The Research Council
of Norway

Rigour and Relevance in Educational Research

Report from the March Seminar 2010

Programme
Norwegian Educational Research towards 2020 - UTDANNING2020

About the programme
Norwegian Educational Research towards 2020 - UTDANNING2020

The Programme for Educational Research – UTDANNING2020 (2009–2018) succeeds the Programme for Knowledge, Education and Learning – KUL (2003–2007) and runs parallel with the Programme for Practice-based Educational Research – PRAKUT (2010–2014). The programme has a total budgetary framework of approximately NOK 344 million.

The UTDANNING2020 programme conducts research on the education sector – from early childhood education and care to doctoral level education. The programme seeks to enhance the knowledge base for policymaking, public administration, professional education and professional practice and to promote research of high scientific merit. A variety of subject areas and research communities are encouraged to conduct research on issues related to the education sector as well as areas of overlap in other sectors.

Programme Board

Kirsti Klette, Professor, University of Oslo (chair)
Lars-Erik Borge, Professor, Norwegian University of Science and Technology
Mikael Alexandersson, Professor, University of Gothenburg
Knud Knudsen, Professor, University of Stavanger
Eivind Heder, Director General, Ministry of Education and Research
Jorunn Dahl Norgård, Senior Executive Officer, Union of Education Norway
Jorun Sandsmark, Adviser, Norwegian Association of Local and Regional Authorities (KS)
Ingrid Helgøy, Research Leader, Stein Rokkan Centre for Social Studies, University of Bergen, (deputy member)
Ingegerd Tallberg Broman, Professor, Malmö University (deputy member)
Observer from the PRAKUT programme

Contents

Rigorous and Relevant Educational Research

Challenges for the new research programme UTDANNING2020

The quality of the educational system and infrastructure is central to every nation's economy, development, social integration and well-being. The quality of education depends and builds on the quality, rigour and relevance of available educational research. It is therefore of critical importance to secure and raise the standards for conducting research in order to improve education. In Norway, the Research Council holds a critical position when it comes to organising and funding educational research. The Research Council has been funding educational research programmes since the mid-1990s, starting with the research programme *Competence, Learning Processes and Value Creation in Worklife* (KUV), which started in 1996, and the evaluation of the 1997 curriculum reform, *Evaluating Reform 97*. However, all research initiatives within the educational sciences have lacked a long-term perspective, sufficient volume for funding and infrastructures that paid attention to processes of synthesising and accumulating research within the education sector. It was therefore a huge step forward when the Ministry of Education together with the Research Council of Norway launched the new research programme *Educational Research towards 2020* – UTDANNING2020. The UTDANNING2020 research programme is designed to address and challenge scientific merits, multidisciplinarity, rigour and relevance in educational research.

The long-term duration of the programme (10 years) facilitates the possibility of addressing factors which are critical to fostering high quality educational research, improving scientific merits within this field of investigation and enhancing the capacity of scholars, enabling them to produce such high quality research.

In order to promote scientific quality and merits, UTDANNING 2020 uses different tools and strategic actions. Funding of high quality research relevant to the educational sciences holds a key position in this tool kit. Through a rich portfolio of varied and intriguing research projects the programme aims to contribute to new insight, accumulate knowledge, support methodological awareness and growth and contribute to fostering research capacity within the educational sciences.

Annual seminars and conferences as mechanisms for knowledge exchange and knowledge building are examples of other activities meant to foster quality in educational research. Within the programme these seminars and conferences are targeting different groups and audiences like policymakers and stakeholders, the teaching profession, researchers and other knowledge brokers.

A special annual seminar –held in March is devoted to addressing scientific and methodological quality in the educational sciences. The first March seminar took place in 2010, and the focus was on rigour and relevance in educational research. This report reflects contributions given at this first March seminar. Scholars from the Scandinavian countries, the US and Switzerland, all from different disciplines, came together to share and reflect upon how we can meet and secure emerging challenges when it comes to promoting quality and relevance in educational research. The six contributions presented in this report discuss rigour and relevance in educational sciences from a variety of perspectives.

Professor Richard Shavelson from Stanford University addresses the question of rigour and relevance in educational research based on his report to the US National Academy of

Science, which he co-wrote with Lisa Town (Shavelson and Town, 2002). A main issue in the *Scientific Research in Education* report (ibid.) is the discussion on how to enhance and secure quality and rigour within research in the educational sciences. What constitutes rigorous scientific research? What is the relevance of educational research, and how might the search for relevance in educational research backfire? In his contribution Shavelson emphasises that quality science is fundamentally the same across all disciplines and fields. Good research is basically about posing significant questions that can be investigated empirically, linking research to relevant theory, using methods that permit direct investigations, providing coherent, explicit chains of reasoning, replicating across studies and encourage professional scrutiny and critique. When it comes to the question of relevance, Shavelson holds a more sceptical position. Relevance is hard to judge, he argues, it is linked to the eye of the beholder rather than an inherent quality of scientific research itself. Shavelson suggests that relevance be placed in the intersection between theory and its usefulness to policymakers and practitioners.

In an interview with Professor of economics Kjell Salvanes and Professor of sociology Knud Knudsen, the two make a plea for extended use of register data and survey data within the educational sciences. The professors think it unfortunate that not more social scientists use empirical data, given the fact that Norway has among the world's best register data available. Used discerningly and in combination with other data sets and resources these registries can provide answers to a vast array of questions, the two professors.

