education, Culture and Science? A very immediate reason is the reorganisation plan or programme for action that the ministry of Education, Culture and Science issued at the end of 2005, called "OCW changes" (*OCW verandert*). One of the main lines of that programme is the aim to make effective policy (a paraphrase of "beleid dat werkt in de praktijk"). To reach that goal the action programme states that we need to strengthen the scientific knowledge basis of our policies or, as it is often called in international discourse, we need evidence based policies. The reason for this speaks for itself. Up to now, a policy proposal may only be based on a single study, while at the same time political pressure may be exerted to implement that proposal. This is not an exceptional state of affairs, neither in the Netherlands nor in other countries, as we learned from the OECD project on evidence-based policy research. The Ministry draw the conclusion that it needed to include researchers and experts in policy-making to share their views and insights with policy advisors in order to bring scientific evidence in. Therefore the Ministry established among other things the Knowledge Chamber.

Mobilising top-ranking officials to minimise overkill, compartmentalisation and process-fetishism

The Knowledge Chamber is a consultative body of the top-ranking officials of both the knowledge institutions and the Ministry. Why does it exactly focus on the top? In answer to this question, we are getting at a second reason for the realisation of the Knowledge Chamber.

The above-mentioned advice of the AWT-Council concluded on the basis of an inventory that only a few departments have formulated an explicit knowledge policy. Without such a policy government runs various risks, according to the Council.

- First there is the risk of an excess of knowledge and information. As the amount of data and information is constantly rising it is becoming more difficult to pick up relevant research, to interpret it correctly and to link it to knowledge already available.
- The second risk according tot the Council is compartmentalisation in knowledge domains. The compartmentalisation between and within departments is reflected in the way the knowledge infrastructure is organised, namely in separated domains. An integral approach is hindered by the compartmentalisation of knowledge.
- The third risk the Council distinguishes is that government officials, especially at the top, concentrate on the process of policy-making rather than on the content of a certain policy. According to the Council, the national government's personnel management nowadays values process-related skills more than expertise concerning content. As a consequence it can happen that (mostly) senior executives lack the understanding to examine the evidence base of policy proposals properly.

It is to minimise these risks that the Ministry has made the Knowledge Chamber into a crucial ingredient of its knowledge policy and follows the view of the AWT, which stresses the need to formulate knowledge policy at the top, starting from a strategic vision of the role of knowledge in policy. The top-ranking officials after all are ultimately responsible that policy is evidence based. Besides that, departments not only need evidence for the formulation and implementation of (short-term) policy measures. They also require perspectives on long-term developments in order to be able to formulate long-term strategies and to prioritise policy issues. This is a typical responsibility of top management.

Modernising government

A third reason for the establishment of the Knowledge Chamber follows from the government programme *Andere Overheid* ("Modernising Government"), which aims to realise "a powerful and decisive government, which puts the community foremost". One of the initiatives of *Andere Overheid* was the restructuring of the system of advisory councils and knowledge institutions (such as planning offices and research institutes). The results of this initiative were set down in a letter from the cabinet to the parliament. In this letter the government stated that the direct interaction between policy makers on the one hand and knowledge institutions and researchers must be improved. This corresponds with a recommendation of the AWT, namely that constant interaction is

needed to further the active use of knowledge by policy makers. The letter also states that the form in which this interaction is organised is up to each ministry to decide. A knowledge chamber is strongly recommended but not prescribed. Every ministry must devise an arrangement that suits the conditions on their domain the best.

A knowledge chamber is not a strictly defined entity. The essence is interaction between policy and research. In its letter the cabinet distinguishes several variants such as a "narrow chamber" in which the interaction focuses on the information needed for current policy programmes, and a "broad chamber" which is not limited to current policy but which also explores long term issues. Differences can also arise in the degree of independence of the knowledge chamber. Is it purely a unit within the civil service and staffed by government professionals or is it organised as an agency at a distance from the ministry? Other variations concern the composition of a knowledge chamber. For instance, the Ministry of Transport, Public Works and Water Management organised their knowledge chamber as a broad meeting of representatives of universities, research institutes, planning offices, social organisations, central government and local governments. The Ministry of Economic Affairs, on the other hand, is inclined to arrange its knowledge chamber as a consultation of the top management with a small group of the most concerned knowledge institutions.

The cabinet expects that by exchanging various experiences with and good practices of knowledge chambers, ministries not only will learn from each other but also that the best practices will ultimately prevail. The development of knowledge chambers is expressly designed as a learning process, a process of weighing up the pros and cons of various models.

The essence: structural consultation on knowledge

In June 2006 a varied group of people gathered in the Ministry of Education, Culture and Science in The Hague. Present were not only the top-ranking officials of the Ministry but also chairmen and managers of various knowledge institutions in the fields of education, culture and science, such as advisory councils, planning offices, the Netherlands Organisation for Scientific Research, the Education Inspectorate and a commission for long term enquiries. It was the first and founding meeting of the Knowledge Chamber, a consultative body on the production of knowledge that serves government policies on education, culture and science. We will now examine the design of that chamber.

The essence of the Knowledge Chamber is a structural consultation between the Ministry of Education, Culture and Science and those knowledge institutes to which the Ministry formally assigns knowledge-related tasks. This essence will be the focus during the first stage of the Knowledge Chamber, the stage during which the Chamber is constructed and tested.

The Knowledge Chamber will meet, in principle, twice a year, in spring and in autumn. The spring meeting will mainly be used to programme and plan the "knowledge agenda" for the following year, which will be reflected in the working plan of the knowledge institutes. The autumn meeting will perform a review of the Knowledge Chamber's activities, addressing questions like "what knowledge did the activities of the Chamber yield" and "how was this knowledge put to use by policy makers". In addition to the semi-annual regular meetings of the Knowledge Chamber there will be one or more

special meetings, which will be organised from a specific angle and aim at broadening the Ministry's perspective.

Participants in the meetings of the Knowledge Chamber will consist of two categories. The regular meetings will be attended by an inner circle composed of the Education Council, the Advisory Council for Science and Technology policy, the Culture Council, the Scientific Council for Government Policy, the Netherlands Bureau for Economic Policy Analysis, the Social and Cultural Planning Office of the Netherlands, the Netherlands Organisation for Scientific Research, the Inspectorate of Education, the innovation consortium SenterNovem and the Consultative Committee of Sector Councils for Research and Development. For the theme-oriented special meetings also an outer circle will be invited, of varying composition.

The Knowledge Chamber must become the nucleus of a network of knowledge workers and policy makers. To promote this, the regular meetings of the Knowledge Chamber could be broadcasted through the Intranet facilities of the involved agencies, possibly offering staff members of the agencies the possibility to intervene and to formulate questions. A digital "home" for the Knowledge Chamber might help realise such a network function.

The essence of the Knowledge Chamber being a structural consultation between the Ministry and the knowledge institutes, in a practical sense this essence will be translated into identifying themes on which knowledge must be accumulated, following a "rolling agenda" – a knowledge agenda which may be brought up to date at each meeting of the Chamber, specifically during its "programming" springtime meeting. During the regular meetings, one or two specific themes may be highlighted – by one of the Chamber's members or by an external expert.

The meetings of the Knowledge Chamber will be prepared and facilitated by the Ministry's directorate for strategy.

Generating validated knowledge

The Knowledge Chamber will deal in *validated* knowledge, which may be used by policy makers. Knowledge stemming from scientific research is pre-eminently a form of validated knowledge. Such knowledge will consist of analytical studies of trends and developments in the educational field, on behaviour and perspective of stakeholders within the field and on the efficiency of institutes. The Knowledge Chamber will help to make such knowledge available and accessible. Another important form of knowledge concerns a perception of whether the instruments which the Ministry designs actually work. Evaluations will follow to determine whether policy aims have been effectuated (possibly coupled to *ex ante* evaluations of newly designed policy instruments).

Other possibilities are still in the future. Thus, the Knowledge Chamber may ultimately come to validate policy proposals by assigning a quality hallmark. Such a hallmark would be a stimulus for the knowledge-orientation of policy makers. Also, the Knowledge Chamber may some day empirically judge the sustainability of major policy projects, *e.g.* through screening the policy documents.

Organising creativity

May be the most challenging part of the Knowledge Chamber's task is the formulating of "knowledge questions". Thus, questions are identified which transcend the traditional boundaries of policy areas. The Knowledge Chamber will examine themes from a future-oriented perspective and from the angle of other policy fields. General knowledge questions which are important for each subject are: what is the actual problem? What are the perspectives of the stakeholders involved with the problem? Which instruments are effective and/or efficient (also in terms of financial costs)? Is it possible to identify effective government interventions to help solve the problem?

But above this basic knowledge questions, the Knowledge Chamber must operate creatively and from an innovative perspective. The Ministry expects from the Chamber sensible ideas on an always uncertain future, creativity, new and surprising perspectives and cross-grained views. These are notoriously difficult to organise. But the Chamber may use innovative debating and presentation techniques which are conducive to creativity. Elements like "digital storytelling" and mobilising new and/or different talents (students, pupils, artists) may be part of this approach. Thus, it is hoped that the essential product of the Knowledge Chamber, validated yet challenging knowledge which transcends traditional policy paradigms, will radiate from the Chamber and permeate both the Ministry and the educational institutions.

The Knowledge Chamber will be judged a success when its activities lead to the actual use of validated knowledge by policy makers – and when the furnishers of knowledge become aware that their efforts count. Of course, this implies that the knowledge institutes should produce *usable* knowledge – that is: knowledge which fits the actual process of policy-making and which makes clear the implications and consequences of the proposed policies. In order to be able to judge the results of the Knowledge Chamber, indicators will be developed to measure the above-mentioned criteria. In any case, the Knowledge Chamber will have to prove itself flexible and capable of adapting itself to new demands.

Chapter 9
The Social Care Institute for Excellence, United Kingdom

Bill Kilgallon, Chief Executive, Social Care Institute for Excellence
United Kingdom

In this chapter, we describe the Social Care Institute for Excellence, which is one of the foundations of the 2000 UK strategy to improve social care. The Institute works on the development of a knowledge base in social care, to provide the underlying knowledge on which other social organisations could build.

Background

Social care supports people who need help with the day-to-day business of living. Social care serves older people, people with learning disabilities, people with mental ill health, people with problems of substance abuse and people with physical and sensory disabilities. It supports families and children. In some cases people have no choice as to whether or not social care gets involved in their lives, such as when there are concerns about the safety and well-being of children.

Adults are supported in the community through home care, sitting services, meals, day services and social work. Some receive support in residential care homes and nursing homes. Children and families are supported at home through a wide range of child protection, social work, early years and other services. Sometimes fostering, residential care, or adoption may be necessary for children. At its best social care can transform people's lives, enabling them to live the lives they choose, in the way they want to. Social care plays an important role in wider policy areas including social inclusion and citizenship. Liam Byrne, the Care Services Minister in 2005 said "Across the breadth of the domestic policy agenda – in health, education, criminal justice and welfare to work… social care is mission central."[1]

Unlike education social care is not a universal service. Access depends on an assessment of need. People using social care services for adults are subject to means testing and may be required to pay for all or part of the costs of the service they require.

The education workforce consists largely of professionally qualified teachers with some ancillary staff. The social care workforce on the other hand is not professionally qualified. Of the over one million people working in social care in the United Kingdom

[1] Liam Byrne MP, Speech to Care and Health conference, 4 October 2005.

only about 80 000 are qualified social workers. The others will have access to training at National Vocational Qualifications (NVQ) level 2 in most settings.

The United Kingdom government in the year 2000 set out a comprehensive and coherent strategy to improve social care. It developed a new structure at national level built on four foundations.

The first was the regulation and inspection of all social care services. All social care services were required by law to register with a new national inspection service – which was designed to inspect all services whether provided by statutory sector, private sector or voluntary sector – non-profit organisations. National minimum standards were established for services against which they were to be inspected. The Inspection service has been modified since it was established and is due to change again. The Inspectorate is funded by government and by charges to those inspected and is semi-independent of government.

The second structure was to establish regulation and registration of staff. New bodies were established to undertake this role. Until then there was no requirement for social workers to be registered in the way that, for example, doctors, nurses and teachers are. All social workers are now registered and of course may be struck off for misconduct. The intention now is to move on from the 80 000 or so social workers in the United Kingdom to the rest of the 1.3 million workers in social care.

The third foundation was the development of an organisation to undertake workforce planning and development, what are now Sector Skills Councils.

The fourth foundation stone set up an organisation to develop a knowledge base for social care, which would provide the underpinning knowledge on which the other organisations could build. This fourth is the task for the Social Care Institute for Excellence, known as SCIE, set up in September 2001.

Stakeholders in social care

SCIE has a complicated network of stakeholders with whom it must work. Social care in the United Kingdom is devolved to the different countries – England, Wales, Northern Ireland and Scotland. SCIE has agreements with the different administrations. There are service level agreements in England, Wales and Northern Ireland and a different arrangement in Scotland.

Social care is commissioned by statutory authorities. In England, Wales and Scotland local government has responsibility for commissioning social care services. In Northern Ireland the National Health Service has that responsibility.

At one time statutory bodies were the main providers of social care but now the majority of social care is provided by organisations in the private and voluntary sector, with some statutory sector provision remaining.

Our stakeholders therefore include policy makers at government level in the different jurisdictions, and at local level. They include those who commission services (there are 150 local authorities in England, 22 in Wales and 5 boards in Northern Ireland) and those who provide services (there are some 25 000 service providers registered ranging from small local voluntary agencies to huge voluntary agencies working across the United Kingdom with thousands of staff; in the private sector ranging from a small residential home run by its owner to large private companies with multi million pound turnover).

SCIE's stakeholders also include people who use social care services and their carers. There are around 1.5 million people who use social care services each day and there are around 5 million people who provide informal care to family members and friends. Of these, 1 million provide more than 50 hours of care a week. Social care staff are also key stakeholders for SCIE.

There is the research and teaching community in social care, and finally the regulators, who are country based, not UK wide. There are different structures both for regulation of services and regulation of staff in the different countries.

SCIE's remit

Our role is to establish a knowledge base in social care, identifying and reviewing the material that constitutes that knowledge base. A parallel organisation was established two years earlier in the health service, the National Institute for Clinical Excellence (NICE), to produce guidance in health care. This was used to endorse the need for SCIE to commission its research externally rather than to develop a fully fledged research capacity in its own right. It would also ensure the full independence of SCIE's review function.

SCIE is also required to establish what works in social care. This involves reviewing practice and establishing from the knowledge base including available research, which interventions are effective. It is then our role to produce guidance for policy and practice which we must then make available as widely as possible to the social care field and support people and organisations in implementing that guidance. Our work is published in traditional form but increasingly is web based. All our publications are free including our website which does not require a password. The aim is to improve the quality of services and for that improvement to be knowledge based.

In establishing SCIE the then minister John Hutton referred to it as "the motor in the engine". It was designed from its outset to be the key source of evidence based policy for other agencies to employ in their work, a touchstone and reference point in a social care arena lacking authoritative bodies of knowledge.

In the beginning, the government considered three options. The first was to have SCIE as a part of a government department. The second option was to have it as a non-departmental public body – a sort of semi-detached organisation like the Inspection services, and the third was to establish an independent body. It chose the latter and so created a non-governmental organisation in England, a charity with independent trustees, fully government funded by means of service level agreements. It also importantly gave it a UK wide remit.

Establishing a knowledge base

One of the key challenges for SCIE is to establish the sources from which it draws knowledge.

SCIE is required to work with all its stakeholders and to do so in a policy context which is emphasising the person who uses social care services as a citizen; in a context where services are encouraged to promote, develop and enhance independence. At an early stage SCIE commissioned and published a report on "the types and quality of knowledge in social care" (Pawson *et al.*, 2003). In particular it explained SCIE's

determination that different kinds of evidence, from a range of sources are recognised, valued and built on.

This meant that SCIE had to consider what types of knowledge we could draw on and how to distinguish good quality knowledge from that which should not be relied on in policy-making and practice. Clearly we draw heavily on the work of researchers and academics involved in social work and social policy; there is a strong body of knowledge in this country and a number of high quality centres of research and teaching. However, we will later see that there is a need for greater involvement in social care research.

The inspection services are now building up a very substantial body of knowledge about the provision of social care services and have invested in the capacity to pull this information together and use the knowledge much more effectively. The Commission for Social Care Inspection (2005) for England has published a very detailed picture of the state of social care.

SCIE is particularly keen that the knowledge that is held by people who use services is included. Increasingly service user groups are demanding involvement in research production and in the United Kingdom the disability movement has led the critique of research that fails to address the need for change in the circumstances of disabled people and fails to involve disabled users. This call for a new kind of relationship between researchers and service users extends beyond the disability field. For example, Shaping Our Lives is a user led organisation working on user defined outcomes of different kinds of community care and the Toronto group is an alliance of researchers and service users established to encourage and support user involvement in research.

Social care has not been effective at capturing practitioner knowledge, nor at effectively involving practitioners in developing the knowledge base. We do not have the tradition of medicine where practitioners are encouraged to be involved in research and teaching and where joint appointments between hospital and university are commonplace. The practitioner/researcher in social care is not at all common. Practitioner knowledge tends to be personal and context specific and therefore difficult to surface and aggregate.

Achieving change

One of the key challenges for SCIE has been to establish itself as a credible source, an authoritative source of guidance. Our independent status is an asset in that respect but may be perceived as a weakness as we have no coercive power. We cannot require any organisation or any practitioner to follow our guidance. We are therefore only able to influence, persuade and support. We must work in partnership so that our work does not remain on the bookshelf or untouched on the web. Partnerships with, for example, the regulators who can use our guidance to inform the standards they will inspect against.

We have had to balance the conflicting demands of stakeholders wanting our work now and having robust quality assurance systems – so that our work has respect from the academic and research community and yet is current and answering today's problems rather than yesterday's.

We work in a political environment – our sponsor departments quite reasonably expect us to work on areas in which there is a strong political and policy interest. Currently in children's services a key issue is that of looked after children – in adult services it is the drive to integrating health and social care. Political timescales are often very short and ministers who often have a very short time in post want quick answers –

often to questions which are far more complex and do not lend themselves to quick solutions.

Absolutely critical to achieving credibility has been genuine involvement of stakeholders in all aspects of our work – so our Board of Trustees reflects the wide spread of our stakeholders from people who use social care services, to managers and academics. In all our projects we involve stakeholders in the advisory and reference groups which oversee the projects; we have a consultative group of 45 stakeholders drawn from across social care which comments on our plans and work programme; we have a network of Practice Partners – organisations which commit themselves to working with SCIE for two years to help develop our work including road testing our products before we launch them.

Examples of brokerage

The first example is our work on foster care that is, looking after children who can no longer live with their birth family. Foster care places the children with another family – it is now the placement of choice rather than residential care. First we commissioned a review of the research available which we published under the title "Fostering Success" (Wilson *et al.*, 2004). This is a scoping review providing a summary of the main trends in research rather than a comprehensive account of all the research that would be available in a systematic review. Its purpose is to alert those involved in fostering to the main messages of research.

We then commissioned a review of fostering practice which was published under the title "Innovative, tried and tested" (Sellick and Howell, 2003) because we looked for what works, whether it was new or well established. We also undertook specific pieces of work on two areas – the adoption of looked after children (Rushton, 2003) – because of a particular policy drive to increase the number of children now fostered who gain the extra security and stability offered by adoption – and then work on resilience – a key factor in children and young people's success in the face of adversity giving practitioners advice on how to build up resilience (Bostock, 2004).

All of this work was then brought together to produce a practice guide for fostering (Social Care Institute for Excellence, 2004). A guide which brings together the knowledge we have from research, the experience of service delivery, the policy and the legislation supporting that policy, into a guide which enables people working in fostering to ensure that their practice is based on the most up to date knowledge – it is a web based resource to allow for updating and development and to enable users to access it at different levels.

The guide is now referred to by the agencies responsible for inspecting foster services – so that foster care providers have a clear guide for practice against which they can be assessed. So knowledge is collected, synthesised, made available and accessible in order to improve the service offered to children and young people.

The second example is central to all SCIE's work. It's a truism, but you can't have evidence-based policy and practice without the evidence.

Our work (Marsh and Fisher, 2005) shows very clearly that the evidence-base in social care is under-developed and in need of urgent strengthening. In comparison with a health spend of 5.3% of total budget, social care spends about 0.3%. In terms of the amount spent per workforce member this translates to £25 per head in social care,

compared with £3 400 in health. If we look at more directly comparable professionals, social care spends £60 per social worker, compared with £1 466 per general practitioner.

Our work on this is an example of SCIE focusing attention on a problem, in a way that would have been difficult for central government to do or for the research sector to achieve. It may be rather obvious to point out that the research sector would have problems of perceived self-interest in calling for research investment. What may be less obvious is that central government would have (and does have) problems about being associated with a call to increase investment in social care research, particularly as it does not control much of the social care budget. Investment in research is a shared responsibility between central and local government, employers, provider agencies, higher education and so on. No-one can clearly exercise leadership in this field so it is convenient and timely for SCIE to do so.

Having placed the issue on the agenda, SCIE has now negotiated authority to undertake a consultation about ways of strengthening research and it is hoped will be able to take forward the issues arising from that consultation.