From the position of the Danish Clearinghouse, Professor Sven Erik Nordenbo discusses how evidence-based research pro-

motes quality and rigour in the educational sciences. Nordenbo claims that no type of evidence or research design has exclusive sovereignty as research evidence. All evidence must be judged on the basis of the type of questions it sets out to answer. Nordenbo claims that the evidence-based research movement reflects constructivist perspectives on governmentality in education. The clearinghouse/meta-analysis movement cuts across the old dichotomy between internal evaluation and free research on the one hand, and external evaluation and directed research in educational research on the other. However, the main challenge when synthesising educational research is the difficulty of assessing whether the variables that are studied in one research project are actually similar to the ones studied in another. This notion of "conceptual chaos" in educational sciences makes the ambition of meta-analyses problematic and destructive, Nordenbo argues.

From the perspective of medicine, Professor Olaf Gjerløw Aasland responds to Nordenbo by asking how we can deal with chaotic and non-linear processes for research assessment. Aasland emphasises how synthesising processes and meta-analyses must be combined with professional insight and judgment in the quest for streamlined research evaluation systems. He says not to claim that all interventions need to be validated by a randomised trial, which means: do not bother with measuring the obvious. Check out new ideas, and do not forget the importance of individual and anecdotal information and evidence, he continues. Aasland sees meta-analyses as one of several tools for strengthening the quality of evidence in research.

Professor Dominique Foray and Professor Ingrid Carlgren both explore the question of how educational research could be
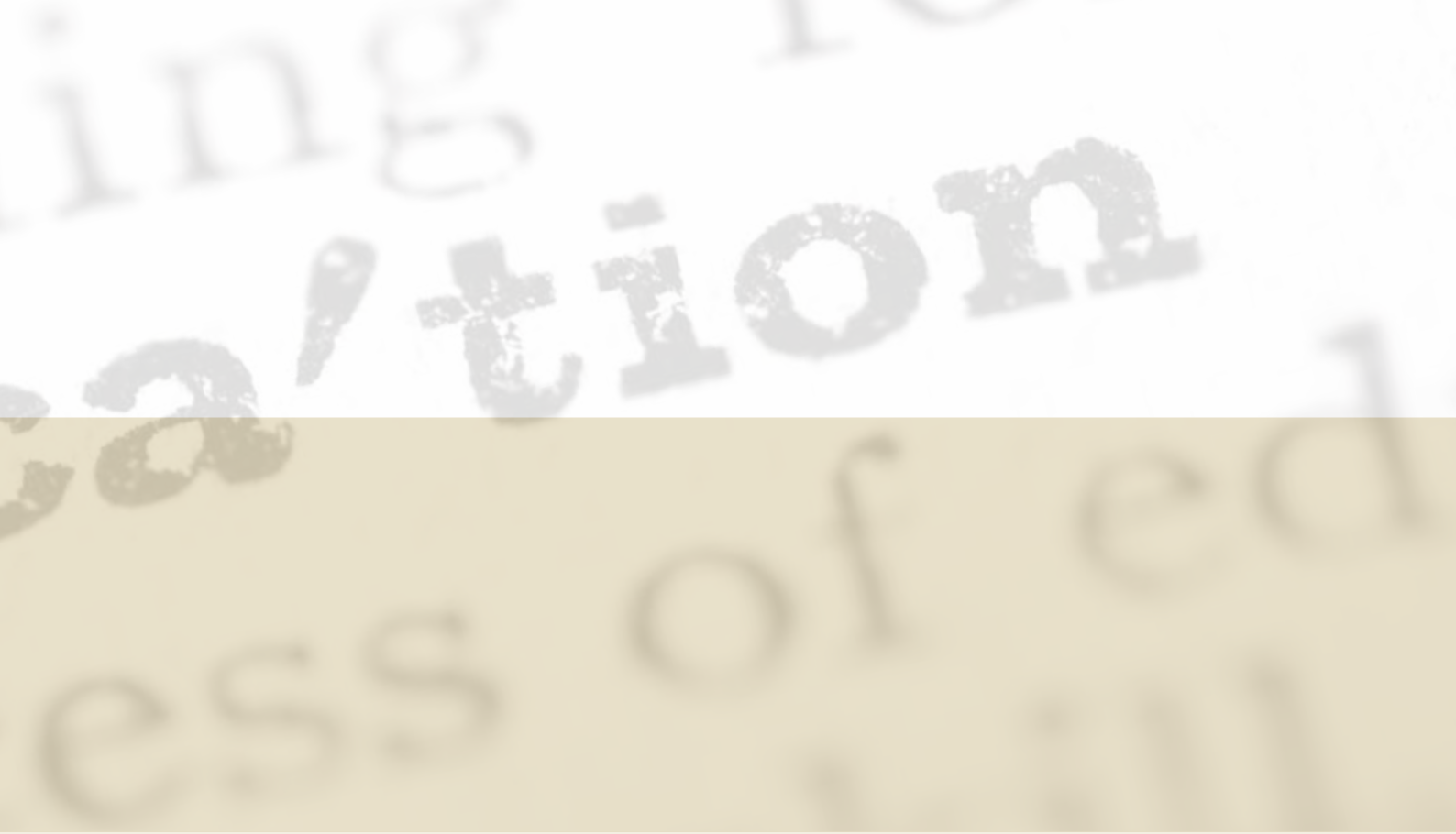
more innovative and progressive, thus contributing to transforming education and the teaching sector into epistemic communities. Carlgren sees two solutions to this dilemma. The first solution is to train teachers so they become deliberate research consumers, while the second is to strengthen the role and volume of clinical, school-based research in education. Her preferred solution is the latter. Carlgren sees the problem in educational research as a consequence of both the volume of research and the focus of research. Lack of funding and resources generate a narrow focus on targeted and detailed studies with the risk of bringing up too many research questions at the same time, she argues. As a consequence, educational research becomes superficial and lacks the applicability and precision necessary for clinical problem-solving at the practice level. Thus the results will not be relevant for teachers or their classroom practice, Carlgren claims. She argues for a differentiation of research activities within our sector and a strengthening of clinical traditions within educational research.

From the field of economics, Professor Dominique Foray highlights the role of innovation when discussing quality and relevance in educational research. Scholars worldwide agree on the crucial role of innovation and mechanisms for disseminating new tools, instructional and institutional practices and technologies when assessing knowledge accumulation. This is, however, not the case for the educational sciences. Educational research and development seldom develops knowledge of immediate value for concrete problem-solving in schools and classrooms, nor do they develop applications, Foray argues. The education sector suffers from innovation deficit and structural limitations as both supply of research and demand for such research in this sector is insufficient and weak.

Information and communication technologies have become a powerful resource for innovation and knowledge growth in the educational sciences, Foray claims. Lack of incentives and market mechanisms, inadequate tools for scaling up, and a reluctant public sector will, however, become bottle-necks when the intention is to derive new and relevant knowledge in the educational sciences from these innovations.

This short report is a plea to marshal all resources necessary to generate the full range of research quality that research within the educational sciences demands. We hope the varied and different contributions here will inspire scholars within the educational field to consider carefully what constitutes rigour and relevance in educational research. And maybe more importantly, we hope that they will want to contribute to and uphold a continuous discussion about quality indicators, the role of evidence and what constitutes analytical and methodological rigour in our field of investigation.

Oslo, January 2011

Kirsti Klette
Chair of the programme board

Reference
Shavelson, R.J. & Towne, L. (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Richard J. Shavelson, Stanford University and SK Partners, LLC

Issues in Conducting Rigorous and Relevant Research in Education [1]

Norwegian educational research is facing challenges when it comes to scientific merit, multidisciplinarity and relevance for practice and policy. The research programme *Norwegian Educational Research towards 2020* – UTDANNING2020 will address these challenges in different ways (The Research Council of Norway, 4 March 2010).

The challenge facing Norway is one that is shared by all countries including my own, the United States. The fact that Norway has committed ten years of research funding not only to improving the rigour and relevance of educational research but also to enhancing the capacity of scholars in Norway to conduct this research is indeed noteworthy and to be applauded.

About ten years ago the US Department of Education recognised a similar challenge:

> There is long-standing debate among scholars, policy makers, and others about the nature and value of scientific research in education and to the extent to which it has produced the kind of cumulative knowledge expected of scientific endeavors. Most recently, this skepticism led to proposed legislation that defines what constitutes rigorous scientific methods for conducting education research (Shavelson & Towne, 2002: 1).[2]

The Department asked our National Academy of Sciences to examine the scientific basis of educational research. The study sought to examine and clarify the nature of scientific inquiry in education and how the federal government can best foster and support it (Shavelson & Towne, 2010).

This paper addresses three issues based on the Academy's report: (1) what is scientific research in education? (2) what constitutes rigorous scientific research? and (3) what is relevance of educational research and how might the quest for relevance backfire? I begin with rigour and discuss principles of scientific research, design of scientific research in education, and disagreements about quality research in education. Then I turn to the question of relevance. Here I stress research that works in the interface between theory and application. Moreover, I warn that relevance is often more in the eye of the beholder (and a product of the political agenda) than an objective, agreed upon quality of research itself.

The nature of rigorous scientific research in education
The Committee on Scientific Principles for Education Research began its deliberations by wrestling with the question of whether scientific research in education differed from scientific research in the social and natural sciences and mathematics. Its initial inclination was to believe that, indeed, scientific research in education differed from other areas of scientific research. Nevertheless, through workshops with scholars in the social and natural sciences and mathematics, none of the distinctions that the Committee considered could withstand the test of careful scrutiny (Towne, Shavelson & Feuer, 2001). In the end, at least as a macroscopic perspective, the Committee (Shavelson & Towne, 2002) concluded that:
- Science is fundamentally the same across all disciplines and fields.
- All fields are characterised by a range of legitimate methods and specialisation depending on the objects of inquiry and context.
- Some differences exist between the social and natural sciences, but they may be more due to different magnitudes of measurement and sampling error than fundamental differences in conducting science.

---

[1] Invited address at the conference on Rigour and Relevance in Educational Research held on 4 March 2010.
[2] Note that the report did not address the question of whether educational research is of poorer scientific quality than scientific research in other disciplines or fields. Rather the report took as its point of departure that all scientific research can be improved and that improvement was the goal of the report.