Conclusion

SCIE is still fairly young. Established in 2001, we have worked throughout with a reforming Labour administration. In one sense, the honeymoon is not yet over.

We have found, however, a strong resonance between our values and those of welfare reform, particularly where we implement a practical form of involvement that delivers the kind of personalised solutions that both government and people who use services are seeking. We might call it democratising welfare.

In pursuing these values, we have found that our power or influence is multiplied. The democratisation of welfare is often portrayed as professionals giving up power in favour of those who use services, as though power is a finite resource. In fact, we have found that sharing power creates power, adding to each other's case for change and for investment. In this sense, brokerage is a creative process, liberating energy and resources, rather than the rather bland definition of the "go-between".

However, there are significant challenges. SCIE's funding is almost entirely from central government (albeit spread over three governments). This makes us vulnerable to political winds. Although this paper endeavours to show we are solving problems for central government and therefore have a useful role, it is unlikely that this will see us through serious adversity. It is therefore vital that we extend our funding sources.

Linked to this, we also need urgently to demonstrate our impact in achieving change. The change we achieve is usually through collaboration and power-sharing, and as such it is often owned by the people we work with, rather than specifically recognised as stemming from SCIE. The active ingredient is a little difficult to detect and demonstrate.

As a first step, we have commissioned an external evaluation of the visibility and utility of our resources, and this reports in March 2007. This will be a vital part of maintaining our position in the agencies charged with improving in social care.

References

Bostock, L. (2004), *Resource Guide 4: Promoting Resilience in Fostered Children and Young People,* Social Care Institute for Excellence, London.

Commission for Social Care Inspection. (2005), *The State of Social Care in England 2004-05,* Commission for Social Care Inspection, London.

Marsh, P. and M. Fisher in collaboration with N. Mathers and S. Fish (2005), *Report 10: Developing the Evidence Base for Social Work and Social Care Practice*, Social Care Institute for Excellence, London.

Pawson, R., A. Boaz, L. Grayson, A. Long and C. Barnes (2003), *Knowledge Review 3: Types and Quality of Knowledge in Social Care*, Social Care Institute for Excellence, London.

Rushton, A. (2003), *Knowledge Review 2: The Adoption of Looked after Children: A Scoping Review of Research,* Social Care Institute for Excellence, London.

Sellick, C. and D. Howell (2003), *Knowledge Review 4: Innovative, Tried and Tested: A Review of Good Practice in Fostering*, Social Care Institute for Excellence, London.

Social Care Institute for Excellence (2004), *Practice Guide 3: Fostering*, Social Care Institute for Excellence, London.

Wilson, K., I. Sinclair, C. Taylor, A. Pithouse and C. Sellick (2004), *Knowledge Review 5: Fostering Success: An Exploration of the Research Literature in Social Care,* Social Care Institute for Excellence, London.

Part Three
Evidence-based Policy Research in Practice
Examples from the Field

Chapter 10
A Large-scale Policy Research Programme: A Canadian Experience

Satya Brink, Ph. D.
Director, National Policy Research, Learning Policy Directorate
Human Resources and Social Development Canada, Government of Canada

In this chapter we look at how a number of factors have provided momentum for a major cultural change in evidence-based policy in Canada. At the same time as longitudinal surveys and methodologies offered opportunities for research on new trajectories, a political will appeared to undertake a concerted policy programme which resulted eventually in the "Children's Agenda".

A major culture change

In the late 1980s the time was ripe for policy action on human development from childhood to adulthood in Canada. However, the evidence was inadequate for supporting expensive policy investments. Much of the research was conducted on past generations and the context had dramatically changed. Single research projects or evaluations, no matter how rigorous, were insufficient for developing a major policy initiative. Therefore, it was time to build a "body of evidence" based on the best sources of national data on the current generation of children. This was the start of a major culture change regarding not only evidence-based policy but also the use of indicators for accountability.

A number of factors aligned to provide momentum major cultural change. Research from natural as well as social science began to show that early childhood development could have impacts on outcomes later in life. Longitudinal surveys and methodologies provided opportunities for research on trajectories that could test the evidence in Canada. There was political will to undertake a concerted policy programme which resulted eventually in the "Children's Agenda".

Policy-driven research demands a long-term view based on desirable outcomes

Governments eager to show policy successes during their short elected mandates often demanded a short horizon for initiatives. The public debates, nevertheless, considered "legacy" proposals and encouraged the consideration of a significant addition to the policy infrastructure of the country, similar to public pensions. Such an investment required a long-term view and a variety of policy instruments to achieve the desired goal. This led over the next decade to the launch of the "National Children's Agenda", which

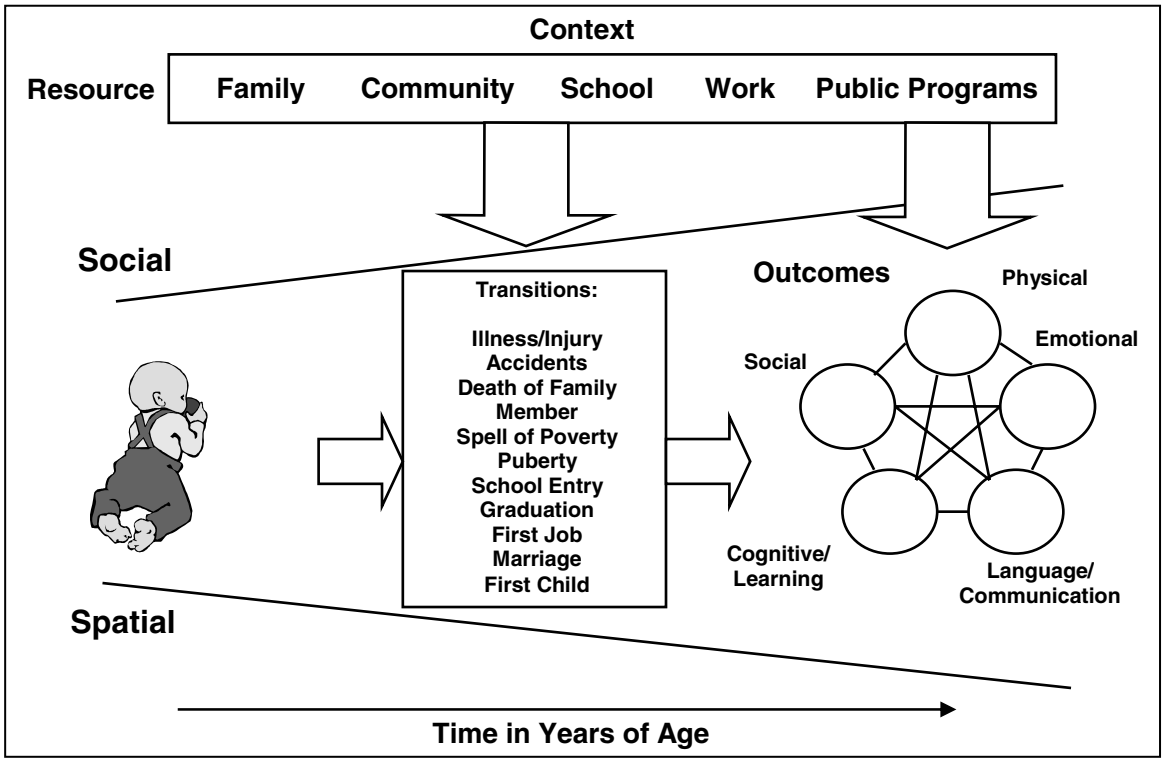
required a package of policy initiatives by two levels of governments. This article describes how this was achieved.

A major breakthrough was achieved in about 1992 when positive human development was made the desired objective rather than *ad hoc* short term goals in childhood (*e.g.* reducing bullying in schools). Human development was defined as a lifelong process by which individuals acquire knowledge, skills and individuality which they use to adapt to the changing environment and for personal and societal benefit. Such a lifelong perspective was important to explain different trajectories of individuals and to relate the differences to opportunities and experiences that could result in positive changes.

Such a broad conceptualisation of human development was directed to the national vision of fostering "good Canadians". Four final outcomes in adulthood were chosen – lifelong learners, productive workers, nurturing parents and engaged citizens.

Such a long-term and holistic view required a different conceptual framework and a major rethink of the sources of data. The conceptual framework (Figure 10.1) had to accommodate the potential developmental pathways of children to adulthood while identifying the contextual factors, resources and determinants, and life events which could affect developmental outcomes. Age appropriate outcomes at each developmental stage were studied but they were linked as a trajectory. The five developmental outcomes selected as leading to the final outcomes in adulthood, were physical development, cognitive development, emotional development, social development and communication.

Figure 10.1. Conceptual framework for data, research and policy for human development



Source: Survey Documentation, HRSDC (1998).

A better understanding of the relation between evidence and policy

Evidence-based policy enables strategic and effective policy decisions for complex, multiple and persistent problems based on reliable research evidence. Its functions are:

- to contribute to the policy debate based on evidence rather than ideology or tradition;
- to identify the seriousness of existing and emerging problems and to estimate their consequences to individuals and society;
- to determine the need and to target clientele;
- to choose between policy options based on evidence of cause and effectiveness of interventions;
- to determine the best time and the most sensitive variable for interventions and to increase the chances and durability of successful results.

The assumption that known and existing problems were serious enough to merit policy action resulted in incremental policies. Because of a high reliance on proxies or risks rather than outcomes, over or mis-targeting were common. Since the emphasis was on children at risk, problems were over-estimated, ignoring the fact that a large majority children were developing normally. Without measured outcomes it was difficult to design policy objectives and to judge their effectiveness. Evaluations were often unable to show that policy effects at the population level. In addition, emerging problems such as obesity grew to be serious issues before they were addressed.

The burden of proof demanded by the public for expenditure of tax dollars required a more rigorous approach. Results from small, single research projects were insufficient. It was no longer sufficient to examine data to see what the current state was but to develop an informed view of what might be. Scarce dollars should be used for interventions that addressed the cause rather than symptoms of the problem to increase the likelihood of problem reduction and durability of result.

Furthermore, it was necessary to build a body of evidence from multiple data sources and multiple analytical methods in order to increase the chances of successful policy investments.

Public investment in national data

In 1992 the Canadian government invested in a national data system that would provide regular, reliable and systematic flows of data. This decision immeasurably enriched the system of surveys that support human capital development. There was a momentous change in the way such data was generated. Stable funding, protected from the risk of budget cuts was assured over time and given to a policy department rather than the national statistical agency in order to ensure that surveys would be policy-driven.

With the shift of policy questions from “who?” and “how many?” to “why?”, “which?”, “how?” and “when?”, the emphasis of analysis was on explanation, size of effect and prediction. This required multi-variate analysis, forecasting models, longitudinal analysis and experiments and data that could support such analyses.

This resulted in a suite of linked surveys, some longitudinal and others cross sectional. The innovative National Longitudinal Survey of Children Youth (NLSCY),

initiated in 1992 with the first data collection in 1996, provided an overall view of the current generation of children based on the conceptual framework in Figure 10.1. A national sample of 25 000 children was surveyed every two years from birth to age 24.

The need for reliable and objective measured outcomes of performance achievement required expensive data collection at the home of the child but contributed to a bank of information on the development progress of children. Such direct assessments of outcomes enabled analysis related to children's assets as well as exposure to risk.

Two types of age-appropriate outcomes were measured:

- *Life events of milestones*: Discrete events or a process that ends with a change in status or condition.
- *Performance achievement*: Acquired knowledge, skills or attributes.

Such a longitudinal survey focused on trajectories is broad but not deep. It was described as a backbone, which was supplemented by other cross sectional surveys and community studies that provided additional in-depth information. For example, in order to understand community effects, community based studies, called Understanding the Early Years were begun. The outcomes of children in a particular community could be compared to national and provincial averages but the distribution of outcomes in the community could also be studied relative to the distribution of services. Such evidence at the community level, anchored in space through the use of maps were powerful instruments for local action integration and for planning of service delivery.

Other sources of data were also put to use. A random control experiment called the Self Sufficiency Project tested (in addition to other research objectives) whether the outcomes of children would be affected if parents were involved in welfare to work programmes. Administrative data on social services and child care were used to understand the child rearing context. Programme evaluations were scheduled.

A policy-driven consolidated policy research programme

The consolidated research programme was essential because it systematically addressed issues related to human development in order to build a body of evidence. The federal government provided research leadership through multiple activities within and outside government.

A major role was the development of key concepts and outcome indicators, in particular, a composite measure of development, the "vulnerability index". The vulnerability index is used to measure age-appropriate multiple developmental outcomes over time to gauge development. Vulnerability is defined as the occurrence of low measured current learning and behavioural outcomes that indicate a higher chance of negative outcomes later in life. Children may experience short episodes or prolonged periods of vulnerability. The developmental trajectory built using the index showed pathways of children who were resilient and those who were not after an episode of vulnerability.

While waiting for the first data release from the longitudinal survey, reviews and syntheses were conducted. Once data became available, research was generated by multiple means. The involvement of the research community was vital, resulting in the growth of interest in human development research. Policy research was largely conducted within government or by a programme of directed research contracting with researchers.

In addition, funding was provided to train young researchers in longitudinal analytical techniques, either through training or by bursaries to graduate students who would use the NLSCY for their research.

The construction of the body of evidence

There were important parallel developments that supported the construction of evidence. E.O. Wilson's concept of consilience links the results from natural sciences with those from research in the humanities (Wilson, 1998). With the advent of magnetic resonance imaging (MRI) major brain studies were conducted capturing the process of "brain sculpting" and "hard wiring" in early childhood. These findings combined with the results of research from the NLSCY began to provide powerful empirical evidence on the importance of early childhood development for success in later life.

The presentation of complex evidence required public education on the relationship between outcomes and determinants. Once the concept of outcomes was understood, information was presented on the fact that not all determinants heightened risk. Some were protective.

- *Determinants of risk* increase the probability of low outcomes, less successful trajectories.
- *Protective determinants* increase the probability of good outcomes, resilience after an episode of low outcomes and positive trajectories.

Over a six-year period, there was a continuous flow of results from high quality, peer reviewed research. The federal government published more than a hundred research reports using the NLSCY alone (see link under References). Two books were published supported by government funding, with research by multiple authors ("Growing up in Canada" and "Vulnerable Children"). Longitudinally, results distinguished between trauma effects, lagged effects and persistent effects.

A body of evidence provided an on-going view of the social and human development of individuals and society so that both preventive and corrective policies were possible. Such a body of evidence was built through multiple lines of evidence to confirm findings. Contradictory findings, on the other hand, called for further research. Moreover, syntheses and meta-analyses done by others were used to consolidate such findings.

The flow of findings using different data sources and various analyses slowly built the case for policy action. Such evidence was essential for both the generation of public acceptance of policy action and the reduction of political risk. It was essential for government to disseminate research results and the consolidated evidence using multiple formats and modes to reach a wide variety of consumers such as parliamentarians, key policy players, professionals and the general public. For instance, parliamentary committees heard evidence supporting legislation on divorce and access to children by divorced parents and grand parents.

Policy innovations driven by evidence

Some policy innovations were possible because the case built by evidence was strong. There were several ideas that have served as a model for other policy fields.

1. *Shift in policy direction*: A packaged approach benefits from synergies and possible interactions across interventions. For example, enhanced maternity leave benefits the baby and its mother in the first year while providing job protection and career retention. Classic debates between the efficiency of targeting versus the fairness of universal policies were abandoned. Rather, there was greater acceptance of hybrid approaches of targeted universalism that would support protective factors on one hand and prevent or risk factors on the other.
2. *Federal research information as a policy instrument*:
   - No other institution, other than the federal government can make the high investment required in national data, which is then made available to researchers.
   - Many professionals and institutions contribute to the welfare of children. Investments by the federal government alone were unlikely to significantly raise the outcomes of children. However, when data from multiple sources was linked and analysed in meaningful geographical context, information became a policy tool. The availability of research information to players other than the federal government resulted in more informed decision-making by them towards achieving the same policy objective. The transaction and co-ordination costs were reduced as all parties had the same information. The resulting interactions were better and therefore, there was less duplication, less unintended consequences and more effective delivery.
3. *Innovative federal-provincial collaboration*: There was a new federal provincial relationship built for the Children's Agenda. Normally, the division of jurisdictional powers prevented joint action. However, because there was such strong public support, the federal government and the provinces negotiated an innovative way to work together. The federal government agreed to transfer payments to families with children, which would result in the reduction of welfare expenditures for provinces. These savings were to be spent on services for families with children. All governments would be accountable by publishing child development outcome indicators achieved by their expenditures. This has become a model for future federal-provincial collaboration.

Concrete results on behalf of Canadian children

The mounting evidence and the power of potential analysis to determine the success of policy investments led to a string of investments. These policies work together as a package (see Table 10.1), with results from the whole being greater than the results of the individual policies.

The value of the data assets in support of both policy development as well as accountability for results has resulted in the availability of stable, sustained funding for the National Longitudinal Survey of Children and Youth. In addition, sub-surveys were added to ensure that additional data could be used for indicators and for in-depth analysis of particular issues. Currently, the children in the survey have reached the age of majority and discussions have begun to start a new cohort.

Table 10.1. The policy package for the national Children's Agenda

Initiative	Start date	Expenditures
• National Longitudinal Survey of Children and Youth	1992	$8 million a year
• National Children's Agenda	1997	
- Centres of excellence		$20 million
- Aboriginal Headstart		$66 million (2003-2004)
• Federal Provincial Early Childhood Agreement	2000	$3.2 billion (2001/2-2007/8)
• Enhanced Canada Child Tax Benefit and National Child Benefit	2000	$9.1 billon (2004-2005)
• Enhanced Maternity and Parental Benefits (10-35 weeks)	2000	$3 billion (2003-2004)
• Child Care expense Deduction	2003	$545 million (2003-2004)
• Understanding the Early Years – 100 communities	2004	$100 million (over 7 years)
• Canada Learning Bond ($3000)	2004	$85 million
• Federal Provincial Agreement on Early Childhood Learning and Care	2005	$5 billion (over 5 years)
• Children's fitness tax benefit	2007	$106 million

Tests for quality of evidence

Evidence does not come cheaply. How do we know that policies based on evidence are more effective and efficient than those that are not? Below are some suggested tests for the quality of the evidence:

1. The need for policy action test: Is public expenditure warranted? Are societal consequences serious? What are the risks of not responding collectively?
2. The reality test: Practical rather than moral imperative to move outcomes in a desired direction could challenge conventions.
3. The public good test: Can policies deliver desirable societal benefits? Can they reduce future public expenditures?
4. The value for money test: Are resulting policies cost-effective?
5. The certainty of result test: Does the policy work under varying situations? Different regions?
6. The durability of result test: Do policies spring board recipients to independence?

References

Brink, S. and S. McKellar (2000), "The National Longitudinal Survey of Children and Youth: A Unique Canadian Survey", *Isuma,* Vol. 1, No. 2, Autumn, p. 111-113 (*www.isuma.net/v01n02/brink/brink_e.shtml*).

Early Childhood Development Activities and Expenditures (2005), Government of Canada Report 2003-2004, Government of Canada, Ottawa.

Early Learning and Child Care Activities and Expenditures (2005), Government of Canada Report 2003-2004, Government of Canada, Ottawa.

Human Resources and Social Development Canada (1996), "Growing up in Canada: The National Longitudinal Survey of Children and Youth", October, Vol. 2, No. 2, HRSDC, Ottawa.

Willms, J.D. (ed.) (2002), *Vulnerable Children: Findings from Canada's National Longitudinal Survey of Children and Youth*, University of Alberta Press, Alberta.

Wilson, E.O. (1998), *Consilience – The Unity of Knowledge*, Knopf/Random House, New York.

Web links

Publications from Human Resources and Social Development Canada using the National Longitudinal Survey of Children and Youth: *www.hrsdc.gc.ca/en/cs/sp/sdc/pkrf/page00.shtml*

Research projects using the National Longitudinal Survey of Children and Youth: *www.statcan.ca/english/rdc/rdcprojectsnlscy.htm*

Chapter 11
Life as Learning – A Finnish National Research Programme

Hannele Niemi, University of Helsinki

In this chapter, we present an example of programmes inviting researchers to create new knowledge on urgent themes in society: the Finnish national research programme Life as Learning *(LEARN), launched by the Academy of Finland for the years 2002-2006.*

The Finnish education system has received attention from all over the world because it came out on top in the first two PISA surveys. Finnish 15-year-olds are number one in terms of skills in mathematics, scientific knowledge, the reading of literature and problem-solving (OECD, 2001, 2004), and only a very few students fall within the lowest PISA categories. Likewise, differences between schools are small. PISA shows that Finland has succeeded in its policy to enhance the equity and quality of learning. It has been a long process, and the long-term development objectives were set almost 40 years ago.

According to researchers (Välijärvi, 2004; Simola 2005; Laukkanen 2006; Niemi and Jakku-Sihvonen 2006), the educational policy has purposefully aimed at equity in education and promoted the common comprehensive school model. In the process, many important decisions have been made, *e.g.* the discontinuation of streaming, the strong allocation of affordable educational resources to lower secondary education and the decentralisation of decision-making powers. Primary school teacher education was also raised to the MA level. Support for weak students was taken care of. Different stakeholders have been invited to express their opinions.