• As in other fields, features of education, taken together, shape scientific inquiry into education, including:
  - Values and politics
  - Human volition
  - Variability in educational programmes
  - Organisation of education
  - Diversity
  - Multidisciplinarity
  - Ethical considerations
  - Relationships between researchers and those engaged in education

Principles of scientific research in education
Although there is no universally accepted description of the elements of scientific inquiry, the Committee found it convenient to describe the scientific process in terms of six interrelated, but not necessarily ordered, principles of scientific inquiry in educational research. The Committee called these "guiding principles" "... deliberately to emphasize the vital point that they guide, but do not provide an algorithm for, scientific inquiry" (Shavelson & Towne, 2002: 52). The principles provide a framework as to how inferences in general are to be supported (or refuted) by a core of interdependent processes, tools and practices. These principles are:

• **Pose significant questions that can be investigated empirically**—science proceeds by posing important questions about the world with potentially multiple answers that lead to hypotheses or conjectures that can be tested and refuted; these questions must be posed so that it is possible to test the adequacy of alternative answers observationally.

• **Link research to relevant theory**—much of science is fundamentally concerned with developing and testing theories, hypotheses, models, conjectures or conceptual frameworks about the physical or social world. Theory guides research and research leads to revision of theory; they are inextricably interconnected.

• **Use methods that permit direct investigation of the question**—research methods, the design for collecting data, and the measurement and analysis of variables in the design should be selected in light of the research question *(not vice versa)*. Methods linked directly to problems permit the development of a chain of logical reasoning from question to method to interpretation.

• **Provide coherent, explicit chain of reasoning**—the extent to which inferences made in the course of scientific research are warranted depends on rigorous reasoning that systematically and logically links empirical observations with underlying theory and the degree to which both the theory and the observations are linked to the questions or problems that lie at the root of the investigation.

• **Replicate and generalise across studies**—replication and generalisation strengthen and clarify the limits of scientific models and theories. While challenging in education, scientific research needs to provide evidence of replicability and generalisability.

• **Disclose research to encourage professional scrutiny and critique**—scientific scrutiny and criticism of research is essential to conducting science and accumulating knowledge. In its absence, research findings take on a life of their own and enter public debate misleadingly.

Designs for scientific research in education
Controversy surrounds the design of educational research and the appropriate methods for collecting and analysing data. At the time of the Committee's deliberations, George W. Bush had just been elected to a first term as President of the United States. As part of his education reform agenda (known as "No Child Left Behind"), the US Department of Education intended
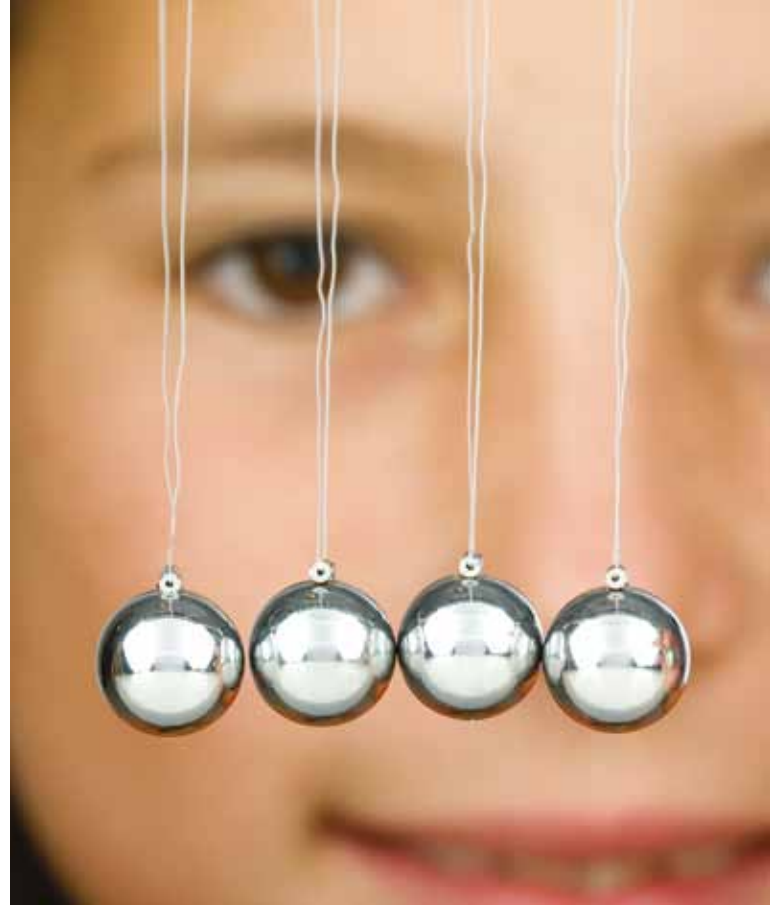
to define scientific research not by a set of principles such as the guiding principles of the Academy's report but by a particular research design: randomised controlled experiments. At the extreme, for the federal government what made research in education scientific was that it used randomised experiments.

This characterisation, in the view of the Academy Committee, was far too narrow. Moreover, the definition of scientific research in education as randomised experiments was a clear case of the "tail wagging the dog." Scientific research employs the set of guiding principles just enumerated and is guided by important research questions, not methods. To let a particular research method define and drive scientific research in education was, in the Committee's view, to get the enterprise backward.

The Committee, then, made it clear that the choice of research method should follow logically from the nature of the question driving the research. That is, the research method chosen must fit the research question posed, not vice versa. To this end, the Committee identified three overarching questions that scientific research pursues: What is happening? Is there a systematic (causal) effect? And how or why is it happening?

• What is happening?
  - Estimates of population characteristics (e.g. percentage of students in the US performing at or above proficient level in mathematics achieve ment)
  - Simple relationships (e.g. correlation between income and achievement)
  - Descriptions of localised educational settings (e.g. eth-nographic study of school and community from the view of students living in poverty)

• Is there a systematic (causal) effect?
  - Causal relationships when randomisation is feasible (e.g. students randomly assigned to experimental and control groups where the former receives self-affirmation in a stereotype threat situation and the control receives other-affirmation)
  - Causal relationships when randomisation is not feasible (e.g. comparison of "traditional" and "reform" reading programmes with extensive information on characteristics of students collected before the study was conducted)
• How or why is it happening?
  - Mechanism underlying causal effect when theory is fairly well-established (e.g. mini-experiments or indepth obser-vation testing mechanism hypotheses)
  - Mechanism underlying causal effect when theory is weak (e.g. exploratory design studies attempting to identify the mechanism)

The Committee pointed out that different types of questions lead to different research designs, data collection, analyses and inferences. The committee noted that random controlled trials were appropriate and the "gold standard" for addressing causal effect questions when feasible and ethical. However, there are many cases where such trials are premature, too expensive, unfeasible or unethical and in those cases there were other designs that could be used to test for causal effects (e.g. Murnane & Willett, 2009). Moreover, the Committee noted that qualitative (e.g. ethnographic) research was just as legitimate as quantitative research; at issue is the logical chain of reasoning from research question to method to data collection to inference, not the method itself.
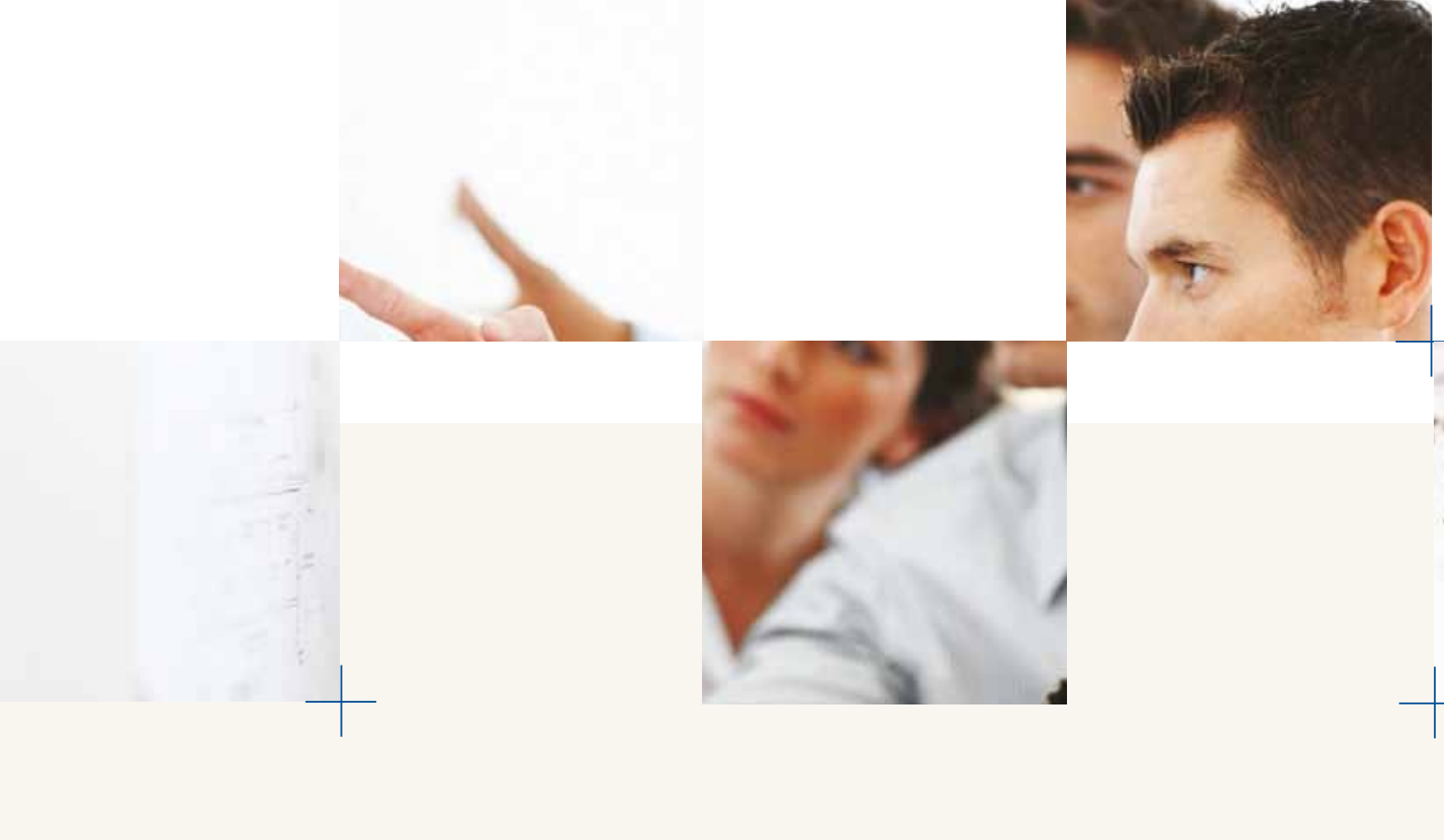
Before turning to the relevance-of-education-research question, I should point out that the Academy's report was quite controversial at the time and still is. Some praised it while others damned it. A 2009 issue of the *Educational Researcher* entitled "Learning from Our Differences: A Dialogue across Perspectives on Quality in Education Research" (Moss, Phillips, Erickson, Floden, Lather, Schneider, 2009) made this clear. To simplify, consider a continuum ranging from a unified and generalisable view of educational research to an interpretive view. The generalisable view holds that scientific research involving prediction and explanation of educational phenomena is possible and fruitful. The other view claims that all research is interpretive and inferences can only be made when in-depth understanding of meaning in context can be achieved—generalisation is illusive. Seven years after the publication of *Scientific Research in Education* the issues surrounding what constitutes quality educational research remain contentious.