The Ministry of Education and researchers of education have been closely collaborating in promoting a common comprehensive school for all. The review group for educational sciences (Educational Research in Finland, 1990, pp. 2-3) assessed that during the last few decades, most of the researchers and professors in education have been working hard towards developing the educational system and teaching arrangements. The research that academic experts have conducted with their colleagues and students at universities has been important input. This has promoted evidence-based policy-making, which has helped to develop the education system in Finland.

Life as Learning – The Finnish case of a national research programme

Important tools for evidence-based policy-making are research programmes which invite researchers to create new knowledge on urgent themes in society. A current

example is the national research programme *Life as Learning* (LEARN) launched by the Academy of Finland for the years 2002-2006 (*www.aka.fi/learn*).

The initiative to establish the programme came from many public and private sector partners in 1999. In the 1990s, Finland had defined itself as a knowledge-based society, and the significance of learning became more urgent than ever. The National Board of Education, the Ministry of Education, the Future Committee of the Parliament and Nokia Corporation took the initiative to contact the Academy of Finland. They emphasised that investment in research on learning and especially in basic research is a key factor when promoting the ideal of a learning society.

In 2000 the Academy of Finland set up a preparatory group to explore the main themes of the programme. The preparatory group organised national seminars and meetings for researchers. It also arranged an international workshop to which key persons from ongoing national research programmes of learning in the United Kingdom (*www.tlrp.org*) and in Norway were invited. Based on what was learned from these events, the Academy of Finland set objectives and themes for the research programme and released a call for proposals in order to start the programme in 2002 (*www.aka.fi/learn*).

Research areas that urgently required new approaches and findings were chosen as the themes. The programme was to focus on redefining the concept of learning and examining social and cultural contexts of learning, knowledge creation, working environments, and new teachership. The preparatory group emphasised the importance of ICT in teaching and learning, but the hope was that this topic would be integrated into the development of teaching and learning environments.

The objectives of the research programme reflect the principles of multidisciplinary, cross-boundary partnerships and the anticipation of the future. The programme aimed to:

- encourage the development of a new research culture and new research partnerships and the creation of interdisciplinary and international research projects around the problems of learning;
- find a way of managing the challenges of lifelong and lifewide learning in order to avoid new forms of exclusion;
- create a solid interdisciplinary research base for developing teaching and learning in different educational and working-life contexts; and
- anticipate future learning needs from the point of view of society, culture and the individual.

The review process had a phase for outlines (116 proposals) and one for full proposals. After the international review the Academy of Finland selected 17 projects with three large consortiums. The acceptance rate of outline proposals was only 15%. An open call for programme co-ordination was also held. The Academy of Finland selected the University of Helsinki as a co-ordination unit, Professor Hannele Niemi as Scientific Director, and Researcher Raija Latva-Karjanmaa as Co-ordinator.

The Academy of Finland was the main funding agency (5.1m euros) of the programme, although Tekes, the Finnish Funding Agency for Technology and Innovation, and The Finnish Work Environment Fund also supported projects in their own focus areas. The Ministry of Education, the National Board of Education, the Centre of Expertise Programme within the Helsinki Region – Culminatum Ltd – and the

University of Helsinki provided the programme with resources for co-ordination activities, seminars and dissemination work.

The accepted projects extended to a broad scale of different disciplines: education, psychology, sociology, technology and engineering, neurology, and economics. The projects also covered a large variety of contexts of human learning in educational institutions and working life as well as non-formal learning settings and virtual learning environments. Learning was approached at individual and collaborative levels. Many projects were collaborative plans connecting researchers in several universities, and some projects also involved business partners. Some projects had an ambitious aim to create new tools for managing changing contexts or to change practices. Most projects also had strong international links.

Co-operation and dissemination throughout the programme

One aim of the programme was to create a new research culture and strengthen co-operation between different disciplines and partners. To increase cohesiveness and mutual interaction within the programme, joint meetings, conferences and social events were arranged each year. The co-ordination unit also arranged forums for researcher training, facilitated joint article and book writings in cross-over projects, and organised researcher meetings with other research programmes close to Life as Learning. Contacts with other European programmes were also offered.

A component of dissemination and partnership with practitioners and policy makers was included from the beginning of the programme. Even at the beginning of the LEARN programme, the projects already had well-grounded frameworks and foundations for their new projects and could offer important scenarios to urgent issues of learning. The national conferences "School and Teachers as Developers of Learning Environments" (2003 and 2005) were designed for teachers, headmasters, teacher educators and key persons in school policy. The conferences "Changing Working Life Contexts" (2004) and "The Social Innovations in Working Life" (2004) were organised in order to create contacts between learning researchers, enterprise and the public sector. These conferences also provided interactive sessions and discussions where practitioners could give their contributions and initiatives. The feedback from the participants was very rewarding. What teachers and practitioners were found to value was that high level researchers informed them about the projects of the Academy of Finland and that they had an opportunity to be partners in the programme.

Life as Learning has published one to three newsletters each year, LEARN periodicals (in Finnish and in English; see *www.aka.fi/learn*), in which the latest news from the projects and conferences have been introduced. Two special issues have been sent to all schools and working-life partners. The national TV and media have been actively involved in the programme's work. Four TV documentaries were released in 2005 and are to be used by digital TV for later broadcasts as well.

The programme has organised two international multidisciplinary conferences, INTERLEARN 2003 and 2005. Both conferences had 200-300 participants, of whom one-fourth were researchers from other countries. Many teachers and practitioners also participated in these conferences.

The co-ordination unit together with researchers drew up a publishing plan at a very early stage of the programme. The programme had a double strategy. One aim was to

publish high quality scientific articles in international scientific refereed journals. Joint writing groups and editorial teams for a special volume of journals were set. Another aim was to produce books and articles for Finnish society. The co-ordination unit arranged negotiations with national publishing companies. As a result, an agreement was made with a company specialised in teaching and learning issues. Further, some universities offered their own publication series for the publishing of new findings. Fours joint books (*www.aka.fi/learn*) have been already published for revising learning in schools and working places, even though the programme has not yet finished.

The dissemination and knowledge transfer has been an ongoing process. It has caused additional work for researchers, but it has also added to their motivation when they have understood the significance of their work to societal partners. This work has also been a learning experience for academic people, and sometimes finding the right way to communicate with practitioners has been difficult.

Strengths and challenges of the programme

The strengths of the project have been in promoting multidisciplinary approaches and cross-boundary co-operation and learning in different learning contexts. Even though the programme has not officially finished, we can see some promising tentative results.

Different generations of learners

Finnish society has different learning generations at the same time. Their needs and capacity to learn new skills vary a great deal. They all have different conceptions of knowledge and learning. How to help different learners to learn new ways of working is a big challenge. The younger generation also must prepare itself to learn several, perhaps three to four vocations or professions in their life course. This sets more and more emphasis on learning to learn skills in schools (Olkinuora and Rinne, 2005). In the area of vocational education new demands are to face occupational de-specialisation, multi-skills and knowledge work, simultaneous and contradictory processes of individualism and a new kind of collectivism at the same time (Heiskanen, 2005).

Learning to learn is decisive

Learning to learn has cognitive and emotional components. It is important how learners see the future and how they conceptualise themselves as learners. They must be capable of adjusting themselves to new environments. Students need a sense of hope in their learning, and motivation plays a key role in their learning paths. Learners use different strategic ways to influence their motivational orientations, and they also need strategies to cope with stress in schools and working life. Learning is more and more about sharing and being connected with networks, and people need collaborative skills when working in these environments. Multicultural groups set special demands on collaboration (Hautamäki *et al.*, 2006; Nurmi *et al.*, 2003; Pitkänen, 2003).

We can have an effect on exclusion

International comparisons (OECD, 2001, 2004) demonstrate that Finnish students have a high level of school achievements and also a high score in learning to learn skills. However, students at risk need special support in their learning. We have evidence that

we can prevent exclusion by allocating for special education and arranging flexible educational structures which give opportunities to continue learning and schooling at any phase of one's life course. We have interesting cases of how the early identification of signs of exclusion is important. An influential factor is the learners' own concept of themselves as learners. At an individual level it is also of great importance if someone (*e.g.* a tutor, friend, or teacher) gives a supportive impulse to a student who is in danger of being excluded. This support can very often happen in a very unofficial way (Martti, 2005; Suikkanen, 2005).

Learning and new technology

Web-based learning changes internal as well as external processes of organisations in knowledge creation. We need to model these processes and make them visible. Earlier roles of teachers and students have changed. In web-based learning there is a need to create richer symbol systems than only textual ones and pay more attention to emotions. Mobile technology opens new possibilities of enlarging learning spaces, and we have many new pedagogical applications available. The important message from the studies of implementations of new technologies is that unaccomplished, untested and unfinished environments estrange users. It may take a long time before they start to try again (Multisilta *et al.*, 2005; Paajanen and Multisilta, 2005). Some projects found evidence that the collaborative development and training simulations created a shared understanding of web-based teaching and studying as a collaborative process and helped to define the new roles and practices of the actors (Smeds, 2005).

New structures in working life

The change in key processes in work, from mass production to customer-intelligent services with mutual co-operation, creates totally new landscapes of learning. Team structures on working life are moving towards more flexible arrangements that can better serve the needs of customers. The nature of work is collaborative, multi-professional, and multicultural, requiring high problem-solving skills and continued learning (Engeström, 2005).

The LEARN programme covered several important areas of learning, but many urgent issues still demand more and deeper analysis of learning processes. We would need much more knowledge about the connections and relationships between processes and learning outcomes. A neuroscience perspective on processes and outcomes could open new ways to support different learners. We would need more research on methodological issues in multidisciplinary projects and how to combine individual processes and societal structures. In the economics of education, new models and more detailed analysis of effects on the investment of learning should be developed. Conceptual, philosophical and value analysis of learning is also needed when promoting learning in different environments and through different methods.

How to add additional value to the programme

Even though Life as Learning has been very successful in many respects, some issues have been threats to its effectiveness. Limited funding resources caused many restrictions on the original objectives of the projects. The Academy of Finland was forced to reduce the budgets of the proposals radically when accepting them into the programme. The

funding was available for four years, which seems to be too short a time for ambitious projects.

To add more effectiveness, the call for proposals could put more emphasis on joint activities and dissemination. This would steer researchers to include these aspects as real components of the projects. At this time it is very much the co-ordination unit's responsibility to get the researcher involved in joint activities during the programme. Joint seminars, meetings and other partnership activities also require resources, *e.g.* travel, accommodation and rent. The project funding mainly only covered the researchers' and the co-ordination unit's salaries. The co-ordination unit had to seek external funding for all dissemination seminars, workshops, and co-operation with stakeholders. Fortunately most co-funders provided some extra resources for these activities case by case, but the lack of funding created uncertainties in long-term planning.

The real gap in dissemination will be seen after the programme is over. The co-ordination ended in early 2006, and the research groups will finish their projects in 2006. Thereafter, neither forums nor resources will disseminate the primary results. The reality is that the best fruits of the programme could be gathered after the programme, but no organised way is available to create this additional value. How to bring major findings to the public awareness and how to inform stakeholders after the funding period is a challenge.

The new initiatives – next steps after the programme

The major funding agencies, the Academy of Finland and the National Technology Agency TEKES, published a document, FINNSIGHT 2015, in June 2006 (*www.finnsight2015.fi/*). It is a proactive national programme for innovation and competence for Finnish research policy. It is based on the work of ten expert panels, which identified the major urgent research fields in Finland: (1) learning and to renew society through learning, (2) services and innovations in services, (3) welfare and health, (4) the environment and energy, (5) infrastructures and security, (6) bio-competence and bio-society, (7) information and communication, (8) understanding and human interaction, (9) materials and (10) global business. Learning was set among the first priorities, with stress on the urgent need to discover how to facilitate the learning of various learner groups and to help organisations create fruitful environments for innovations and competence building. The main themes focus on a combination of Learning – Brain – Technology.

When drawing scenarios and future directions for research on learning, the FINNSIGHT 2015 report often refers to the network of CICERO Learning. This research network which stands for **C**ross-disciplinary **I**nitiative for **C**ollaborative **E**fforts of **R**esearch **O**n Learning, was established in 2005 by the University of Helsinki to promote multi-disciplinary research on learning. To a large degree, the Life as Learning research programme also contributed to the emergence of the CICERO Learning network. The network's new co-ordination unit is located on the premises of the University of Helsinki (for more information visit *www.helsinki.fi/cicero*).

The CICERO Learning network focuses on promoting nationally and internationally recognised high-level research on learning over the boundaries between different scientific fields, universities and corporate lives. The core areas of research of the

network are: (a) learning and the brain, (b) learning throughout life and in different contexts, (c) technologies of learning, and (d) learning and society.

Through conducting cutting edge research on learning, CICERO Learning aims at innovations and synergies between the research community, business and industry. An important goal of the network is to maintain Finland's position as one of the leading countries in innovations and industrial development.

As the new initiatives as well as the outcomes of the Life as Learning research programme demonstrate, promoting learning research is a long-term process, and knowledge-based societies urgently need learning research. Without systematic high-level basic research we cannot answer to those challenges.

References

Educational Research in Finland (1990), "Academy of Finland. Publication 1/90", Valtion painatuskeskus (in Finnish), Helsinki.

Engeström, Y. (2005), "Refining the Theory of Expansive Learning: Lessons from Bumpy Roads towards Co-configuration. Programme and Abstracts. Interlearn – Multidisciplinary Approaches to Learning", Conference, University of Helsinki, Life as Learning Research Programme, 55, 1-2 December, Finlandia Hall, Finland.

Hautamäki, J., S. Kupiainen, P. Arinen, A. Hautamäki, M. Niemivirta, P. Rantanen and P. Scheinin (2006), "Learning-to-learn Assessment in Finland – Versatile Tools to Monitor and Improve Effectiveness and Equity of the Educational System", in R. Jakku-Sihvonen and H. Niemi (eds.), *Research-based Teacher Education in Finland – Reflections by Finnish Teacher Educators*, Finnish Educational Research Association Turku, pp. 189-202.

Heiskanen, T. (2005), "Strategies for Cooperation and Learning in Network of Public and Private Sectors", *LEARN Periodical 2005*, University of Helsinki, pp. 16-17.

Laukkanen, R. (2006), "Finnish Strategy for High-level Education for All", paper presented at the conference "Educational Systems and the Challenge of Improving Results", University of Lausanne, Lausanne, Switzerland, 15-16 September.

Martti. S. (2005), "Struggling against Marginalization", *LEARN Periodical 2005*, University of Helsinki, pp. 16-17.

Multisilta, J., R. Koskimaa, H. Ruokamo and S. Tella (2005), "The MOMENTS Project (2002-2006) – Major Outcomes", *LEARN Periodical 2005*, University of Helsinki, pp. 18-19.

Niemi, H. and R. Jakku-Sihvonen (2006), "Research-based Teacher Education", in R. Jakku-Sihvonen and H. Niemi (eds.), *Research-based Teacher Education in Finland – Reflections by Finnish Teacher Educators*, Finnish Educational Research Association 2006, Turku, pp. 31-50.

Nurmi, J., K. Lonka, M. Niemivirta, K. Salmela-Aro and J. Vuori (2003), "Motivation and Self-regulation in New Learning and Teaching Environments", *LEARN Periodical 2003*, University of Helsinki, p. 5.

OECD (2001), *Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assessment (PISA)*, OECD, Paris.

OECD (2004), *Learning for Tomorrow's World. First Results from PISA 2003*, OECD, Paris.

Olkinuora, E. and R. Rinne (2005), "From Teaching Society to Learning Society? The Varying Meaning of Lifelong Learning among Different Age Groups and Generations", *LEARN Periodical 2005*, University of Helsinki, pp. 16-17.

Paajanen, M. and J. Multisilta (2005), "Productivity of Knowledge Work: Case Study of On-the-job Learning in Knowledge Work", in S. Tella, H. Ruokamo, J. Multisilta and R. Smeds (eds.), *Teaching, Studying, Learning. Information and Communication Technologies in Transdisciplinary Contexts* (in Finnish, abstracts in English), Rovaniemi, University of Lapland.

Pitkänen, P. (2003), "Learning Intercultural Competency in the Workplace," *LEARN Periodical 2003*, University of Helsinki, p. 5.

Simola, H. (2005), "The Finnish Miracle of PISA: Historical and Sociological Remarks on Teaching and Teacher Education", *Comparative Education*, Vol. 41(4), pp. 455-470.

Smeds, R. (2005), "Co-evolution Paths in E-learning", *LEARN Periodical 2005*, University of Helsinki, pp. 6-7.

Suikkanen, A. (2005), "Marginalisation", *LEARN Periodical 2005*, University of Helsinki, p. 13.

Välijärvi, J. (2004), "The System and How Does it Work – some Curricular and Pedagogical Characteristics of the Finnish Comprehensive Schools", *Educational Journal*, The Chinese University of Hong Kong, Vol. 31, No. 2, 2003 & Vol. 32, No. 1, 2004, pp. 31-55.

Web links

www.aka.fi/learn: Life as Learning Research Programme, Academy of Finland, Finland.

www.finnsight2015.fi/: FinnSight 2015: Exploring the Outlook for Science, Technology and Society. Academy of Finland and the National Technology Agency Tekes.

www.helsinki.fi/cicero: CICERO Learning. Multidisciplinary Research Network.

www.tlrp.org: Teaching and Learning Research Program (TLRP), Economic and Social Research Council, Great Britain.

Chapter 12
The United Kingdom's Teaching and Learning Research Programme[1]

Andrew Pollard, Director, TLRP

In this chapter, we look at the United Kingdom's Teaching and Learning Research Programme (TLRP), which aims to contribute new knowledge for the improvement of learning and aspires to improve the quality of the educational research that will be available in the future.

The Teaching and Learning Research Programme (TLRP) is the Economic and Social Research Council's largest research programme and provides coordination for over 500 researchers in some 60 project teams and over 20 initiatives of cross-programme thematic analysis across the United Kingdom. The first projects began empirical work in 2000 and the last project is presently expected to end in 2011. The total budget in the autumn of 2006 was some £37m and drew contributions from a wide range of UK government bodies.

The origins of the Programme can be traced to the mid-1990s when educational research was heavily criticised for being small scale, irrelevant, inaccessible and low quality. Whether or not these criticisms were entirely justified, researchers faced major challenges in demonstrating the value of investment in this field. Fortunately, thanks initially to the imagination and commitment of the Higher Education Funding Council for England, a new opportunity was created though TLRP.

TLRP's overarching strategy has been to support research which is of both high quality in social scientific terms and of high relevance in terms of policy and practice – to satisfy the criteria of "Pasteur's Quadrant". At the same time, considerable effort has gone into impact work, capacity-building across the field of educational research and in "bridging" between the worlds of researchers, policy makers and practitioners. Over time, we believe that a greater respect for, and understanding of, the complementary forms of expertise has been developing.

[1] TLRP researchers are extremely interested in sharing experiences and learning from those engaged in similar work in other countries internationally. Please see *www.tlrp.org/international* for details of some links and get in touch if you would like to develop an association with the Programme. In the first instance, please contact: Alan Brown, Associate Director, TRLP at alan.brown@warwick.ac.uk. Andrew Pollard can be contacted at a.pollard@ioe.ac.uk. A more extensive account of TLRP strategies and of some of the challenges and opportunities with which it has engaged is available at: *www.tlrp.org/dspace/handle/123456789/380*. This is the text of the Annual Educational Review Guest Lecture, 2005, by Andrew Pollard.

The success of the overall initiative is indicated by the steady growth of the TLRP budget (now almost four times greater than the initial figure), by the model of user-researcher collaboration set by the Programme, the eagerness of researchers to become involved and by the take up of findings by practitioners and public bodies. A recent independent review of TLRP concluded that the quality and relevance of the research was high (see *www.tlrp.org/manage/documents/NFER_Final_TLRP_Report_March_2005.pdf*). There are no grounds for complacency though and "evidence-informed policy and practice" remains beset by challenges. We do, however, feel that we have been making some progress.

Aims

TLRP's aims emphasise the positive contribution being made by research on teaching and learning. More specifically, they are:

- *Learning*: TLRP aims to improve outcomes for learners of all ages in teaching and learning contexts across the United Kingdom.
- *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. The Programme is concerned with patterns of success and difference, inclusion and exclusion through the lifecourse.
- *Enrichment*: TLRP commits to user engagement at all stages of research. The Programme promotes research on teaching and learning 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. This work is the particular focus of the Programme's research capacity building strategy.
- *Improvement*: TLRP develops the knowledge base on teaching and learning and contributes to the improvement of policy and practice in the United Kingdom. The Programme works to maximise the impact of its research.

TLRP's overall development is driven by six key strategic commitments:

- User engagement for relevance and quality.
- Knowledge generation by project teams.
- Knowledge synthesis through thematic activities.
- Knowledge transformation for impact.
- Capacity-building for professional development.
- Partnerships for sustainability.