Relevance in educational research

Everyone agrees that educational research needs to be relevant. But not everyone agrees as to what constitutes relevant research for which audience and for what purpose.

Some have argued that relevance should not be a criterion for scientific research, including research in education. Such a requirement would stifle discovery, creativity and innovation (see Stokes, 1997). I believe Stokes (1997) presents ample evidence to the contrary.

I would argue that relevance is not a property inherent in scientific research (in education). Rather, relevance is relative to the prevailing cultural, societal, and political context operating at the time the research is being carried out.[3] What is relevant

to some audiences for some purposes today may not be so tomorrow. And what is relevant to some audiences is irrelevant (or even worse) to another audience holding fundamentally different values and beliefs. A case in point is the current emphasis on achievement testing in the US. This emphasis has been spurred by the fiscal incentives offered by President Obama's administration. All of a sudden, what might be considered cutting edge but irrelevant research in assessment and psychometrics has now become extraordinarily relevant, both in pre-college and higher education in the United States. Relevance, then, depends on the research question (and answer!), the prevailing context, and the audience with the power to find it useful (fitting an agenda) or not (inconsistent with an agenda).

Having put relevance in perspective, I strongly believe that educational researchers and policymakers should not demand relevance. Indeed, I believe that much can be done to increase the relevance of educational research to policy and practice. I build my case on the work of Donald Stokes (1997).

In his review of path-breaking scientific discoveries, Stokes (1997) provides the key to relevance. Asking whether scientific breakthroughs followed the pattern of moving from basic to applied research, Stokes concluded that the firmly ingrained belief was not, in fact, the predominant pattern. Rather he found that path-breaking scientific research was a consequence of two interacting factors: the quest for fundamental understanding (theory building) *and* considerations of use (practice). This led him to sketch the now famous Pasteur's quadrant to show his point (Figure 1). The quadrant was dubbed Pasteur's quadrant because Pasteur sought not

---

[3] A similar conclusion has been reached, for example, by scholars involved in climate change research and its use for policy-making (e.g. Schneider, 2009).

just a theory of disease in his research but also the prevention of disease. In Stokes' words (1997: 63):

Work directed toward applied goals can be highly fundamental in character in that it has an important impact on the conceptual structure or outlook of a field. Moreover, the fact that research is of such a nature that it can be applied does not mean that it is not also basic.

Note that the quadrant representing basic research with little application is dubbed Bohr's quadrant, as Neils Bohr was engaged in very basic research on the structure of the atom. The quadrant that represents virtually sole focus on use with little theory is dubbed Edison's quadrant. Thomas Edison was an engineer who focused on solving practical problems. Finally, no one occupies the null quadrant, as might be expected.

Quest for Fundamental Understanding		
	Bohr	Pasteur
		Edison
	Bohr	

Figure 1. Stokes' quadrants of scientific research.

For me, then, relevance falls at the intersection of theory building and use. More specifically, much of educational research should fall at this intersection. Note, however, that what might be of concern to practitioners and policymakers today with regard to use may not be of concern tomorrow. This said, there remain fundamental questions at the intersection of theory and use—such as the education of a diverse student body—

that are as enduring as finding ways to prevent milk (and wine!) from causing illness through a pasteurisation process. And it is these questions which fall at the intersection of understanding and use that the Academy's committee believed to be the important questions and the ones that should constitute the basis for relevance in educational research.

Relevance: a two-edged sword
Relevance, especially in an area such as education—where beliefs and values define, for example, what constitutes the "good life" and how and what children should be taught— can be a two-edged sword. Producing research that fits societal or political expectations may fall prey to manipulation. That is, relevance depends on the intersection of research and context. Research can be dubbed relevant or irrelevant and be used or abused depending on various factors such as political agendas. For example, Tom Colburn, a Republican senator from Oklahoma, proposed prohibiting the National Science Foundation from "wasting any federal research funding on political science projects" because he believes the research to be irrelevant and expensive. Political scientists rallied in opposition. They pointed out that one of the year's Nobel Prize winners, Elinor Ostrom, had been a frequent recipient of the very programme now under attack by Coburn.

My concern, at least some part of it, is captured in Figure 2. The concern is that research can get warped to fit the procrustean bed of values and beliefs rather than values and beliefs being modified by scientific findings. In what follows, I provide several examples in addition to that of Coburn where values and relevance conflict with scientific research.

Figure 2. When data are forced to fit theory, relevance becomes a two-edged sword.