Because of the duration, scale and complexity of TLRP, these elements are managed simultaneously – for instance, with some projects being commissioned just as others

complete. However, as the Programme matures, there is also a progressive change in the balance of activity, with more emphasis being placed on knowledge synthesis, transformation and impact. Explicit strategies to underpin post-Programme *sustainability* are being developed in relation to capacity-building, the use of ICT to support research development in the field and the deepening of partnerships with cognate bodies.

User engagement for relevance and quality (*www.tlrp.org/users*)

Project teams work closely with practitioners and others in their research sites and also to link up with key national organisations with potentially "high leverage" for dissemination and impact activity. Such relationships are reflected in the membership of project "Advisory Groups".

The Directors' Team maintain links with high-leverage user organisations in each educational sector and in each part of the United Kingdom. TLRP also works directly with governments in each part of the United Kingdom to maximise the use of its research. TLRP has been represented by the Director on significant national bodies for the coordination of education research in Wales (Education and Training Research Liaison Committee of the National Assembly for Wales); England (National Education Research Forum and the Department for Education and Skills Schools Research Advisory Group); Scotland (Management Committee of the Applied Educational Research Scheme). The Programme has also sustained excellent links with senior government officials in Northern Ireland and presented at the 2005 conference on the restructuring of teacher education.

Knowledge generation by project teams (*www.tlrp.org/proj/index.html*)

In 2000, TLRP started by funding four networks of projects. A second phase brought in nine larger projects and this was followed by funding of twelve more. At the same time, focused funding initiatives have made specific provision for teams in Scotland, Northern Ireland and Wales – and for some high priority topics (such as widening participation in higher education and concerning technology enhanced learning). Additionally, there have been five different types of investment in capacity-building – ranging from fellowships, training to e-resource development. With researchers from a large number of universities actively involved, the responsibility is devolved but the Directors' Team provides critical friendship and encourages collaboration across projects.

Knowledge synthesis through thematic activities (*www.tlrp.org/themes*)

The Programme's strategy for thematic development is a major focus of work as the initiative matures. The portfolio of initiatives to add value through cross-Programme analysis includes: consultancies, thematic groups, thematic seminar series, conferences, workshops, thematic meta-tagging of outputs and sectoral reviews. A conceptual framework is used to organise and integrate this work.

Knowledge transformation for impact (*www.tlrp.org/pub/index.html*)

TLRP's impact strategy is a multi-level one, in which we try to produce research findings in forms which are tailored to specific audiences. We both produce many of our

own publications and also work extensively with user bodies to maximise impact. Outputs include "Research Briefings" (summarising findings), "TLRP Commentaries" (applying findings to contemporary issues), practitioner applications (classroom activities drawing on research), books (in two series with Routledge), journals (including many special issues), reports, etc. TLRP also uses an electronic repository and meta-tagging system called D-space. This has been adopted for deposition of all project publications and has significantly improved the availability of outputs via the Internet and major databases. Downloads from the website (*www.tlrp.org*) have considerably exceeded our expectations, with several TLRP Commentaries on public policy leading the way.

Such Commentaries are an interesting example of our approach. They are attractively designed in a glossy, four-colour A4 format and are designed to provide a concise and accessible evidence-informed commentary on a contemporary, up-to-the-minute issue. The timescale of this is important, for most research takes years from commissioning to publication. We wanted to find a "rapid response" vehicle which could present the unique insights of researchers into public discussion as it happens in our fast-moving democracy. Each Commentary is thus the product of a group of researchers and users working together in a very short timescale with editors, designers, printers and distributors standing by. Documents on Personalised Learning, Teaching and Learning, Science Education have been produced in this way and another on Neuroscience and Education is in production. Although such documents gloss the detail of the evidence-base, we have found them to have a considerable impact in highlighting key issues. For example, the English Department for Education and Science has recently published a document, "2020 Vision", setting out a new approach to the personalisation of teaching and learning in schools which is very consistent with the directions indicated by TLRP research. We certainly cannot claim cause and effect but it is evident that the contributions which the research programme makes have been considered seriously and many of the key themes which TLRP has been working on are reflected in the report – for instance, on pupil voice, learning how to learn, engaging parents and carers, new technologies, professional learning and researcher-practitioner partnerships.

TLRP's commitment to transforming and disseminating research knowledge in partnership with other bodies is an extremely important foundation for all our impact work. In particular, we have identified a small number of high-leverage user organisations in each educational sector and in each part of the United Kingdom. Key users, such as the General Teaching Council (England), the National College for School Leadership, the Learning and Skills Development Agency, the National Institute for Adult Continuing Education and Department for Education and Skills, have been extremely helpful in the dissemination of results through their communication systems. TLRP also works directly with governments in each part of the United Kingdom to maximise the use of its research. For instance, the Programme held its 2006 Annual Conference in partnership with policy makers in Scotland and this is to be followed up with the placement of Research Fellows bridging government departments and the Programme itself.

Capacity-building for professional development (*www.tlrp.org/capacity*)

Capacity-building is an intrinsic part of TLRP's work. Indeed, in all phases of TLRP funding, it has been a criterion for project selection, and this work is supported, monitored and reported on each year. Particular attention has been paid to skill and career development of contract research staff, with special events each year. Additionally, with

support from the Department for Education and Skills, TLRP has funded five Research Training Fellowships, which enable senior practitioners to study part-time for PhD's in association with TLRP projects.

From 2002-2005 the Programme's Research Capacity Building Network (RCBN) provided cross-Programme training services in the research methods which were felt to be particularly appropriate in the study of teaching and learning. Each included: research design issues, the use of large-scale data-sets, and the combination of quantitative and qualitative methods. RCBN also initiated a journal, Research Capacity Building. A new strategy for capacity-building provision has been adopted for 2005-2008. This is based on an explicit attempt to embed processes for the development of research expertise within the social practices of educational researchers and is intended to complement other provision through the National Centre for Research Methods and Research Methods Programme. Additionally, there are close working relationships with the Applied Research in Education Scheme in Scotland and with relevant UK learned societies, such as the British Educational Research Association. A particular feature of TLRP's new provision is the promotion of a range of e-resources for research training which will be freely available to institutions and groups across the United Kingdom.

In short, the objective of TLRP's capacity-building work is to work *with* the academic community and to support the development of new forms of commitment and provision for the professional self-improvement of educational researchers.

Partnerships for sustainability

Despite its size, TLRP is still small in relation to the challenge and range educational research. It is also only expected to exist for a limited period. For such reasons, we have sought to develop close working relationships with other organisations. A least five different forms of partnership can be identified.

First, we seek expertise from which we can benefit. Such partnerships exist with the British Education Index (BEI) for electronic knowledge management, the Cambridge Centre for Applied Research in Education Technology (CARET) for development of an advanced ICT infrastructure, and Routledge for book and journal publications. Second, as indicated above, we work with key user organisations which are generous enough to promote TLRP work through their communications systems, and thus lend us some of their leverage as we attempt to maximise impact. Third, we work with partners where cooperation enables us to be more effective – for example, a recent TES special supplement was co-funded with National Institute of Adult Continuing Education (NIACE) and Learning and Skills Development Agency (LSDA). Indeed, TLRP aspires to contribute to a series of sectoral reviews, developed in partnership with others. For example TLRP contributes to the Nuffield Foundation's 14-19 Review, and Esmee Fairbairn Trust's Primary Education. Similar work is undertaken with the Institute for Employment Research on work-based learning. Fourth, as indicated previously, we work where we can with the government bodies which help to form policy regarding education research within each country of the United Kingdom.

For some years too, TLRP has contributed to the work of the National Education Research Forum (NERF), through the membership of first Charles Desforges and, more recently, myself. This has been valuable both in the expression of support and reservation for particular initiatives, as appropriate. For example, TLRP has contributed actively to

discussions and development work on the establishment of a National Education Evidence Portal (NEEP).

Finally, we work with organisations which may, in due course, take on some of the resources, assets or commitments of TLRP into the future. Indeed, one of our informal goals is to "give everything away" by the end. The most important legacy organisations are seen as being among the professional research associations – and, in particular the British Educational Research Association (BERA). TLRP participants are active in membership and a number of key positions are held by colleagues who bridge both organisations. There are many areas on which cooperation is developing, including joint capacity-building activities and the possibility of eventual transfer of tools and components from of TLRP's IT infrastructure. On the capacity-building front, we are also very pleased to be working with the research committee of the University Council for the Education of Teachers (UCET). Other associations with which we have collaborated closely include the Society for Research in Higher Education (SRHE), with major contributions being made to their annual conferences.

Conclusion

TLRP represents a major opportunity for UK educational research. It aims to contribute new knowledge for the improvement of learning – but it also aspires to improve the quality of the educational research which will be available in the future.

We feel we have some successes – but we also experience many challenges and frustrations. For example, with such a wide range of aspirations, we find it impossible to follow up on all the opportunities which present themselves. In such a complex society too, it is extremely difficult to know how to focus our limited resources and track impact with precision. Most fundamentally, we struggle between the assumptive worlds of researchers, policy makers and practitioners in the knowledge of considerable differences in their daily experiences, cultures, priorities, accountabilities and incentivisation systems. Sometimes, things get lost and, to our eyes, opportunities are missed.

Overall though, TLRP has provided an opportunity for educational researchers in the United Kingdom to regroup after the critiques of the mid-1990s. The work produced now is generally well received and we believe that the research community is justified in being more confident in facing the future (though the plight of teacher-educators engaged in research is a rather different story). The policy climate is more open than it once was and, even allowing for some selectivity, there is greater respect for evidence. We feel that we have both benefited from such developments and contributed to them.

The Programme has recently been funded to take forward a more focused group of projects on Technology Enhanced Learning to 2011 and beyond, but its broad portfolio will end by mid-2008. Books and other publications will continue to emerge for some years. Additionally, UK leadership in knowledge creation in education will to be taken up by ESRC investment in a small number of Research Centres on more focused topics, whilst the capacity-building role of TLRP will be taken forward through the National Centre for Research Methods or other specific initiatives. The Programme will, therefore, pass the educational baton on to new forms of research organisation. We hope that these new research teams will benefit from a climate in which educational research is seen as an important contributor to high quality policy-making and professional practice in an open, democratic society.

Chapter 13
Policy-driven Research and Evidence-based Educational Innovation in Singapore

Professor David Hogan, Dean, Centre for Research in Pedagogy and Practice,
National Institute of Education, Nanyang Technological University, Singapore[1]

In this chapter, we describe the way Singapore is pursuing the objective of promoting evidence-based policy and planning in order to comply with the vision of "a nation of thinking and committed citizens capable of meeting the challenges of the future, and an education system geared to the needs of the 21st century".

Context

In 1965 Singapore achieved independence as a postcolonial nation state, but it was more state than nation. In the 42 years since, Singapore has undertaken a distinctive and remarkable successful programme of national development, becoming not only an economic powerhouse in the Asian region, but an influential, prosperous, orderly, cohesive, multi-racial, global city and nation-state. In this endeavour, education has played a pivotal part. From the beginning, the state provided a free, well-funded universal system of public education: currently, education accounts for 4% of Singapore's GDP. In the same year, secondary schools had a retention rate of 95%. Between 1970 and 2004 literacy rates jumped from 68.9% to 94.2%; during the same period, the percentage of university graduates in the population increased from 1.9% to 12.1%. These achievements are also evident in exceptional performance in international assessments in Mathematics and Science. In the Trends in International Mathematics and Science Study (TIMSS) assessment for example, 4th and 8th grade students from Singapore consistently scored in the top place in Mathematics in 1995, 1999 and again 2003. In Science, 4th grade students were 7th in 1995 and 1st in 2003, while 8th grade students were 1st in 1995, 2nd in 1999 and 1st in 2003.

The recession of the mid-1980s made it very evident that the global economy was changing rapidly and the only way for Singapore to continue growing its economy, especially under the threat of equally attractive low-cost labour in other parts of the region, was to both upgrade its existing labour force and prepare a future labour force that is well-equipped to meet the challenges of a New Economy. Although the discourse

[1] The author wishes to thank Professor Gopinathan of CRPP/NIE for his helpful comments on the many papers that provide the (unseen) background for this paper. The views expressed in the paper are the author's and the author's alone and have no official CRPP or NIE status.

around "knowledge-based economies" (KBE) and globalisation was not widely established then, Singapore was, in many ways, one of the first global cities and a midwife of the KBE – an economy where knowledge is constantly created and exchanged and production and services are based on knowledge-intensive activities.

Since the influential report of the Economic Committee (1986), *The Singapore Economy: New Directions*, which highlighted the need for creativity and broad-based holistic education to provide sufficient skill base for Singapore to move up the economic ladder into higher value industries such as high technology-based manufacturing, financial, banking and service sectors, policy makers in Singapore have wrestled with how to produce the kind of workers that would thrive in a KBE. Over the years, the accelerating pace of globalisation and criticality of graduating Singapore into a knowledge-based economy have brought together high-level committees, including the *Committee on Singapore's Competitiveness* (1998) and the *Economic Review Committee* (ERC, 2003) chaired by the current Prime Minister, to evaluate and make recommendations on critical issues that bear on Singapore's continued economic prosperity. These issues include the organisation of work associated with the knowledge economy, the changing capital formation requirements for the knowledge economy (to wit, "knowledge" capital, "imagination" capital, "emotional" capital, and "social" capital), and the growing inequality associated with the growth of such an economy (Brown and Lauder, 2003). In general terms, the ERC committed Singapore to the following macroeconomic policy settings:

- a *globalised economy* where Singapore is the key node in the global network, linked to all the major economies;
- a *creative and entrepreneurial nation* willing to take risks to create fresh businesses and blaze new paths to success; and
- a *diversified economy* powered by the twin engines of manufacturing and services, where vibrant Singapore companies complement multinational corporations (MNCs), and new start-ups co-exist with traditional businesses exploiting new and innovative ideas.

The Ministry of Education (MOE), too, has been strongly committed to the development of an education system that prepares young people for the worksites of the knowledge economy, promotes innovation and creativity rather than simply learning and memorisation, recognises and rewards a plurality of talents rather than a singularity of merit (namely, performance on high-stakes assessment), provides a broader diversity of choices and pathways for students in and through schooling, and generally prepares young people to successfully negotiate the more complex institutional demands of a rapidly globalising and "post-modern" world, and to do so without a loss of civic attachment or a clear normative framework.

The new policy settings were initially announced in the launch of the *Thinking Schools, Learning Nation* (TSLN) initiative in 1997. In the past decade since the launch of TSLN in 1997, educational policy in Singapore has been dominated at the broadest level by a vision of "a nation of thinking and committed citizens capable of meeting the challenges of the future, and an education system geared to the needs of the 21st century" (*www.moe.gov.sg*). Specifically, this vision has centred on the pursuit of five strategic objectives:

- Strengthen *capital formation* appropriate for a small but ambitious and highly successful knowledge economy through improved pedagogy, learning

environments and student outcomes (*Thinking Schools, Learning Nation; Teach Less, Learn More; Innovation and Enterprise; IT Masterplan 1 and 2, Engaged Learning*) across the curriculum, but permitting greater choice and diversity and recognition of diverse "talents" without sacrificing the major gains and achievements of the past, including national performance in international assessments (*e.g.*, TIMSS).

- Maintain *meritocratic forms of social organisation*, including the organisation of schooling, in order to promote elite recruitment into public administration and optimal allocative and productive efficiency in the labour market.
- Support and maintain *traditional social identities* but not at the cost of racial harmony through a variety of initiatives, including, in education, the bilingual language policy.
- Promote the moral and civic development, emotional well-being and capacity for full and effective participation in the institutional and community life of Singapore (*National Education, Social and Emotional Learning, Desired Outcomes of Schooling*).
- Prevent the growth of a permanent *underclass*.
- Promote evidence-based policy and planning.

The Singapore core research project

In pursuit of these objectives, in 2002 the MOE and the National Institute of Education (NIE) in Singapore announced the establishment of a Centre for Research in Pedagogy and Practice (CRPP) at the NIE with an initial five-year renewable grant of some SGD$49m (USD$31.8m).

Since its establishment, CRPP has pursued three primary objectives:

- To describe and measure patterns of classroom pedagogy (curriculum, assessment and teaching) in Singaporean schools.
- To measure the impact of pedagogical practices on student outcomes controlling for student characteristics.
- To identify opportunities for the improvement of pedagogical practice through a carefully designed and evidence-based intervention (or innovation) strategy.

Core Research Programme

For the first three or so years (2003-05), CRPP's research activity centred on the Core Research Programme. As Luke, Freebody and Lau (2003) indicated in their initial research proposal to the MOE:

> *"The Core Programme is the foundation for CRPP's research, providing a multi-dimensional baseline of descriptive, observational and intervention-based data. This programme employs a variety of design and analytic strategies, over short-, medium- and long-term time spans. The research addresses questions that are consequential for classrooms, schools and policy-making bodies." (p. 4)*

The Core Programme begins from an analytic map of the broad variable pathways from diverse linguistic/cultural communities and socioeconomic backgrounds to and through schooling. This will generate a picture of the social, demographic and cultural factors that shape school performance and outcomes and assess whether and to what extent these patterns fit the meritocratic ideals of the system (*Panel 1*). At the same time, the design focuses on the practices of pedagogy defined broadly to include knowledge, instruction and assessment: on both the everyday patterns of classroom talk and work, and on how system policies, school structure and leadership, teacher training, belief and attitude, curriculum, assessment influence and motivate teachers' work (*Panels 2, 3 and 4*). The design also expands the definition of educational outcomes from conventional indicators of achievement (year level retention, marks and grades, test and examination performance) to include student artefacts (*Panel 5*) and a broad array of social, economic, civic and psychological outcomes and life pathways (*Panel 6*).

The aim of the Core Research Programme, then, was to provide a rich description and comprehensive overview of pedagogical practices and student outcomes over variable levels of schooling in Singapore. In so doing, it attempted to capture the complexity of a system in a way that an experimental design, for example, cannot. Instead, methodologically, the resultant Core design is:

- *Multi-method*: The different panels enable the blending and triangulation of quantitative (survey, observational) and qualitative (observational, discourse analytic, interview) data.
- *Multilevel/hierarchical*: Samples of students, classrooms and schools are nested across panels, and linked to a comprehensive population database on achievement and socio-demographic background.
- *Cross-sectional and longitudinal*: Cross-sectional samples and multi-year repeated measures are combined.
- *Representative and generalisable*: Schools, teachers and students are selected from random stratified samples.
- *Multidimensional*: Multiple outcomes – cognitive and social outcomes – are assessed through high-stakes assessment results, conventional assessments in English and Mathematics, evaluation of student artifacts, and longitudinal surveys.

Table 13.1 briefly describes the six panels that together make up the Core Research Programme.

Table 13.1. Core Panel Design (2004-2005)

Panels	Sample	Key focus
Panel 1: **Student background/ achievement**	Entire school population from 1993-2002+ (500 000 students pa).	Modelling impact of SES, race and MT on student achievement in high stakes assessment in primary, secondary and postsecondary levels.
Panel 2: **Teacher and student survey**	Sample (n=19 000) primary and secondary students in random stratified sample of schools. Sample linked to Panels 3, 4 and 5 and linked to Panel 1. Sample of teachers (n=4 000) in same primary and secondary schools across all subjects.	**Students**: Modelling impact of classroom pedagogy on student achievement in Math and English controlling for student characteristics. **Teachers**: mapping pedagogical capacities and teaching practices. Also school climate and leadership.
Panel 3: **Classroom observation and coding**	2004/2005: Sample of 1 200+ lessons in Math, English, Science, Social Studies, Chinese, Malay and Tamil in 56 schools using the Singapore Coding Scheme.	Structure and distribution of classroom pedagogical practices with respect to knowledge, teaching and assessment.
Panel 4: **Discourse analysis of classroom interaction**	Audio-taping and selected video of lessons drawn from Panel 3 above.	Structure of classroom talk, patterns of social interaction, language patterns and knowledge construction.
Panel 5: **Analysis of student work**	Same sample as Panel 3 and 4 above.	Teacher assessment tasks and student work artifacts (worksheets, homework, projects) produced in response. Both evaluated for intellectual quality by expert teachers using rubric drawn from Panel 3.
Panel 6: **Longitudinal survey of student experiences, choices, pathways and attainments**	Three samples of students (Primary 4, Secondary 1, Postsecondary 1) (N=28 500) in 100 schools and postsecondary institutions tracked for an initial period of 3 years.	Longitudinal measures of life experiences, patterns of social participation and attainment and life goals, choices and pathways. Includes standardised assessment in English and Math.

Specific Focus Projects

Since the middle of 2003, CRPP has designed and implemented over 120 Specific Focus Projects (SPFs). Luke, Freebody and Lau explain (2003):

> *"While the Core Programme directly addresses the key questions shown above, the Specific Focus Projects are aimed at addressing questions about particular facets of classroom, school, and system practice, and at adding substance and detail to the findings developed from the Core Programme. Generally of shorter duration and with more specific curricular foci, the set of Specific Focus Projects will display a mixture of methodologies, analyses and time-spans." (p. 5).*

SPFs are both conventional research projects (using both quasi-experimental designs and design-experiment designs) and innovation projects generally focused on interventions in domain-specific fields – literacy, English language, Mother Tongue (Malay, Chinese, Tamil), Mathematics, Science, IT, Social Studies and Drama. In Mathematics, for example, during 2004 and 2005, CRPP funded a number of SPFs.

Evidence-based innovation programme

The "Intervention" Programme was intended to answer two general questions: "How can students' learning be enhanced? And how can students' application of knowledge to new task settings be enhanced?" (p. 7)

During the second half of 2005, CRPP staff began to review the research findings from the Core Research Programme and the Specific Focus Projects and identified, designed and began, at the beginning of 2006, to implement some 15 intervention projects within an intervention framework based on a number of key principles:

- CRPP interventions focus on:
  - promoting student engagement;
  - developing disciplinary and transdisciplinary understandings; and
  - developing valued social competencies (work, citizenship).
- By building teacher capacity in:
  - curriculum design;
  - assessment literacy (formative, authentic);
  - evidence-based "reflective pedagogy";
  - pedagogical realignment at the classroom and school level (*e.g.*, through "backward mapping", professional deliberation);
  - recognising, valuing and supporting student diversity.
- By promoting organisational change, specifically, the organisation of the school as a professional learning community:
  - professional reflection/deliberation:
    - individual;
    - collective (year level, subject);

  - evidence-based decision-making:
    - school-wide student database;
    - continuous formative assessment;
  - distributed leadership;
  - school-based, pedagogically focused and effective professional development.
- And by promoting appropriate forms of pedagogical alignment:
  - curriculum, assessment and teaching;
  - balance of tight and loose coupling:
    - tight coupling of enacted curriculum and assessment;
    - loose coupling of assessment and instruction/teaching;
    - multidimensional (including centrally moderated school-based authentic assessment).

These interventions will not be completed until the end of 2007 or later. While these are impressions and not hard data, we have been struck by a number of conclusions:

- *Teaching situations* are inherently problematic, messy, indeterminate, non-routine, uncertain, unstable, unique, reflexive, fluid, unpredictable, non-standardised and agentic…even in Singapore! The character of the teaching situation has important consequences for the nature of schools as organisations, for the regulation of pedagogical activity and for understanding processes of pedagogical innovation. However, in Singapore, compared to many other systems, the national high-stakes assessment system assures a tight coupling of pedagogy to system priorities, although it also constrains the opportunity for pedagogical innovation in schools.
- *Good teaching* cannot be bureaucratically scripted. While teaching can – and often is – viewed as a rational technical activity or "science" subject to general laws that can be developed into rationalised (pre)-scripted pedagogical (or practical) algorithms designed to achieve specified goals, such a view of teaching ignores the inherently messy and deeply agentic character of the classroom situation. It is thus more useful to think of good teaching as a complex reflective practice requiring continuous and ongoing inquiry, individual and collective reflection, and principled practical judgment in ever-changing classroom circumstances. Generally it requires significant "teacher change".
- *Teacher change* depends on a number of enabling factors. First of all, we have found that teacher change depends hugely on *teacher commitment and sense of agency*. This requires the active involvement and support of teachers in the identification of pedagogical challenges, solutions and strategies. Teachers have variable levels of *commitment* to, and ownership of, the process of pedagogical change. Many teachers in Singapore see little or no reason why pedagogical practices should change – after all, they suggest, Singapore has done exceptionally well in international assessment, the system is well funded and managed, pedagogical practices well tested and culturally appropriate, their own histories testament to the ability of the system to promote high levels of student

achievement and social mobility. Besides, pedagogical change is uncertain and risky and hard work technically and emotionally. Others recognise the need for pedagogical change but argue that significant pedagogical change is difficult in the current assessment environment. Teacher change is a matter of challenging and altering teacher beliefs and conceptual understandings, developing commitment to specific professional norms and processes, supporting the development of specific kinds of professional identities and attachments, and helping teachers cope successfully with the emotional and technical demands of teaching and pedagogical innovation. Second, teacher change is a matter of building *technical capacity* – developing content and pedagogical content knowledge, particularly at the conceptual level; developing skills in classroom enquiry and collaborative reflection and planning; and developing the capacity for informed and principled *pedagogical judgment*. Finally, teacher change depends on giving teachers ample *opportunity* to observe and practise desired pedagogical innovation and to be coached, mentored and otherwise supported in ways that facilitate sustained teacher change and to be supported by the school administration and colleagues in the school without fear of penalty if specific innovations fail to deliver desired results. Ironically, successful innovation depends on acceptance of risk, uncertainty and failure (OECD/CERI, 2004).

- *Organisational change*. Successful pedagogical innovation depends on *organisational and cultural change*, including changes in patterns of teacher belief, values and identities and the development of appropriate organisational supports (de-privatised practice, developed forms of classroom inquiry and knowledge production, collective reflection, and strong and distributed leadership).

Reporting: towards a knowledge management and innovation system

One of the key commitments of CRPP to the MOE is to provide timely and useful advice to the Ministry, and the teaching profession more broadly, on CRPP's research and intervention findings. We do this in a variety of ways:

- Annual technical reports to the Ministry of Education summarising CRPP's research and intervention findings.
- Preparation of policy-friendly summaries of research and intervention findings for senior policy makers and professional audiences.
- Annual presentations of research findings to the Minister and senior MOE officers.
- Annual presentations of research and intervention findings to principals and school staff involved in CRPP research and intervention projects.
- Periodic presentations to mid-level MOE officers, principals and teachers.
- Presentations at academic conferences (*e.g.*, we gave 42 presentations at the 2006 AERA meeting in San Francisco).
- Publications/chapters in peer-referred journals and books.
- Participation in policy conversations with senior MOE officers.

- Editorship of two peer referred journals (the *Asia Pacific Journal of Education* and *Pedagogies: An International Journal*).
- Publication of a professional journal (*SingTeach*) for the teaching profession in Singapore.
- Periodic presentations to NIE teaching staff and senior management in part to inform teacher education and professional development programmes.

For the last half dozen or so years, the OECD's Centre for Educational Research and Innovation (CERI), has suggested that contemporary schools, as they are currently organised, are not appropriately designed to successfully address the manifold and complex institutions demands of modernisation, modernity and knowledge-based economies. Contemporary schools, CERI argues, are not yet "Schumpeterian" institutions although the successful "schools of tomorrow" will be radically different institutions from today's schools (OECD/CERI, 2004, p. 11). Above all, the "schools of tomorrow" will need to be institutions that are especially adept and effective, not merely in transmitting knowledge to the next generation, but in producing, disseminating, applying and institutionalising knowledge that increases the effectiveness of contemporary schooling and promotes the development of knowledge societies. Given the pivotal role that knowledge production and innovation plays in organisational improvement, and the critical role that education plays in shaping the future of Singapore more broadly in a rapidly globalising world, not the least of challenges confronting the NIE in Singapore, and schooling more broadly in Singapore, is the development and institutionalisation of effective knowledge management and innovation systems in educational institutions of all kinds in Singapore.

In this task, CRPP/NIE can both be a model to other educational institutions and a strategic partner with the Ministry in the development of schools as knowledge management and innovation systems across schools at all levels.

Five areas have been identified below as a framework of research issues to improve our understanding of knowledge and learning processes in education and in a broader context of the knowledge economy and society. First, the way in which knowledge and learning are managed by modern organisation and in the education system. Second, ways in which this knowledge can be identified and measured, whether by the organisations themselves or by policy makers and the wider public. Third, specifically in education, how improved knowledge management may create organisations that become more effective at learning and innovating than they have been in the past. Fourth, the challenge to R&D systems within education to become a more effective part of knowledge management in this sector, potentially creating new structures that bring them close to policy-making and practice. Finally, the pursuit of a specific breakthrough in the knowledge used by education, by bringing together brain specialists and learning specialists to pursue a better understanding of learning processes. (OECD/CERI, 2000, p. 98)

Not the least of the challenges teachers, researchers and schools will face is a radical rethink of the relationship between teacher knowledge and effective innovation in classroom practice that will require teachers to abandon privatised forms of professional practice in favour of collaborative and reflective partnerships with fellow practitioners and researchers.

Conclusion

I want to conclude with a brief consideration of the problem of transfer: To what extent is the Singapore model of knowledge management and innovation "transposable"? Are there general lessons to be learnt?

In general terms, I can see no reason why the key features of Singapore's emergent system of knowledge management and innovation are not transposable to other jurisdictions. However, there are some particular features of the Singaporean context and its specific institutional arrangements that have functioned to support the knowledge management and innovation system in quite distinctive ways. For example, Singapore has a highly centralised system of school governance, resulting in a system of very tight coupling between instruction and policy, strong policy leverage over instructional practice, and secured by a powerful and complex (some would say over-determined) regime of bureaucratic, discursive, cultural, cognitive and performative controls over instructional practice.

The Singaporean educational system is also relatively small and only modestly differentiated institutionally, with considerable uniformity of pedagogical practice in Singapore across levels of schooling, subjects and streams. The institutional and governance relationship between NIE and the Ministry is unusually close and effective. NIE is the sole provider of teacher education and a major provider of in-service training in Singapore. The government has demonstrated an exceptional willingness to invest a considerable amount of public funds in research and innovation, and it does so because it is deeply committed to rapid and appropriate levels of capital formation that will enable it to negotiate the knowledge-based economies, and 21st century institutional arrangements more generally, effectively.

And finally, there is broad acceptance within NIE of the importance and value of accepting government funds for strategic policy-directed research at the expense of traditional solo research by academics following their own interests. These conditions are distinctive and important, but they are not, in my view, individually unique or collectively necessary for the creation of an effective system of knowledge management and innovation in other cultural contexts.

References

Brown, P. and H. Lauder (2003), "Globalisation and the Knowledge Economy: Some Observations on Recent Trends in Employment, Education and the Labour Market", Working Paper, School of Social Sciences, Cardiff University.

Hogan, D. (2007), "NIE Research, Development and Innovation Plan: Background Working Paper", unpublished paper, National Institute of Education, Centre for Research in Pedagogy and Practice, Singapore.

Luke, A. (2005), "A CRPP Intervention Plan: Moving from the Core to Pedagogical Change", unpublished paper, National Institute of Education, Centre for Research in Pedagogy and Practice, Singapore.

Luke, A., P. Freebody and S. Lau (2003), "Core Program Proposal: Case for Support", unpublished paper, National Institute of Education, Centre for Research in Pedagogy and Practice, Singapore.

Luke, A. and D. Hogan (2006), "Reworking What Counts as Evidence in Educational Policy: The Singapore model", in J. Ozga, T. Popkewitz and T. Seddon (eds.), *World Yearbook of Education 2006: Education Research and Policy*, Routledge, London.

OECD/CERI (2000), *Knowledge Management in the Learning Society: Education and Skills*, OECD, Paris.

OECD/CERI (2004), *Innovation in the Knowledge Economy: Implications for Education and Learning*, OECD, Paris.

Part Four
The Politicians' Perspective

Chapter 14
Research-based Policy-Making: The Need for a Long-term Perspective

Johnny Nilsson, Former State Secretary, Ministry of Education and Science, Sweden

Policy makers often hunt for evidence-based educational research that can prove which educational practices are superior to others. In this chapter, Johnny Nilsson from Sweden gives a few examples which show that this process is possible and profitable, but difficult, and usually needs to be carried out over the long term.

In Sweden the use of empirical research in relation to policy-making has a long history. In several parts of that history empirical research has had the character of evidence-based research with an impact on the process of policy-making. There seem to be at least six different ways in which evidence-based policy research may have an impact. It can:

- create more clear illustrations of the policy in use;
- question and challenge the dominant policy;
- act as a catalyst;
- control how well the policy is implemented;
- legitimate the policy;
- construct policy.

The Swedish way to use research as a basis of policy-making processes for education has changed over the years. During the period between 1940s and the early years of the 1990s empirical research was used as an integrated part of the work of state commissions. For instance within the experimentation with a new comprehensive school system carried out during the 1950s municipalities were invited to shape new solutions on the organisation of schools and empirical research followed the effects of the different ways to organise schools and classes. Municipalities were invited again during the mid 1980s to create solutions on the way upper secondary schools could be arranged to have larger effects. Researchers followed the new solutions, reported the findings and new policies for the upper secondary schools were created. In earlier years state commissions worked through the problems of the educational system in depth, taking years to study the problems and to deliver solution proposals.

Imbalance between the tempo of policy-making and of research

As the tempo of the political processes have risen during the end of the 20th century the old model of slow working state commissions has been left. Today state commissions seldom work for periods longer than a year. Although work in these commissions usually is based on a solid research ground, the commissions nowadays seldom initiate research. More efforts within the education system are spent on evaluations where the experiences of new solutions gradually are followed leading to step-by-step policy changes. This has been the aspiration of the social democratic governments of Sweden that I worked for. But to use empirical findings produced as evaluations of policies that have already been practiced may face many difficulties. An example from the last years may illuminate this.

In the mid-1990s the social democratic government which had replaced a coalition government in 1994 took an adult education initiative. Originally it was not very much based on empirical research. The overarching goal of the incoming social-democratic government was to reduce unemployment by half. Unemployment was at unprecedentedly high levels at the time. And the adult education initiative was a cornerstone of the government's programme to reduce unemployment.

The process started in October 1995 with the so-called Growth Bill where an education initiative was announced. At the same time the Commission for the Promotion of Adult Education working at the time was asked to come up with suggestions for the contents of an initiative to be more fully presented in Spring 1996. In the meantime two government bills specified more clearly the training and the number of training slots that were to be financed. To start with 100 000 training slots were to be provided. In the Budget Bill for 1997 the accompanying study grant system was introduced and the Economic Spring Bill of 1997 opened up for even more study seats to be provided (up to 140 000).

There were no clear evaluations underpinning these decisions. There were no evaluations of the effects of municipal adult education made at that time. There were, however, some indications that municipal adult education was inefficient and had to be renewed. Labour market training at the high levels now operated was thought not possible to expand further. To the general arguments in favour of an education initiative belonged: that Sweden should compete with high skill levels not low wages; that education would increase the productivity and wages of low skilled individuals; and that forecasts showed an increasing demand for more highly skilled labour.

However, even before the programme started to be implemented, in July 1997, the idea came up that the adult education initiative was to be accompanied by evaluations. The first traces of these evaluations appear first in the Bills presented in 1999. This may not be surprising given that it was decided to ask the sitting Commission for the Promotion of Adult Education to initiate independent evaluations. In May 1997 the Commission received the instructions and the budget for the independent evaluations. A tender process was concluded in 1997. Therefore in March 1998, the Commission could only report to the Government about the planning of evaluation. A second report was produced in March 1999 and the Commission's final report in March 2000. They contained evaluations results but, of course, the labour market effects for the participants were still rudimentary. This was due among other things to the fact that many individuals continued to higher education after the adult education initiative. The evaluation programme was therefore not of much use even when the political conclusions were to be drawn from the adult education initiative. The programme was a five-year programme

and ended in 2002. Well in advance the stakeholders had to be informed as to what would happen after the programme. That was communicated in February 2001 (Bill 2000/01:72).

However, the up-coming evaluations are mentioned innumerable times in the Bills during the period. There was a genuine interest to learn from the evaluations and some disappointment that results were so slow to show up. There was some frustration as to the many aspects to the evaluation of this huge programme – its implementation and macro- and microeconomic effects for the individuals, the schools, the teachers, the local employment offices and labour markets, the municipalities, etc. The interest in evaluations could also have been motivated by an expectation that they would provide justifications for the programme and by this time evaluations had more or less become a political must. The example shows that there are many obstacles to using empirical policy research in such a way that it can feed back important findings that can be used in new adjustments of policy. Although you might have high ambitions to use empirical research as a basis for the policy making process, the tempo of policy-making and the tempo of empirical research are not adjusted to each other. Policy work needs long term empirical research that is ongoing along with all the quick moves that modern politics require.

The long-term perspective

Today comparative research is part of the policy landscape. The OECD-driven PISA countries face challenges that are new, as the quality of the internal education is seen in the light of the quality in other countries. The political debate on what development steps need to be taken in the federal state of Germany, with demands on shared standards for the schools in some subjects, might be the most obvious contemporary example. The time that it takes to move from the catalytic moment to a new and stable practice that may come out of such a political process is usually rather stretched out, as the example given above on the adult education initiative shows.

The long term engagement that needs to be a basic feature of the political system can be found in another Swedish example of evidence-based policy research. In the 1970s a state committee reviewed the Swedish system. Among other activities the committee made experiments where a dozen schools were trusted to use their resources on basis of local decisions instead of centrally taken decisions, which were the common pattern at that time. A scientific study (Kilborn and Lundgren, 1974) was linked to this experiment in which the experiences of the schools were described and interpreted. One of the main proposals of the committee argued for more autonomy in decision-making of the local school and an overall decentralisation of the whole educational system in Sweden. The study that followed the experiment pointed at the importance of school plans linked to local evaluation as important conditions if the decentralisation strategy were to be put into practice. When the government later on worked out proposals based on the committee work in the end of the 1970s it proposed that each school had to produce a local working plan and to evaluate its own efforts once a year. As was shown by a follow-up study (Ekholm (1987) of these decisions carried out during 1980-1985, they were only partially implemented in local schools during the first five years after the political decision.

Since 1991 the National Agency for Education of Sweden (NAE) carries out national evaluations of the school system, used by policy makers at the national level. During the 1990s this agency showed that the responsible bodies – the Swedish *kommuns* and their schools – still did not succeed very well in linking local school plans with local

evaluation work. The government used this information and to require the schools to make local evaluation based on local working plans in 1997. When new investigations were made in 2001 by the NAE (2003), the policy was found to be almost fully implemented among the schools. The experience of the Swedish system shows that it may take as long as a quarter of a century to go from an evidence-based policy initiative to an institutionalised stage in schools.

Some of the politically initiated changes that occur in a society are accompanied by strong streams of argumentation. At the moment when decisions are to be taken sometimes veritable volcanic eruptions of arguments break out. Ideas are confronted with counter ideas as well as with threatening future visions from different participants. Sometimes evidence-based research enters the eruptive situation using a calmer and sober tune by reminding the fighting parties about what really is known and about what is not known. In Sweden there have been recent examples of this use of evidence-based research in the political debate about "free standing schools", as charter schools are called in the Swedish context. Evidence has been presented that these schools cause some social segregation, achieve a little better results in school subjects, and reach equal results in social development variables compared to ordinary schools driven by the Swedish *kommuns* (Myrberg, 2006). In the inflamed debate, the "free standing" schools seem to cause a catastrophic rise in social segregation as seen by the opponents of this idea. Many of the "free standing" schools themselves have argued that they achieve much better results than other schools both in school subjects and on social development variables. Empirical studies produce evidence that shows that these sayings are false and that the actual picture indicates no change in the policy of "free standing" schools. The existence of this kind of schools is authorised by the actual research.

Interpretations of research findings are important

The dream of the use of evidence-based policy research is that the research will present solutions to problems that exist. Examples of when this kind of dream has been used as a ground for actions can be found. The political efforts to use research as the true basis for policy recommendations in the United States (*What Works* published in the 1980s) is such an example. Another example is the strong belief of some educational policy makers in several countries on the findings of the early research on effective schools and how these findings could be copied in less effective schools. Neither the use of selected theses taken from educational research nor the use of findings from studies of specific effective schools have been a success. The use of the evidence that research presents usually needs to be transformed several times before the research results can be transformed into policy. Policy makers seem to hunt for evidence-based educational research that can prove educational practices that are superior to others. Educational research shows that there are many successful practices and that the success to a large extent depends on the context in which different methods are used, and as such turning these context dependent findings into policy demands interpretation. Some of the examples given in this paper show that this process is possible and profitable, but difficult, and usually needs to be carried out over the long term.

To summarise, I think that we can gain a lot from creating a good cooperative climate between systematic research and politics. I see that we can develop the relations between researchers and politicians in many productive ways. Of course politicians cannot rely too much on research, at least not in the short term, as I have seen many research reports that have had little value for politicians. The reports may have been interesting, but they have

been presented at the wrong time. To a large extent, politics is a question of timing. Proposals need to be presented when the time is ripe for them and many research products are presented much too late to fit into the needs of modern politics. They are more slow afterthoughts than vital challenges.

In the field of education I see many areas where research work could be of help to the political process. In Sweden, the numerous multicultural schools that we have would be such an example. Why is it that school results are lower for all categories of students in Malmo schools with more than 50% of the school population coming from another country, while students of all categories succeed much better in schools in Lund 25km away where the proportion of migrants is less than 20%? Why are students coming from higher socio-economic backgrounds less successful in some cities compared to other cities? When we get researchers to illuminate questions like these, they will feed politicians with important material.

Under some circumstances, it would be better for researchers that are interested in the educational field to behave more like researchers in the field of national economics. Educational researchers are often too occupied with pedagogical and didactical questions that mainly are of interest to the professions working at school. Few of them concentrate on problems that deal with the management and steering of education. Relevant questions would be, for example: how is the governing process practiced? Who is really influencing decisions that are taken? In what way do national and local political decisions about education cooperate or conflict? In what way do political decisions reach the every day life of teachers and students? Educational researchers need to show an interest in structures and government in the same way that researchers of national economics do. If they do not, someone else will play this role. Moreover, I have seen too many examples of national economic researchers that have acted as amateurs within the field of education. The result does not raise the faith in the kind of research that these researchers do, and does a disservice to education.