I draw the first example from scientific research on the impact of charter schools in the US on students' achievement. As is well known, people in the US have strong and deeply ingrained beliefs and trust in the "power of economic markets" to lead to beneficial (at least profitable) outcomes. Of course, not all Americans hold this belief, but it surely is prevalent. Charter schools have been proposed as a palliative to what ails US education. Charter schools—public schools freed of most of the restrictions of educational codes and teachers' unions—fit the power of markets belief. These schools are viewed as bringing competition and choice into education. Regular public schools would have to compete with charter schools for students. This competition, so the argument goes, would lead regular public schools to shape up and improve student achievement or go out of business because students could, in theory, go to charter schools (if there were enough of them). As we shall see, even if charter schools do not produce achievement differences, school choice is to be valued above scientific evidence; scientific research is irrelevant.

In reality, substantial evidence shows that charter schools, even given their freedom, are no better on average than regular public schools, and they even may not be as good at producing mathematics and reading achievement. Charles Murray, a co-author of *The Bell Curve* and conservative policy researcher, admitted to the accuracy of my interpretation of these findings. He went on to say in a New York Times op-ed article from 5 May 2010:

> ...all I can say is thank heavens for the Milwaukee results. Here's why: if my fellow supporters of charter schools and vouchers can finally be pushed off their obsession with test scores, maybe we can focus on the real reason that school choice is a good idea. Schools differ in what they teach and how they teach it, and parents care deeply about both, regardless of whether tests scores rise.

Murray went on to say that even knowing that charter schools may not produce better achievement test scores than regular public schools:

> I would still send my own children to that charter school in a heartbeat. They would be taught the content that I think they need to learn, in a manner that I consider appropriate... The supporters of school choice need to make their case on the basis of that shared parental calculation, not on the red herring of test scores (New York Times, 5 May 2010). Theory – in this case values – trumps scientific data when convenient.

A second example of relevance being defined by values, expressed as theoretical expectations (or in this case legal interpretation), is drawn from the University of Michigan affirmative action case. For years US colleges have argued for a policy of affirmative action—having somewhat different criteria for admitting underserved children, often with a minority background, who have grown up in conditions characterised by poverty and discrimination. These children are given credit for their accomplishments, perseverance and resilience under unequal societal conditions. However, the American public disagrees with this policy on the grounds that it discriminates against other children (e.g. non-minority children growing up in poverty). The debate is heated, and California abolished the University of California's affirmative action programme in a statewide vote.

The National Association of Scholars, a conservative organisation of university faculty, filed a lawsuit on behalf of plaintiffs who sought to overturn the University of Michigan's affirmative action policy for undergraduates. The Association argued that there was no empirical support for the direct effect of affirmative action on valued student outcomes (e.g. achievement, self-esteem). Of course the Association's claim turned on their interpretation of legal precedent. The data showed that structural diversity (the number of minorities on campus) was not statistically related to outcomes (Figure 3). Moreover, according to the Association's interpretation, legal precedent required this relationship to be present. This interpretation

supported the association's prior beliefs that affirmative action was an improper policy. Specifically, the Association argued that the University of Michigan's research showing that affirmative action did have a positive effect on outcomes was arrived at through an improper statistical analysis. The two positions are depicted in Figure 3. It turns out that the University of Michigan's modelling of the data was statistically proper. Having more minority students on a campus does not directly affect outcomes. Rather, diversity operates through the educational experiences that such diversity affords to produce positive outcomes. Whose relevance, then, counts?
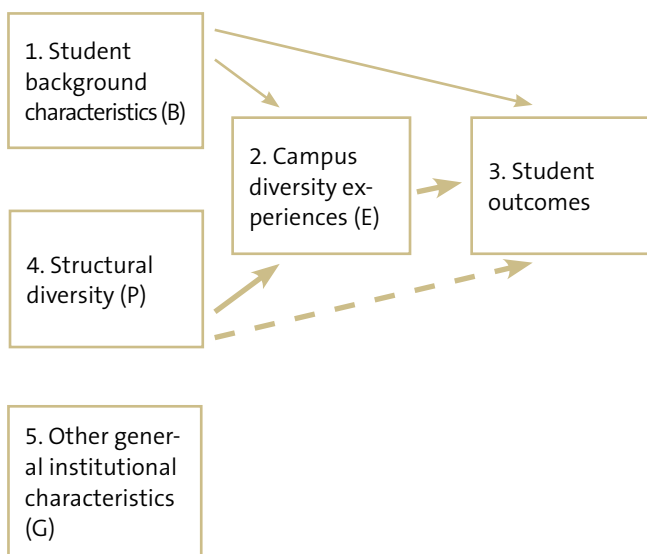

Figure 3. Competing statistical models showing the impact of affirmative action (structural diversity) on student college outcomes.

Concluding comments

Like all countries, Norway faces the challenge of determining and supporting the scientific merit, multidisciplinarity and relevance of educational research as a means of improving practice and policy. Norway's commitment of ten years of funding to support and improve the quality of educational research as well as to enhance the education and training of researchers from multiple disciplines is praiseworthy. I believe it is possible to conduct rigorous scientific research in education. To this end, I have set forth the findings and advice arrived at by the Committee of the National Academy of Sciences in the US that may be useful to the Norwegian programme. I also believe it is possible to improve the relevance of educational research for policy and practice. My recommendation is to support research at the intersection of theory and usefulness to policymakers and practitioners. However, I have also warned that relevance is context bound. Moreover, it can be used as a two-edged sword where scientific research can inform beliefs and values or where beliefs and values can warp scientific findings. For the next ten years I will be a keen observer of the "Norwegian Experiment". I wish it great success; the world stands to learn from what you learn.