One way to create a climate where we can get more educational researchers to contribute politically useful research would be to support more frequent interaction between researchers and politicians. As a politician I look forward to analysing questions that jump up on the political agenda together with researchers in a regular way. To get a state-of-the-art report from researchers about actualities that we as politicians wrestle with, could probably help the researchers to find new blind spots that could lead to more research. In the Swedish context, we have chances to create such regular exchanges between politicians and researchers when we start to use our council for educational research in a more active way. I look forward to doing so in the future.

References

Ekholm, M. (1987), "School Reforms and Local Response: An Evaluation of School Reviews in 35 Local Management Areas in Sweden 1980-1985", *Compare*, Vol. 17, No. 2.

Kilborn, W. and U.P. Lundgren (1974), *Forsok med friare resursanvandning*, MAP-gruppen 1974:2, Pedagogiska institutionen, Goteborgs universitet, Goteborg.

Myrberg, E. (2006), *Fristående skolor I Sverige. Effekter på 9-10-åriga elevers läsförmåga*. ("Independent schools in Sweden – effects on third-grader students reading achievement"), Acta Universitatis Gothuburgensis, Göteborg Studies in Educational Sciences, Göteborgs universitet, Göteborg.

NAE (2003), *Årsredovisning 2002*, Skolverket, Stockholm, p. 47.

Chapter 15
Evidence-based Policy: Yes, but Evidence-based Practice as Well!

Maria J.A. van der Hoeven, Minister of Economic Affairs, Netherlands[1]

In this chapter, Maria van der Hoeven from the Netherlands develops her argument for an evidence-based approach to education: the high social importance of education, in combination with limited resources, demands that policy and practice are based on the best possible insights into "what works".

Introduction

A surprising request! For this publication on evidence-based policy research into the field of education policy, Ministers were invited to share their views on this issue. Alongside the vision of researchers, civil servants and other specialists, the political point of view was also considered to be of importance. And that is quite understandable: a combination of the various standpoints would deliver the best for education. It is for that reason that I would hereby like to respond to the request.

After many years as a member of Parliament, I have been Minister of Education, Culture and Science in the Netherlands since 2002. When developing policy I have always tried to strengthen the development of an evidence-based approach, assisted, among other things, by the activities of the OECD in this area. In this respect, I consider myself particularly lucky to have the combination of education and science in my portfolio. I can therefore dare to say that I speak from experience.

My argument for an evidence-based approach to education is simple: the high social importance of education, in combination with limited resources, demands that policy and practice are based on the best possible insights into "what works". This already applies within national borders but even more so across borders. I am therefore pleased with an international initiative such as this one that enables us to learn from the experiences in education (policy) of other countries. The Ministry of Education is pleased to have acted as host for one of the conferences in this OECD project.[2] I have, however, also encountered limits to this approach; reason enough to examine these as well.

[1] From 2002-February 2007, Maria van der Hoeven served as Minister for Education, Science, and Culture.

[2] "Linking Evidence to Practice: Continuing Discussion on Evidence-based Policy Research in Education", The Hague, The Netherlands, 14-15 September 2005, *www.oecd-conferences-ocw.nl/*

I begin with a brief outline of the policy context for the benefit of the readers. The next paragraphs concern a more solid knowledge base for national policy and educational practice respectively. In my closing section I draw up the provisional balance.

Brief outline of the policy context

The main objective of every Minister of Education and of all education policy is to improve the quality of education. Education, being a human activity, can always be improved. I do not want to achieve this by regulating the sector more stringently but by doing precisely the opposite. I would like to encourage the ambitions of teachers and pupils/students and allow room for professionalism. This would not of course be obligation-free: education institutions must be directly accountable for their efforts and outcomes to the stakeholders – pupils and their parents, other youth institutions, students, the business community and the government, mainly through the Education Inspectorate. This choice is not so much based on evidence as on values like responsibility, freedom of choice, encouraging people to make the most of their talents – in short, a social vision as politicians can be expected to have. I am convinced that this vision will benefit the sustained quality of education.

Against this background, I am trying to encourage evidence-based working at two levels: more evidence supporting national education policy and more evidence supporting educational practice. In the title I tried to stress the importance of evidence-based practice. In the Dutch context this is very important, because education development takes place less at the national level and more at the level of the practice of education itself.

More solid knowledge base for national policy

A realistic vision

I am a firm supporter of a more solid knowledge base to support our national policy. That said, I would add that attention is needed for a realistic view of the issue. It is perhaps at the national level that the limits of an evidence-based approach are felt most strongly. I would like to list the four major limits that I perceive to policy development.

The first limit: education is an extremely complex world. The question is whether we are always capable of properly understanding the causal relationships. A familiar example of this is the question of whether the uniform reduction of class sizes across the entire line would lead to better learning outcomes for pupils. This is a subject on which the specialists find it difficult to agree.

The second limit: are we always capable of identifying possible unintentional side effects in advance? What would happen, for example, to the labour market for health care and the police force if (only) teachers were to be awarded significantly higher wages?

The third limit: education is not a laboratory in which you can experiment with various groups of children in a controlled environment. Parents want to have the best education for their children *now* and are not always prepared to participate in experiments or to wait for their results.

And this brings me to the fourth limit: do the often critical citizens of today really want to wait until politicians have gathered sufficient evidence before improvements are implemented? In my experience, politicians are often not given enough time.

Despite these limits, I remain a dedicated supporter of a more solid knowledge base for education policy under the motto: everywhere where it can be done, it should be done. Below, I will outline a number of concrete initiatives that I have taken over the past years. These initiatives are for the long term: they involve changing the attitude of many stakeholders and guaranteeing an adequate knowledge infrastructure. This international OECD project and a recommendation from the Netherlands Education Council (experts in educational practice and education science[3]) were welcome sources of inspiration in this respect.

A few new initiatives

The development of knowledge for policy and education is not a new phenomenon. The Ministry and the education sector naturally have a long history in this area. However, from time to time it is sensible to review its organisation and yield. This has been happening for several years now and it will still take some time before a new working method is firmly rooted and can bear fruit.

The first step is to formulate explicitly the Ministry of Education, Culture and Science's need for knowledge at a central level. In addition, there is a need for a greater focus on the social issues of education policy for the years to come.

A second step is needed to fulfil the Ministry's need for knowledge. The so-called *Kenniskamer* ("Knowledge Chamber") was established to strengthen the relationship between policy and research. Advisory Boards, planning offices and research institutions meet within this Knowledge Chamber to focus their programming more explicitly on the knowledge needed to develop policy and to introduce more cohesion into the development of knowledge. For those of you who are interested: more details of this Knowledge Chamber are provided in Chapter 8 of this volume.

Although intended for local educational practice, at the national level the knowledge infrastructure of education also needs to be adjusted to fit the role that will be assigned to this professional group in the chosen governance approach. If we ask school managers and teachers to base their practices as far as possible on proven effectiveness, then we have to ensure that they are actually able to do so. Does education research answer their questions satisfactorily? We want evidence to be practice-based just as practice should be evidence-based. And how is the knowledge gained communicated? This is why we are reviewing the way in which knowledge is produced and made accessible. This involves both knowledge originating from scientific research as well as knowledge from validated educational practice. There is currently no blueprint for a single large institute or a digital desk. Rather, this is more likely to become an organic growth towards a better harmonisation of supply and demand in programming and towards the improved accessibility of knowledge.

[3] Recommendation "*Naar meer evidence based onderwijs*" ("Towards more evidence-based education"), Education Council, 2006. For an English summary visit: *www.onderwijsraad.nl*

More solid knowledge base for educational practice

The heart of educational practice

From the governance approach outlined above, in the Netherlands the heart of education development lies in educational practice: in schools, with school managements and teachers who consider this as part of their profession. However, they must be able to provide the choices they make in this regard with a knowledge base that is as solid as possible. The improvements they wish to adopt must have already proven their effectiveness in scientific studies or as good practices at other schools. Schools often pursue this end, but do not always succeed in achieving it. It is definitely not always the case that there is correct and sufficient knowledge available in a smart way. And not all school managements and teachers have developed this attitude to a sufficiently high degree.

In the past years, therefore, I have taken several initiatives aimed at furthering this development. These initiatives involve changing attitudes so that school managements and teachers can reflect more on the impact of their choices, on producing the relevant knowledge and improving access to it.

A few new initiatives

Pilots have been implemented with what we term "academic training schools". At these schools, it is not only "ordinary" teachers who are trained in cooperation with teacher training programmes, but a link is also established between teaching and carrying out research within the school aimed at the further development of the school. These pilots could lead to a new position of "research teachers" with the responsibility of strengthening the knowledge base of the development of their own schools. This would involve both the systematic gathering of practical knowledge from their own and other schools as well as scientific knowledge. These research teachers could add impetus to the contact between educational practice and education research; in this way they could contribute to a more evidence-based approach to the development of education at their school. An additional but no less important consequence is that this would offer senior teachers a new career opportunity. And that is an enormous benefit in a time of teacher shortages.

As a general measure, we have introduced the position of "lector" into universities of higher professional education. Lectors act as a link between the universities of professional education, businesses and other knowledge institutions. The teaching sector can also benefit from this. The education world has access to these lectors from teaching training programmes. Together with schools in the region, they form so-called *kenniskringen* ("knowledge networks") for teachers to pass on practical experiences and to link up with research. The teachers who are trained as research teachers at the academic training schools described above play an important role in this respect. There are several lectors now in the secondary vocational education branch who are working towards encouraging a more research-oriented attitude among teachers as well as working on the relationship between school and professional practice. With these two initiatives we are promoting a more evidence-based professional attitude and concrete work is being done on producing and communicating the relevant knowledge.

In addition to these two national initiatives, I would also like to mention here the Dutch contribution to the "Schooling for Tomorrow" project which is also being run by the OECD/CERI. We chose as our theme: *Kennis delen voor innovatie* ("Sharing Knowledge for the Purpose of Innovation"). We invited both the primary and secondary education branches to indicate themselves how the interaction between research and education, and the sharing of knowledge between schools could be improved. A number of the initiatives referred to above were brought together in the Dutch contribution to this project: working with the knowledge networks built up around lectors/researchers, the development of education by schools and a knowledge institution together, and the academic training school for research teachers. At the end of 2007, the study group will draw up the balance of these activities in order to arrive at a number of recommendations. The answer to the question of what the education sector itself thinks about a well-constructed and properly-functioning knowledge infrastructure is important to education policy in the Netherlands, but it could also be of interest to the education systems of other countries.

As Minister for scientific policy I consider it my job to link education research with educational practice and to improve the use of research outcomes. This puts me in the position of being able to allow more room for disciplines other than the traditional education sciences. In particular, for example, promising initiatives from the fields of neuroscience and cognitive science could provide a more solid knowledge base for a number of aspects of learning and teaching. I expect that when setting priorities for future research these disciplines will have a greater chance of success as they increasingly demonstrate that they take questions from educational practice seriously and devote attention to the applicability of research results.

In conclusion

Collaborating on this publication has been a learning experience for me and I would like to compliment the OECD/CERI highly for offering us this platform for sharing knowledge. The request to provide a political point of view regarding a more evidence-based approach to education and education policy was a welcome motivation to re-examine our own approach in this area. All things considered, over the past years in the Netherlands a lot has been set in motion to promote and better facilitate an evidence-based approach. Both objectives are of a long-term nature. The spheres of educational practice and education research do not know one another very well and they only change slowly.

This publication will hopefully work in two directions. The combined experiences of different countries could soon be the source of a further strengthening of our own approach. Similarly, the Dutch approach as I have outlined above could be a source of inspiration to other countries.

Chapter 16
The Importance of Evidence-informed Policy Research in Education
A perspective from Wales

Jane Davidson AC/AM, Minister for Education, Lifelong Learning and Skills,
Welsh Assembly Government

In this chapter, Jane Davidson from Wales describes the challenging educational reform agenda the country is developing, in particular for the youngest and the most vulnerable children. She shows that this education policy is strongly evidence informed and heavily influenced by international practice.

Introduction

I am very pleased to have the opportunity to contribute to this very timely publication on evidence informed policy research in education from the OECD/CERI. My hope is that this will be an invaluable resource supporting educational policy development and in due course improved educational outcomes within national boundaries. I was delighted that the OECD chose to draw on Wales' experience in a seminar on evidence informed policy during 2006. This publication offers the prospect of bringing our experience to the attention of an even wider audience.

Wales is a small country but we have very big ambitions for education, training and lifelong learning. We are developing a challenging reform agenda which encompasses all phases of education starting with improvements in child care services for our youngest and most vulnerable children. And we are harnessing the opportunities we have been given since the devolution of government that took place in the United Kingdom in 1999 to ensure that our agenda is well matched to Wales' distinctive needs. Since 1999 Wales, like Scotland and Northern Ireland has had responsibility for nearly all areas of education policy. Those powers will be further strengthened by changes to the process of law making being introduced in 2007.

Education policy in Wales is both strongly evidence informed and heavily influenced by international practice. We are developing a distinctive system of education, training and lifelong learning, not a parochial one. If we are to achieve our ambitions with the limited resources that we have as a small nation, we need to draw on the growing international evidence of how we can improve learning outcomes. But we also want other countries to observe and wherever possible learn from us. I hope that this article will provide a further contribution to what is becoming an increasingly fruitful international exchange of ideas and experience.

The Learning Country

Devolution has provided us with the opportunity to create a distinctive education and lifelong learning system in Wales and our ambition is that it should be world class: that is why we boldly call ourselves "The Learning Country".

We set out our vision in September 2001 in our 10-year strategy "The Learning Country". In 2006, 5 years into the strategy, we published "The Learning Country: Vision into Action" as an updated strategic statement. This allowed us to reaffirm our commitment to improved outcomes, report on progress and set out our main priorities from now until 2010.

Through "The Learning Country" programme we want to:

- introduce a radical Early Years education and care programme for 0-7 year-olds;
- put the needs of learners first and encourage their full participation in the policy-making process;
- raise standards across the board and particularly for our most disadvantaged;
- support practitioners and develop their pedagogy;
- provide a more rounded, skills-based curriculum that gives far greater flexibility to schools;
- enhance social inclusion, including much greater participation in post-16 education and training, thereby removing barriers to learning;
- transform 14-19 education and training;
- carry out tri-level reform of our education system, through a new relationship between the Assembly Government, local authorities and schools/colleges.

In short we want to create a skilled and creative nation with opportunity for all and based on policies made for and in Wales.

Evidence informed policy

The Government to which I belong is deeply committed to evidence informed policy-making. Across all of the sectors for which we are responsible we commission and act upon high quality educational research, draw upon inspection evidence, value practitioner knowledge and conduct regular evaluations of our major policies. We are particularly fortunate in Wales to be able to draw upon qualitative evidence from all phases of education from our education and training Inspectorate Estyn.

A considerable challenge for the Welsh Assembly – as with all governments – is to get best value from scarce resources. We believe this should be tackled through thinking smarter, using this research-based knowledge and other evidence in a powerful way.

I would like to illustrate this commitment with some particular examples.

Early years

Early years education and care is one of the foremost areas where we are committed to significant education reform in Wales. We have looked carefully at international

research evidence on the outcomes of early years interventions. This has convinced us that the years before formal schooling are critically important to a child's personal and social development and to their attitudes to learning later in life.

To that end we have piloted the "Foundation Phase" – a new approach to the curriculum, learning and assessment for 3-7 year-olds. This has been piloted in 41 schools and early years settings throughout Wales, including maintained schools; voluntary and private nurseries; playgroups and child minders. In September 2008, we will roll-out the Foundation Phase to all schools and settings in Wales and implementation will be complete by 2011.

The Foundation Phase reflects the research evidence that children learn through well-planned play and a curriculum based upon areas of learning rather than separate subjects. But powerful though that evidence is, one of the key ingredients contributing to the success of the policy to date has been to provide opportunities for classroom practitioners and others to see at first hand how similar policies have been implemented elsewhere. Through funding that we have made available through the General Teaching Council for Wales many practitioners have visited areas which are successfully utilising this approach such as Reggio Emilia in Italy and New Zealand. We see this as being the congruence of educational research and practitioner research of a type that we are especially keen to develop in Wales.

We have also been anxious to ensure that the implementation of this policy should be informed by a powerful evidence base. We therefore commissioned leading researchers from the internationally regarded Effective Provision of Preschool Education (EPPE) Project at London's Institute of Education, assisted by colleagues from the University of Wales Institute Cardiff, to evaluate the pilot. The final evaluation report was published in 2006. Its findings were very encouraging, confirming that there is overwhelming support for the new framework and that it meets the needs of children, their parents and practitioners. It confirmed that the Foundation Phase provides a broad and balanced basis for children's learning and development, and that the emphasis the Foundation Phase puts on play and active learning has had a positive effect on children's progress.

The Report also highlights a number of issues concerning training, staffing levels, qualifications and funding that will be addressed through a workforce action plan. This will consider the additional staffing that will be required, the training needs of existing staff and the qualifications needs and career paths of our new Early Years Professionals.

Pupil assessment

The proposal set out in "The Learning Country" that produced the most favourable response was that to discontinue statutory assessment tests for 7-year-old pupils at the end of Key Stage 1.

It had been apparent for a number of years that teachers are extremely competent in their own assessment of this age group, so I was confident in being able to remove this element of national testing.

Key Stage 1 tests were undertaken for the last time in Wales in the Summer of 2002. Results of teacher assessment in 2003 maintained the high levels of performance in previous years with over 80% of pupils achieving at least level 2 in each subject. This trend continued in 2004 and 2005.

Research on the National Curriculum in Wales in 2004 revealed that there was an appetite in our schools for removing national testing for 11-year-olds and 14-year-olds at Key Stages 2 and 3. There has been a growing perception in the United Kingdom that pupils are tested too often during the early key stages. This had been exacerbated by widespread use of commercially available tests in addition to statutory assessments.

Many teachers, particularly in our primary schools, felt under pressure to "teach" to the tests, with a consequential negative impact on the wider curriculum notwithstanding the fact that the statutory guidance stressed that both measures (*i.e.* teacher assessment and the tests) had parity of esteem. This meant that neither teachers nor pupils were making the best use of educational opportunities during the school year.

I therefore asked Professor Richard Daugherty of the University of Wales Aberystwyth and a leading member of the United Kingdom's Assessment Reform Group, to chair a Review Group to look into this area. Explicit in the Group's remit was a requirement that any proposed system should have the interests of pupils as its primary focus.

I received the Review Group's final report in May 2004. It set out a number of proposals which we were able to accept and which were subsequently supported through a consultation exercise. These were:

- assessment for learning should be at the core of our assessment system, allowing opportunities for the whole child to develop and flourish;
- good and effective use should be made of teachers' own judgements;
- at the end of year 5 pupils aged 9-10 should take diagnostic assessments in literacy, numeracy and enquiry skills. The information derived from these assessments should be used by the Year 6 teachers to prepare pupils for the transition to secondary schools by Y7 teachers in receiving and progressing pupils;
- Year 6 teacher assessment should be retained;
- at Key Stage 3 (age 14) it was recommended that we move away from national tests and put greater emphasis on teacher assessment;
- we should introduce a system where schools could achieve accredited centre status using high quality assessment procedures to support this new approach.

We are also drawing on the work of the Assessment Reform Group and in particular their publication "Assessment Systems for the Future" in preparing our revised and skills-based National Curriculum for 7-14 year olds from 2008.

We believe that having drawn on this powerful research and practitioner evidence, we are in the vanguard of assessment reform in schools in the United Kingdom and more widely. At the same time we recognise the importance of benchmarking our performance against developments in other countries.

This is why we have also committed ourselves to participating in the Programme for International Student Assessment (PISA). The response from secondary schools invited to take part has been very positive. Schools were able to enter their pupils for tests in either Welsh or English according to their choice or that of the pupil. 2006 was the first time that Wales has taken part in PISA as an independent participant. We will know at the end

of 2007 how well we compare internationally. I am confident that we shall show up positively.

The Welsh Baccalaureate

The Welsh Baccalaureate Qualification (WBQ) aims to create an overarching and unifying post-16 qualification intended to maintain depth of study whilst encouraging breadth. The WBQ grew out of concerns here in Wales that the traditional advanced level academic programme followed by many post-16 students was too narrow and not sufficiently inclusive. Post-16 qualification reforms introduced in England and Wales in 2000 were generally felt not to have succeeded in ensuring breadth and balance in post-16 programmes of study. There was also concern in Wales about the lack of parity of esteem for vocational and academic qualifications.

The WBQ is based on a "Core plus Options" model, with the optional part consisting of existing qualifications such as GCSE, GCE, VCE, GNVQ, BTEC National Certificates and Diplomas, or NVQs – all of which form part of the wider UK qualifications framework. There are two distinctive features of the WBQ. The first is that it is an overarching qualification and the second is that it has a Core consisting of four components:

- *key skills* – the development of key skills will be embedded in each candidate's programme, by design in the core component and within specific key skills programmes and through signposting in the Options, leading to assessment and certification;
- *Wales, Europe and the World* – in which the focus is on political, economic, social and cultural issues in Wales, and which sets them in the context of Europe (including the United Kingdom) and the wider world. This component includes a language module;
- *work-related education* – which enhances understanding of the world of work, the importance of enterprise and entrepreneurship and which contributes to careers education and guidance. There is a requirement that all candidates will work with employers and contribute to team enterprise activities;
- *personal and social education* – which includes equal opportunities, social inclusion and sustainable development, and which aims to promote active citizenship. There is a community participation element for all candidates.

The Core has been designed to develop all the key skills, including the wider key skills. It is intended that the study of the key skills will also be contextualised within the other components of the WBQ.

The Welsh Assembly Government has funded the development of the WBQ since 2001. Altogether 31 schools and colleges were involved in the pilot phase. Three cohorts of students have begun the 2-year programme in each of September 2003, September 2004 and September 2005. The pilot phase runs until 2007. The WBQ is being developed and piloted by the Cardiff-based Welsh Joint Education Committee (WJEC). The project director is Keith Davies. Further details on the WBQ are available at *www.wbq.org.uk*.

Ongoing project evaluation has been built into the pilot from the outset. Internal evaluation has been provided by the Centre for the Study of Education in an International Context (CEIC) at the University of Bath. The Centre has provided ongoing formative

evaluation as well as reporting on a regular basis to the WJEC WBQ Team and to the Welsh Assembly Government. The reports produced by the internal evaluation team have included a series of eight which focused on different themes relating to the pilot. The areas covered have included aspects of teaching and learning; the development of key skills; management and organisation within centres; and staff training and support. These reports created an invaluable resource which both identified and critically evaluated best practice from across the pilot settings and which all other participants in the pilot were able to draw upon.

The WBQ pilot has also been externally evaluated by the Centre for Developing and Evaluating Lifelong Learning (CDELL) at the University of Nottingham. The evaluation was based on the collection and analysis of a combination of quantitative and qualitative data. These included questionnaire surveys of staff, students and parents; case study visits to centres; and interviews and consultations with a range of other key stakeholders.

The Nottingham team concluded that the WBQ model was suitable for rollout to 14- to 19-year-olds in all schools, colleges and workplace providers in Wales and made recommendations for future planning and action aimed at ensuring that this could be achieved successfully. The evaluation highlighted the enormous amount of progress that had been made during the course of the pilot, whilst also drawing our attention to some important challenges surrounding more general implementation. The evaluation report is available at *www.nottingham.ac.uk/centres/cdell/ltsn/*

This evaluation evidence made a crucial contribution to the decision I announced in autumn 2006 that the Welsh Baccalaureate Qualification would be rolled out to all schools, colleges and work-based learning providers in Wales from September 2007. In fact, from September over a third of schools and colleges in Wales will be offering the new qualification.

Devolution of student finance

Higher education tuition fees of up to £3000 were introduced in England following the Higher Education Act of 2004. That same Act devolved responsibility for determining such matters in Wales to the Welsh Assembly Government. This offered the prospect of a wide range of benefits for Welsh learners. A commitment to widening access to higher education is a well established aspect of our policies in this area. The new powers gave us the opportunity to devise and implement a tuition fee and student support regime to suit Wales's needs and which would benefit our poorest students. It also provided the opportunity to develop a system which would be user-friendly, enable applications and administration through the medium of Welsh and which in due course could encourage the uptake of subjects which would benefit the Welsh economy.

The Assembly Government was, however, anxious that the political debate surrounding the role to be played by tuition fees in Wales should be informed by evidence and by as objective as possible an understanding of the implications for Wales of the tuition fee arrangements that were being introduced in England.

In order to promote this deeper understanding more widely and in order to ensure that the policies adopted in this area were informed by the best available research evidence – including research specifically commissioned for this purpose – we asked Professor Teresa Rees from Cardiff University, one of Wales' leading social science researchers, to chair an evidence based review. The remit was to produce recommendations for the Assembly Government as to the tuition fee and associated student finance policies that

would be most appropriate in a Welsh context. The review was launched in July 2004 and reported the following year.

The Review examined the evidence from around the world on how students have reacted to the introduction of deferred tuition fees. Research specifically commissioned for the review looked at the attitudes of Welsh students towards taking on a loan burden to cover the cost of fees. The Review also considered the financial implications for the higher education sector in Wales of the policies introduced in England given the two very significant cross-border flows of both staff and students between our two countries. The evidence base generated by the Review played a crucial part in influencing the political discussion which preceded the final decision on this matter.

Following the Review the policy we have adopted in Wales involves:

- deferred flexible fees of up to £3070 to be introduced from autumn 2007;
- a fee remission grant of £1845 for Welsh students studying in Wales which provides an incentive for greater numbers of students to study in Wales;
- a Welsh Bursary Scheme that will allow institutions to tailor bursaries to meet their needs but which has a common, means tested bursary made available to all students;
- loans to cover student fee liabilities which only become due for repayment following graduation and when individuals are earning over the repayment threshold.

Students are also able to access grants and loans to assist with living costs as well as support targeted at those with additional needs and commitments.

The Review also highlighted the importance of continuing evaluation of the new arrangements in Wales, partly in order to ensure that the introduction of higher rate fees did not serve as a disincentive for certain groups of students.

The Rees Review also recommended that the impact of the new tuition fee arrangements on part-time students should be the subject of a separate study. Part-time study has become an increasingly important aspect of higher education in the United Kingdom over recent years. Theresa Rees was anxious to ensure that we should guard against the risk that the new arrangements might have unintended consequences in this area.

The Government agreed this recommendation and we commissioned Dr Heather Graham, Director of the Open University in Wales to conduct a separate independent review. As a result of that further review we have introduced a range of measures designed to provide additional support for part-time students, particularly those engaged in small volumes of study. We have decided that the statutory student finance system administered through Student Finance Wales should offer a package of financial support for students studying at half time and above. Below that level we believe that support is better targeted through institutions.

We have provided an additional £10.6 million per annum to support part-time study. The additional funding is designed to enable institutions to strengthen part-time provision without the need to charge higher fees. Wales will begin implementing the new arrangements brought about by the Graham Review for the academic year 2007/08 onwards.

A combination of research and evaluation evidence was, therefore, central to the decision to introduce flexible tuition fees in Wales. They remain central to our implementation of the student finance arrangements in Wales. The outcomes of this further policy research and evaluation will be published in due course.

Areas for further work

These are a few examples of where Wales has successfully harnessed evidence to inform policy. There are many more! I would however like to highlight three further policy areas which are still being developed but where I see the evidence-based approach as having a critical contribution to make over the next few years.

Flying Start

The Foundation Phase is providing evidence-based learning for the three to seven age group. The other pillar of our Early Years policy is for our youngest children and we call it Flying Start. It will fund high quality services for children between 0 and 3. It will build on and complement existing valuable work begun under Sure Start. £46 million has been made available for Flying Start programmes between 2006 and 2008.

It will be based on international evidence of the interventions that support improved outcomes for children in the long term. Local children and Young People's Partnerships will choose the most deprived primary school catchment areas, and in those areas families will have an entitlement to a prescribed menu of services free, good quality childcare for 2-year-olds, additional health visiting, language and play programmes, and the best evaluated parenting programmes.

We have now issued detailed Flying Start guidance on childcare and parenting. This has been based on thorough reviews of existing evaluations. We are also commissioning an original review to define good quality health visiting, to support learning and development right from the earliest stages.

As this programme gains momentum we will continue to commission further high quality research on the effectiveness of interventions so as to inform professional practice.

RAISE

For the Welsh Assembly Government everything that we do in the field of education, lifelong learning and skills is underpinned by our conviction that these areas have a major role to play in achieving social justice. We are determined to tackle and remove the links between deprivation and low educational attainment.

It gave me great pleasure, therefore, to be able to announce in 2006 a major new programme to tackle disadvantage and low attainment in our schools.

We have called this programme – which will involve a total spend of £16m in 2006/07 – RAISE (Raising Attainment and Individual Standards of Education). This funding will be used to support our most disadvantaged pupils and to offer them opportunities they would otherwise be denied.

We are already drawing upon international evidence in the fields of school improvement and educational transformation in framing this exciting programme. We will be commissioning action research and evaluations of the programme in due course.

We intend to link the work going on in individual schools and education authorities with international evidence on tackling the links between disadvantage and low levels of educational attainment. We would be very keen to work with OECD partners in this area.

Practitioner pedagogy

We believe there is a crucially important role for practitioners in achieving our aims. Practitioners in this context includes the teachers, lecturers early years workers, teaching assistants, work-based tutors, youth and community workers and all the other education professionals that we now have in our system.

These practitioners are absolutely vital to the success of the vision set out in "The Learning Country". This is why I consider the work we are doing to encourage, disseminate and network innovative learning and teaching practice – what we call our "Pedagogy Initiative" – to be of such great importance.

Whilst we do not want to define a prescriptive approach to pedagogy, it is clear from a growing body of evidence that the following characteristics are associated with successful practice in our schools and colleges:

- passion for learning which infuses organisations and workplaces;
- constantly looking outwards for new ideas and schemes;
- learner-centred work;
- student work which is focused on problem-solving and enquiry approaches;
- assessment which is used for learning; and
- students being part of a wider learning community, where they support each other and their teachers.

I want all our children born today to experience learning in such environments and to work with practitioners who are regularly reviewing their own teaching styles.

We are now turning our attention to taking this initiative forward over the next five years. It is being informed and shaped by leading international thinking. I have appointed a team of pedagogy champions – current practitioners who will provide leadership across all phases of education. And we will build practitioner networks so that key messages can be disseminated to education settings throughout Wales.

The pedagogy initiative aims to provide independent and authoritative guidance on what appears to work best. We have no expectation that this will produce, in all cases, straightforward answers. It is certainly not our intention to produce some form of official pedagogy. It is, however, important that we use research evidence to produce a robust methodology for the important work we are undertaking. In this respect, I constantly reflect upon the view of Professor John MacBeath that we have probably found out more about learning in the last decade, than was discovered in the previous 2000 years. I am equally convinced that there remains plenty still to discover over the next decade and beyond.

Working together

Be sure that the message from Wales is that we are a nation very ready to join with others in developing the strongest possible networks for evidence informed policy.

As a learning country we will continue to want to learn from others, wherever they are, as we build our vision. A number of countries have contributed to our work to date:

- some of our early years planning has been informed by Finland, Italy and New Zealand;
- some of our bilingual development proposals have been informed by Canada and the Basque country;
- some of our assessment and teacher development agenda is being informed by New South Wales in Australia.

I hope we can look forward to many other countries making a contribution over the years to come.

Wales is also very keen to build international networks and to play a stronger role within the OECD family. It was the Welsh bard Elfed who said "it is good to love one's country, but it is far, far better to love the world". That is my maxim and that of my Government. It's not a bad one for international educational research and the work that OECD sponsors in this area.

Chapter 17
Promoting Evidence-based Policy in Education: The Case of Poland[1]

Jerzy Wisniewski, Expert
CASE – Centre for Social and Economic Research (Poland)[2]

In this chapter, Jerzy Wisniewski describes the current state of Poland's evidence-based policy in education. He explains that, following the country's accession to EU and the involvement of its researchers in international surveys, this policy has made real progress and is now embodied in the Centre for Social and Economic Research (CASE).

Background

In 1989 Poland initiated changes which may be described as revolutionary. The economic system was transformed through the establishment of free market, as were the principles underlying the State and public life through the implementation of democratic procedures. Only education was changing at a slow pace. This may be explained by two reasons. Firstly, this is the very nature of education which may not be changed overnight. Changes in curricular contents require, for example, new curriculum frameworks and textbooks to be developed and teachers to be trained. The introduction of revised curricula in schools should be coordinated with the educational cycles, with new curricula only gradually replacing the existing ones as successive cohorts are promoted to the next years. Secondly, it was not entirely clear which direction changes and reforms of the

[1] This text is an outcome of a short policy seminar which brought together ten Ministers and Vice-Ministers of Education holding their post in different periods between 1989 (the beginning of the democratic transition in Poland) and the present day. The seminar was organised by CASE – Center for Social and Economic Research. "The CASE Education Policy Seminar" was designed as a stocktaking and trailblazing exercise which would ultimately help to put in place a mechanism facilitating the development of evidence-based educational policy. This initiative was inspired by three factors: a continuing need for the educational reform to be supported with the expertise available; the CERI-OECD project "Evidence-Based Policy Research"; opportunities and financial support provided by the Human Capital Operational Programme financed by the European Social Fund between 2007 and 2013.

[2] CASE is an independent, international and non-profit institute founded on the idea that research-based policy-making is vital for the economic welfare of societies. Established in Warsaw in 1991, CASE today is an internationally renowned institute drawing on the talents of prominent economists and driving the creation of a network of partner institutions in transition countries. CASE carries out policy-oriented research and development assistance projects, specialising in questions of European integration, post-communist transition, and the global economy.

EVIDENCE IN EDUCATION: LINKING RESEARCH AND POLICY – ISBN-978-92-64-03366-5 © OECD 2007

education system should take. This was combined with a shortage of experts, both within and outside the Ministry of Education, who could propose the direction and agenda for changes.

First reformatory efforts aimed primarily to make schools free from the taint of ideology – to remove it from curricula and textbooks, in particular those for history and the mother tongue. Efforts were made to promote foreign language learning, mainly through the establishment of foreign language teacher training colleges as an alternative to five-year teacher training programmes offered at higher education institutions.

General lines for change in governance were defined, with the responsibility for the administration and financing of schools to be gradually taken over by local authorities. Higher education institutions were granted extensive autonomy.

These changes were taking place in response to immediate needs (filling in "blank pages" in history; tackling the shortage of teachers) or were related to the general direction of changes in the functioning of the State, for example the decentralisation and delegation of powers to local authorities. No comprehensive long- or medium-term strategy for the development of education was developed. The Ministry was not prepared to provide strategic leadership because it served only as an administrator under the previous regime, while the communist party structures were the decision-making nerve centre.

Research base

The research base of the Ministry was weak and had no contacts with the West. Like in other countries of the Soviet bloc, researchers were concerned with "pedagogical sciences", which meant reflecting on most effective ways to educate a future citizen of a communist state. Fortunately, teachers were not bothered by that work and sought to convey sound knowledge and a reasonable system of values to their pupils.

Higher education institutions did not conduct educational research either, because necessary institutional structures were (and still are) non-existent. Teachers were (and still are) trained in faculties providing programmes and training professionals in specific fields (mathematics, biology, modern languages, etc.). Faculties of education, or "faculties of pedagogy", focused mainly on training teachers for pre-primary education and initial stages of primary education. Interdisciplinary research was not undertaken, because researchers – locked in the rigid organisational structure of their faculties and departments – failed to see the links between education and, for example, employability, labour market, economy, etc.

In the early 1990s, financial support was made available by the European Community within the framework of the Phare Programme. Phare projects funded in the field of education and training were targeted mainly on higher education and the vocational education and training system. Apart from support for changes in these areas, the projects offered an added value by promoting project methodology as a working method. It forced those involved to define objectives, inputs and outputs, to develop monitoring indicators and tools, and to evaluate the outcomes achieved. The project development process required that links should be identified with the economy and labour market. Most projects involved foreign experts and were implemented in co-operation with foreign institutions. This provided an opportunity to exchange experience, access to research and a channel to follow policy debates in the (then) twelve EC Member States.

From the very beginning of co-operation, experts from the Member States involved in the Phare projects drew attention to the absence of a strategic vision for the development of education linked with economic and social reforms and to the weakness of the research base of the Ministry of Education.

This was reflected, for example, in the report prepared under the Phare/UPET Programme in 1994:

> *"The key Ministry of National Education (MoNE) departments are neither structured nor staffed to carry out their new innovative and pro-active roles. There is no permanent secretariat working solely on behalf of the Committee for Reform or the Executive Council.*
>
> *There is no single section within MoNE responsible for ensuring that the decisions of the executive are informed by research, supported administratively, implemented and evaluated.*
>
> *Outside MoNE there are only two institutions readily available to implement policy: CODN (National In-Service Teacher Training Centre) and IBE (Institute for Educational Research).*
>
> *While there are outstanding individual Polish educationalists there is no national network of ready expertise available to MoNE. Nor is there a significant budget to buy in such outside assistance."* (Jan Potworowski, "Final Report on Assistance to the MoNE Policy and Evaluation Development", Phare, Ministry of National Education 1994.)

It is worth noting that, while 12 years have passed since that judgment was made, the Institute of Educational Research has not undertaken yet any work to support the development of educational policy. Despite the introduction of an ambitious reform of the school education system in 1998, including structural changes, curriculum redesign and the establishment of external examinations, the Institute has not contributed on a regular basis, for example, to the monitoring of reform implementation.

OECD and reform

Issues such as the establishment of appropriate structures and a support system for the development of educational policy and, more broadly, a HRD strategy were also raised in the recommendations of the 1995 OECD review:

> *"According to its major function of basic, initial education and skills formation, the Ministry of National Education should be serviced, as soon as possible, by a strategic unit. Among major permanent tasks the unit should be in charge of:*
>
> *a) proposing alternative visions/scenarios of the development of the education and training system;*
>
> *b) developing and maintaining a good statistical indicators' unit or keeping close contact with such unit;*
>
> *c) synthesising key outcomes of educational R&D and evaluation research and advising relevant units on priority research activities to be implemented;*

d) preparing, publishing and disseminating to various stakeholders at regular intervals, an overall state of education and training in Poland which would bring together and interpret quantitative information based on the data so collected;

e) ensuring that, within the decentralisation policy, the above data and surveys would fully cover specific local trends in the framework of broader social and economic development.

The examiners consider that, faced with such lack of basic information , the Polish authorities should have reacted rapidly in establishing or re-establishing some major "think tanks" which could have helped the various stakeholders to get a preliminary appraisal of the situation and some perspective directions for the future. Several ministerial research institutions were disbanded, even within the Ministry of National Education, but those which survived or developed as independent institutions did not seem to be equipped in terms of human, or even material resources, to cover such a complex issue as the current state and likely future of HRD in a transition society. (Reviews of National Policies for Education: Poland, OECD, 1996.)

Indeed, soon after the OECD reviewers presented their recommendations, a unit to be responsible for strategy was set up within the Ministry of National Education. However, its tasks were actually limited to the design of a curricular reform. It soon became clear that this unit was most vulnerable in terms of consequences of political changes. Following the elections, each time the unit was reorganised, its staff replaced and its remit often changed. The short time span for planning and action – from elections to elections – made it very difficult to develop long-term strategies. The only factor that has remained unchanged, despite changing governments and parliamentary coalitions, is the lack of funds for research.

As mentioned earlier, in 1998 the Ministry designed and implemented a comprehensive reform of the school education system within a very short period of time (several months). Curricular reform was the central part of changes. The Minister established core curricula, and schools were free to choose from curricula available on the market or develop their own curricula. In order to measure learning achievements, a system of national tests and examinations was introduced, covering all pupils at the end of successive stages of education (primary, lower secondary and upper secondary education). The responsibility for preparing, administering and marking examinations was given to the newly established central and regional examination boards. In the new structure of the school education system, 8-year primary school and 3-year to 5-year secondary schools were replaced with 6-year primary school followed by 3-year lower secondary school and 3-year to 4-year upper secondary schools. At the same time, the responsibility for the administration of schools was delegated to local authorities. The implementation of the reform was accompanied by changes in the promotion and remuneration system for teachers.

The reform was designed in a very short period of time and, though it built on the work done earlier (including the OECD review of educational policy), its implementation was not preceded by any reliable and comprehensive analysis. Thus it was even more important to ensure proper monitoring of the reform implementation. This task was entrusted to the Institute of Public Affairs, an independent non-governmental institution. Over several years, the Institute produced a number of reports which served more as a basis for public debate than for changes in the policy of the Ministry.

Various papers and reports were also prepared by other non-governmental organisations (*e.g.* Polish Children and Youth Foundation, J.A. Komeński Foundation), on their own initiative and often with support from foreign partners. However, there was no institution or unit to collect and analyse such materials. Neither was there any well-designed system in place to contract research, endorse its findings and ensure appropriate follow-up.

Some hopes for capacity-building were pinned on the involvement of Polish research teams in international surveys such as IALS, IEA Civic Education or, last but not least, PISA. Regrettably, these opportunities were only partly used. Due to limited funding, the Ministry contracted only the necessary minimum set – as defined by the international consortium co-ordinating the survey – of services: developing the Polish-language version of questionnaires, collecting data and preparing a short report. This enabled the experts directly involved in the survey to acquire new important competences. However, the surveys were not accompanied or followed by wider debate on methodology, findings, etc. involving representatives of the research and higher education sectors or prospective users of their findings: policy makers, social partners, etc.

Effect of EU accession

Another chance was offered by Poland's accession to the European Union and the access to EU Structural Funds. Like the programming process for the use of Phare funds in the first years of the transition period, preparations undertaken to use Structural Funds necessarily involved the development of strategies for changes, the identification of objectives and the development of detailed actions plans. However, compared with the support under Phare, the financial resources available were much more substantial. Regrettably, potential opportunities were again limited by external factors. Poland acceded to the European Union at the mid-point of the so-called programming period for the years 2000-2006, and thus had practically only 2-3 years for the implementation of projects. The Ministry concentrated more on effective and fast ways to spend the money than on developing a long-term strategy.

Entirely new opportunities were created by programmes which will be implemented as part of the new Financial Perspective 2007-2013. The preparatory work, including the development of first principles and strategy concepts, started already in 2004. This coincided with the critical debate in the European Union on the Lisbon Strategy and its implementation, which gave an impetus to place more emphasis on educational issues, in connection with the competitiveness of the economy on the one hand and the promotion of employment and greater social cohesion on the other hand.

The Human Capital Operational Programme will be the only programme financed by the European Social Fund between 2007 and 2013, supporting projects which aim to enhance employment and social cohesion, to develop competencies and to improve the quality of the education system. Planned activities include "the implementation of research projects in the field of education". Moreover, the programme "will ensure co-ordination of data collection, which will make it possible to draw up consistent recommendations for national educational policy".

[Detailed description of priorities, a working document of the Ministry of Regional Development, 2006]

Agenda-building

At that time, on the verge of implementing the Human Capital Operational Programme, CASE – Center for Economic and Social Research – came up with an idea to use the experience of former ministers in order to propose some solutions and to create a basis for future dialogue and co-operation between those who carry out educational research and policy makers. 15 years experience has shown that a specific "language barrier" was a major obstacle to the development of such co-operation. A team of experienced policy makers with high standing may become a good mediator, translating expectations of decision makers into the language of research topics and, vice versa, interpreting research findings so that they could be embedded in policy decisions. At the same time, the high standing of individual members and the entire team, based on their experience and will to co-operate despite different political backgrounds, ensures effective leadership of the project.

The idea was put into action by the "CASE Educational Policy Seminar".

The initiative was well received by the present management of the Ministry of National Education. The seminar was attended by two vice-ministers currently in office.

Before the meeting, the seminar participants received a list of key issues for the debate:

- Who is and should be the main user of educational research?
- Is there a need for a brokerage institution providing answers to decision makers' questions on the basis of analysis of available research findings or through commissioned research?
- What should be the scope of responsibilities of such an institution?
  - Collecting data and information, conducting analyses, preparing (periodical) reports?
  - Running an educational research clearing house?
  - Supervising on-going research and ensuring its quality: methodological standards, international comparability?
  - Conducting research?
  - Tendering for research?
  - Developing terms of reference for research projects and participating in the evaluation of tenders?
  - Carrying out activities to promote the development of research (capacity-building)?
- What should be the thematic scope of its activities?
  - Learning and learners: curricula, methodology, teacher training, learning strategies, self-learning, resources?
  - Governance, management, financing, organisational arrangements, quality assurance?

  - Outcomes: school and individual learning achievements, labour market, economy, social capital?

- What should be the status of such an institution? Where should its funding come from? To whom it should report?
- Inter-sector links: the ministries responsible for higher education, labour, health and economy, and the Central Statistical Office?
- Links within the education system: the Central Examination Board, the National In-Service Teacher Training Centre, the Institute of Educational Research, higher education institutions?

Clearly, it was not possible to answer all of the above questions during one relatively short meeting. However, as a result of the discussion, consensus was reached about a number of issues:

- The *research problem is important*. All participants declared their readiness to contribute to the project, to share their expertise with the Ministry of Education which is responsible for strategy development – and "there is no escape from it".
- There is a need to establish a unit which will act as a (*knowledge) brokerage agency*. It should be an independent unit, but closely co-operating with the Ministry. It would propose a list of key issues, identify sources of information, formulate research topics and define the framework for conducting research, and summarise and interpret research findings for policy makers.
- No *new institution* should be created *within the structures of public administration*.
- Such a unit will need to cope with contradictory expectations:
  - to engage in theoretical reflection and to propose practical, readily applicable solutions;
  - to be independent, but serve policy makers, responding to their *ad hoc* expectations;
  - to co-ordinate consultancy services, while not monopolising them.
- The unit *should not carry out research itself*, because this would involve a conflict of interest (contracting institution – service provider).
- The *thematic scope* of the unit's activities should be *sufficiently wide*, because education is not an end in itself. It serves the purpose of encouraging economic growth and enhancing the competitiveness of the country and its regions, contributes to the development of human and social capital, boosts employability and is a key factor promoting social cohesion. Thus, in addition to the Ministry of Education, the customers of "brokerage services" should include:
  - other ministries, in particular those responsible for higher education, and labour and social policy;
  - regional and local authorities;
  - schools, continuing education providers and higher education institutions.

- It is particularly important to provide *reliable information* which would *inform decisions* as well as *inspire public debate*. Such informed debate may give an impetus to, and exert pressure on, decision makers to take specific action. At the same time, public debate is an effective mechanism for public consultation, legitimising and providing support to reforms proposed by the Government.
- There are numerous reports and publications which are not based on reliable research. In many cases, research covers narrow-scope issues, small social groups or a small number of institutions. It is often based on small research samples selected according to questionable criteria. In times of rapid changes, research findings become obsolete quite rapidly. Moreover, some research topics lose relevance, while others emerge and grow in importance. All this makes it difficult or even impossible to draw general conclusions from partial findings.
- At the same time, as reliable information is not available, decision makers often rely on stereotypes and anecdotic knowledge.
- Tasks of the unit
  - *maintaining the continuity* of educational policy;
  - conducting *ex-ante evaluation of new proposals* from policy makers, based on reliable diagnosis and carried out with regard to their implementation;
  - *assessing the impact* of new proposals on the society, economy and labour market – anticipating "side effects";
  - *facilitating the involvement of key partners* – local authorities, teachers' trade unions, NGOs – in the development and implementation of an education strategy;
  - *building public consensus* around the reform agenda. The media and the Internet would be very useful for this purpose.
- Action plan
  - drawing up a list of *key issues* which may set directions for long-term development of education. Clarifying concepts, the language of debate. Formulating questions corresponding to these issues.
  - stocktaking:
    - institutions and organisations which carry out or may carry out educational research;
    - reports, papers, studies, publications, research findings – scope, reliability, relevance;
    - databases;
    - international research and surveys;
  - identifying gaps. Proposing research topics on the basis of key research and surveys.

- The participants recognised the need for internal discussion on educational research within the researchers' community itself. However, according to them, the community is not ready yet to do so.

The unit should blaze a trail in the field, creating and promoting best practice through its activities. Over time, it will become common practice, a routine approach, which will turn into a procedure, and finally a standard – a normal way of supporting the development and implementation of educational policy.

During the discussion, the participants referred to the following passage from Alice's Adventures in Wonderland:

"Cheshire Puss" "Would you tell me, please, which way I ought to go from here?"

"That depends a good deal on where you want to get to" said the Cat.

"I don't much care where–" said Alice.

"Then it doesn't matter which way you go", said the Cat.

"–so long as I get SOMEWHERE", Alice added as an explanation.

"Oh, you're sure to do that", said the Cat, "if you only walk long enough".

Everyone agreed that it did indeed matter which way Polish education was going and that there was no time to walk long enough – because problems to be solved would not wait.

Biography

Adrienne Alton-Lee is the Chief Education Adviser for the New Zealand Ministry of Education's Iterative Best Evidence Synthesis (BES) Programme. Her role is to strengthen the evidence-base informing policy and practice in education and to provide medium term strategic advice to government. Dr. Alton-Lee is a Fellow of the International Academy of Education. She was formerly a teacher, classroom researcher, Professor and an Associate Editor of *Teaching and Teacher Education*. She has published in leading educational journals including the *Harvard Educational Review,* the *Elementary School Journal,* the *International Journal of Inclusive Education* and the *American Educational Research Journal.*

René Bugge Bertramsen is the Deputy General Director for the Danish University and Property Agency within the Danish Ministry of Science, Technology and Innovation. Since 1999 he has been involved in reforms aiming at enhancing the quality of the Danish educational R&D system (such as the establishment of the Danish Pedagogical University – DPU – and the R&D centre Learning Lab Denmark). Mr. Bertramsen was responsible for the University Act of 2003 which gave Danish universities a new governance system, *i.e.* boards with external majority and employed rectors, deans and department heads. In 2006-2007 he was responsible for a merger process where government research institutes were integrated with the universities and a number of single-faculty universities were merged with larger multi-faculty universities, including the merger of DPU with multi-faculty University of Aarhus.

Robert Boruch, Professor, University of Pennsylvania (USA). Dr. Boruch is current co-chair of the Steering Group of the International Campbell Collaboration, and principal investigator for the Institute of Education Sciences What Works Clearinghouse, which is designed to be a central and trusted source of information on evidence about what works in education. Dr. Boruch is an elected Fellow of the American Academy of Arts and Sciences, the American Statistical Association, and the Academy for Experimental Criminology. He has received awards for his work on evaluation policy, randomised trials, and on privacy of individuals and confidentiality in social research. Dr. Boruch's academic background is in psychology, statistics, and mechanical engineering, with degrees from Iowa State University and Stevens Institute of Technology.

Satya Brink is currently Director, National Learning Policy Research, Human Resources and Social Development Canada. She and her team are responsible for developing evidence in support of policy development for lifelong learning for the Government of Canada. This work includes analysis on outcomes for each age group and type of education as well as the impacts of earlier learning on subsequent learning. In her previous post, she was responsible for research on human development based on two major Canadian longitudinal surveys. During this time she and her team produced a major body of evidence based on the National Longitudinal Survey of Children and Youth which influenced major new initiatives of the Canadian government in support of children and their families.

Tracey Burns is a research and policy analyst for the Centre for Educational Research and Innovation, OECD, Paris. Previous to this she worked on social determinants of health across the life-span with Charles Ungerleider & Associates in Vancouver, Canada. As a Post-Doctoral Fellow at the University of British Columbia, Dr. Burns led a hospital-based research team investigating newborn infants' responses to language. Tracey Burns holds a BA from McGill University, Canada and PhD from Northeastern University, USA. She is the recipient of various awards and honours, including the UBC Post-Doctoral Fellowship, a student-nominated university teaching award, and the American Psychological Association Dissertation Research Award.

Thomas D. Cook is the Joan and Serepta Harrison Chair in Ethics and Justice and Professor of Sociology, Psychology, Education and Social Policy at Northwestern University, where he is also a Faculty Fellow at the Institute for Policy Research. He has a BA from Oxford University and a Ph.D. from Stanford University. He is interested in causal methods for the social sciences and in the joint effects of neighborhoods, schools, peers and families on how young people develop socially and cognitively. He is a Fellow of the American Academy of Arts and Sciences and the Margaret Mead Fellow of the American Academy of Political and Social Science. He has been awarded the Myrdal Prize for Science by the American Evaluation Association, the Donald Campbell Prize for Innovative Methodology by the Policy Sciences Organisation, and a Distinguished Research Scholar Prize of the American Psychological Association. He is the author or editor of 10 books and over 150 chapters and articles.

Jane Davidson is the Assembly Member for Pontypridd and former Deputy Presiding Officer for the National Assembly (Wales, United Kingdom). Since October 2000 she has been the National Assembly Education and Life-Long Learning Minister responsible for all aspects of education, training and lifelong learning. Educated at Malvern Girls' College, Birmingham University and the University of Wales, Jane has taught English, Drama and Physical Education. She is also an experienced youth worker and former Cardiff City Councillor. She was a member of the Arts Council for Wales and its Lottery Board, and Head of Social Affairs at the Welsh Local Government Association before her election to the Assembly. Jane has had a keen interest in education and youth work and is enjoying the challenges of the Education and Life-Long Learning portfolio.

Stephen Gorard holds the Anniversary Chair in Educational Studies at the University of York (United Kingdom), and directs the Centre for Research into Equity and Impact in Education. He is currently leading an Economic and Social Research Council (ESRC)-funded project promoting the use and understanding of randomised controlled trials in public policy (*http://trials-pp.co.uk/*), and was the originator of the ESRC's Research Capacity-building Network. He has published widely about the research process in social science, but his substantive work focuses on issues of equity, especially in educational opportunities and outcomes, and on the effectiveness of educational systems. Recent books include "Teacher supply: the key issues", "Adult learning in the digital age", "Overcoming the barriers to higher education", and "Schools, markets and choice policies".

David Gough is Professor of Evidence Informed Policy and Practice and Director of the Social Science Research Unit (SSRU) and its Evidence for Policy and Practice Information and Coordinating (EPPI) Centre, Institute of Education, University of London, United Kingdom. Previously he worked at the University of Glasgow and Japan Women's University. He directs the Methods for Research Synthesis node of the ESRC National Centre for Research Methods Node and research projects for the Department of

Education and Skills, the Teacher Training and Development Agency, the Social Care Institute of Excellence, and the Department for Work and Pensions. Dr. Gough is editor of the journal *Child Abuse Review* and associate editor of the journal *Evidence and Policy*.

Rebecca Herman, a principal research scientist at American Institute for Research (USA), specialises in setting standards for the quality of educational research and reviewing research based on those standards. As the project director for the What Works Clearinghouse, she is responsible for the US Department of Education's flagship project to identify effective educational programmes and practices. Dr. Herman was project director of the *Educators' Guide to Schoolwide Reform*. She provided congressional testimony and many invited presentations on this and related work. Dr. Herman holds an M.A. and Ph.D. in sociology from Johns Hopkins University.

Maria J.A. van der Hoeven is the Minister of Economic Affairs (Netherlands). Maria J.A. van der Hoeven was born in 1949. She was trained as a primary teacher and taught at schools of home economics and junior secondary commercial education. Thereafter she was head of the Adult Commercial Vocational Training Centre in Maastricht and of the Limburg Technology Centre. From 1991 to 2002 Ms. Van der Hoeven was a member of the House of Representatives for the Christian Democratic Alliance (CDA). She has held a variety of social and cultural posts. Ms. van der Hoeven served as Minister of Education, Culture and Science from 2002 until February 2007. She was appointed as Minister of Economic Affairs in early 2007.

David Hogan is currently Professor and Dean of the Centre for Pedagogy and Practice at the National Institute of Education, Nanyang Technological University in Singapore. Between 2004 and 2006 he was Vice Dean for Research at CRPP. Prior to that he was Professor of Education at the University of Tasmania in Australia, and before that he held appointments as Assistant and Associate Professor at the University of Pennsylvania in Philadelphia. He completed his PhD in the history of education at the University of Illinois in 1979. His current research interests focus on the intersections between research, policy and practice, pedagogical theory, curriculum theory and design, the design of knowledge management of innovation systems in schools, multi-level and longitudinal modeling of student outcomes, citizenship and education, and education and social theory.

Bill Kilgallon, OBE, has been the Chief Executive of the UK's Social Care Institute of Excellence since 2003. Prior to that he was Chief Executive of St Anne's Community Services from 1978 to 2002, an organisation he founded in 1971, which works with single homeless people and people with learning disabilities, mental health problems and alcohol and drug problems across Yorkshire and the North East. He was Chair of the Leeds Teaching Hospitals NHS Trust, the largest NHS Trust in the country from 1998-2002 and Chair of the Leeds Community & Mental Health Services NHS Trust from 1992-1998. Bill Kilgallon served as a member of Leeds City Council from 1979-1992 where he chaired the Social Services, Housing and Environment Committees. He has led independent inquiries, including one into alleged abuse in a local authority children's service and one into the management of an NHS hospital for people with learning disabilities.

Hannele Niemi is Professor of Education (1998-) and Vice-Rector for academic affairs at the University of Helsinki, Finland (2003-). She has been Professor of Education in Oulu, Turku and Tampere Universities (1987-1998). She has been a member of the Standing Committee of Social Sciences of ESF, the Council for Society and Culture in the Academy of Finland, and the Scientific Council of the University of Helsinki. She is a Steering Committee member of the British national research programme on teaching and

learning (TLRP). She was Director of the Finnish national research programme "Life as Learning" 2002-2006. Dr. Niemi has been Chair or a researcher in many national and international evaluation projects for development of educational research and teacher education. Her main research interest areas are teachers' professional development, moral education and technology-based learning environments.

Johnny Nilsson is the Former Secretary of State for Education in Sweden.

Andrew Pollard is Director of the Economic and Social Research Council's Teaching and Learning Research Programme (*www.tlrp.org*), the UK's largest coordinated initiative for educational research. As a teacher, his career started in Yorkshire primary schools and he has worked in teacher education or research at Oxford and Bristol Polytechnics and the Universities of the West of England, Bristol, Cambridge and London. He is presently based at the Institute of Education London. Andrew Pollard has published widely, including work on longitudinal ethnography and analysis of social factors in teaching and learning, learner perspectives, and resources for teacher education and school practitioners. He is at present working on an analysis of learning experiences through secondary education.

Rien Rouw is senior policy advisor at the Dutch Ministry of Education, Culture and Science (Department for General Strategic and Economic Advice). He is secretary of the Knowledge Chamber.

Tom Schuller is Head of the Centre for Educational Research and Innovation (CERI), OECD, Paris. Formerly Dean of the Faculty of Continuing Education and Professor of Lifelong Learning at Birkbeck, University of London, his latest books are *The Benefits of Learning: The Impact of Education on Health, Family Life and Social Capital* (RoutledgeFalmer, 2004) and *International Perspectives on Lifelong Learning* (edited with David Istance and Hans Schuetze, Open University Press, 2002).

Hans Stegeman is senior policy advisor at the Dutch Ministry of Education, Culture and Science (Department for International Policy). He is member of the OECD's Education Policy Committee.

Charles Ungerleider is Director of Research and Knowledge Mobilisation for the Canadian Council on Learning. From 1998 until 2001, Dr. Ungerleider served as Deputy Minister of Education for the Province of British Columbia, Canada. Prior to this he was Associate Dean for teacher education (1993-1998) at the University of British Columbia. Dr. Ungerleider has studied and written about educational policy and governance, student assessment, inter-group relations, and the impact of media on Canadian society. His most recent book *Failing Our Kids: How we are ruining our public schools* provides a critical analysis of the state of public schooling in Canada, the key part schooling plays in fostering Canadian values, and how public schools are treated by parents, professionals, and politicians.

Jerzy Wiśniewski is a consultant in education, and public administration and an expert of the Center for Social and Economic Research (Poland). From 2003-2006 he served as head of Strategy and Structural Funds of the Ministry of Education. He was also Director General of the Polish Ministry of National Education at the time of launching the reform of the education system, as well as the head of the International Department of the Ministry of Education and project manager in the Foundation for Public Administration Development. He was a member of the CERI/OECD Governing Board as well as the OECD team reviewing the educational system in Lithuania, advised the Ukrainian Ministry of Education on the reform of the system, and led the team reviewing the VET system in Croatia (with the European Training Foundation).

Also available in the CERI collection

Understanding the Brain: The Birth of a Learning Science
330 pages • June 2007 • ISBN: 978-92-64-02912-5

Demand-Sensitive Schooling? Evidence and Issues
146 pages • November 2006 • ISBN: 978-92-64-02840-4

Think Scenarios, Rethink Education
200 pages • April 2006 • ISBN: 978-92-64-02363-1

Personalising Education
128 pages • February 2006 • ISBN: 978-92-64-03659-8

Students with Disabilities, Learning Difficulties and Disadvantages – Statistics and Indicators
152 pages • October 2005 • ISBN: 978-92-64-00980-9

E-learning in Tertiary Education: Where do We Stand?
290 pages • June 2005 • ISBN: 978-92-64-00920-5

Formative Assessment – Improving Learning in Secondary Classrooms
280 pages • February 2005 • ISBN: 978-92-64-00739-3

Quality and Recognition in Higher Education: The Cross-border Challenge
205 pages • October 2004 • ISBN: 978-92-64-01508-6

Internationalisation and Trade in Higher Education *– Opportunities and Challenges*
250 pages • June 2004 • ISBN: 978-92-64-01504-3

Innovation in the Knowledge Economy *– Implications for Education and Learning*
Knowledge Management series
96 pages • May 2004 • ISBN: 978-92-64-10560-3

www.oecdbookshop.org

OECD PUBLICATIONS, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(96 2007 08 1 P) ISBN 978-92-64-03366-5 – No. 55655 2007

Knowledge Management

Evidence in Education

LINKING RESEARCH AND POLICY

Education policies and systems in all OECD countries are coming under increasing pressure to show greater accountability and effectiveness. However, the information readily available for policy making is often unsuitable, either because the rigorous research required for policy needs has not been conducted, or because the research that is available is contradictory and does not suggest a single course of action. It is crucial that educational policy decisions are made based on the best evidence possible.

Evidence in Education: Linking Research and Policy brings together international experts on evidence-informed policy in education from a wide range of OECD countries. The report looks at the issues facing educational policy makers, researchers, and stakeholders – teachers, media, parents – in using evidence to best effect. It focuses on the challenge of effective brokering between policy makers and researchers, offers specific examples of major policy-related research, and presents perspectives from several senior politicians. This book provides a fresh outlook on key issues facing policy makers, researchers and school leaders today.

CERI publications

Giving Knowledge for Free: The Emergence of Open Educational Resources (2007)

Understanding the Social Outcomes of Learning (2007)

Understanding the Brain: The Birth of a Learning Science (2007)

Demand-Sensitive Schooling? Evidence and Issues (2006)

Think Scenarios, Rethink Education (2006)

Personalising Education (2006)

OECD*publishing*
www.oecd.org/publishing

ISBN 978-92-64-03366-5
96 2007 08 1 P