References

Moss, P.A., Phillips, D.C., Erickson, F.D., Floden, R.E., Lather, P.A., & Schneider, B.L. (2009). *Learning from our differences: A dialogue across perspectives on quality in education research,* 38(7), 501-517.

Murnane, R. J., & Willett, J.B. (2010). *Methods matter: Improving causal inference in educational and social science research.* NY: Oxford University Press.

Schneider, S.H. (2009). *Science as a contact sport: Inside the battle to save Earth's climate.* Washington, DC: National Geographic.

Shavelson, R.J., & Towne, L. (2002). *Scientific research in education.* Washington, DC: National Academy Press.

Stokes, D.E. (1997). *Pasteur's quadrant: Basic science and technological innovation.* Washington, DC: Brookings Institution Press.

Towne, L., Shavelson, R.J., & Feuer, M.F. (2001). *Science, evidence, and inference in education: Report of a workshop.* Washington, DC: National Academy Press.

Illustrasjon: Jon Solberg, Forskningsrådet

Foto: Helge Hansen

Siw Ellen Jakobsen

Social scientists need to consult the hard figures

Although Norway has perhaps the world's best registry data, much of the knowledge produced by Norwegian social scientists is based on interpretation. A sociologist and an economist take a closer look at this.

"Important sources of knowledge are seriously under-utilised in Norway," said Professor Kjell G. Salvanes of the Norwegian School of Economics and Business Administration in his presentation at the kick-off seminar for the Research Council's new educational research initiative, the UTDANNING2020 programme.

Our insight into the Norwegian schools is still rather limited. Dr Salvanes fears that educational research could be missing out on an exceptional opportunity to learn more.

"We know that teachers are important, but we don't know what makes a good teacher. We know that the drop-out rate from upper-secondary school is high, but we don't know what happens to those who drop out. The answers to these questions may be found in the unique registry data that we have ready and waiting in Norway. It is possible to conduct high-level research on this in Norway and Scandinavia. The biggest problem is 'selling' these ideas to educational researchers," said Dr Salvanes.

The professor thinks it is unfortunate that more social scientists do not use empirical data given that Norway has among the world's best data registries. Used discerningly, and preferably in combination with other sources, these registries can provide answers to a vast array of questions.

Dr Salvanes was invited to discuss this issue with Professor Knud Knudsen, a member of the UTDANNING2020 programme board.

Dr Knudsen has years of experience as a sociologist. Together with Gudmund Hernes, he wrote one of the key reports from the first Norwegian survey of living conditions on the topic of education and inequality. Since then he has held a position at the University of Bergen, and he is currently employed as a professor of sociology at the University of Stavanger.

Dr Knudsen, do you agree with the Dr Salvanes that the social sciences in Norway have based their scholarship mostly on interpretation that is not grounded in empirical data?

Dr Knudsen: I think Dr Salvanes is on to something, but I want to refine his point. A lot of good quantitative research is being conducted, including in Norwegian sociology. The use of registry data is on the rise, and certain groups of sociologists are quite active in this regard. But I think that segments of the Norwegian social science community continue to be hampered by the debate about positivism in the 1970s and influenced by the later wave of post-modernism. There are conflicting points of view about the most important research questions and the most useful methodologies. Of course these issues are important, but a lot of time is spent discussing them over and over again. I don't think the problem lies in the distinction between quantitative and qualitative methods; both approaches are valid. But some researchers have landed in between the two methods. They conduct studies using a very small sample but want to make broad generalisations at the same time. It is difficult to accumulate knowledge and – not least – it is difficult to extrapolate clear implications for public policy when this knowledge is based on sparse data.

Dr Salvanes: In my opinion, the main focus should be changed. There have been major developments in methodology in recent years that can help to distinguish between various explanations of a phenomenon. Large segments of the Norwegian social science community have not incorporated these into their work, not even many of the researchers who actually use registry data. This was made evident in the popular scientific programme on biology and society, *Hjernevask* ("Brainwash"), broadcast on Norwegian public television in the spring of 2010. Even more important than the fact that gender researchers tripped up is that key social researchers – including those who are oriented more towards empirical analysis – believe that we cannot say anything of import about what we cannot observe. As a result, weighty conclusions are drawn without considering alternative explanations. It is a weakness that these researchers do not take an interest in methodologies that along with better data sources will help them to distinguish between various explanations of their findings.

Do you believe that Norwegian social science has been in decline?

Dr Salvanes: Yes, in some areas. The University of Bergen, where I studied sociology, had a good, empirically oriented social science community. The pioneering researchers there formulated and tested alternative explanations. In the 1990s the social sciences changed their focus from facts, data and guarded explanations to understanding and interpretation.

The Research Council is now seeing a change back again. The organisations that commission and pay for research want more knowledge based on registries and numbers. Is there a danger that we will once again be too single-mindedly focused on facts and figures?

Dr Salvanes: The use of registry data definitely has its limitations. We can by no means find all the answers there. In many cases registry data must be supplemented with other data sources, including qualitative data. But registry data provide a good opportunity to accumulate knowledge over time and across disciplines since the data apply to the entire population. In my view, sociologists, anthropologists, psychologists and economists all have sound traditions and they should serve as an inspiration for each other. This applies to subjects outside the social sciences as well. It is good to employ different perspectives in research, but the problem is that so many social scientists are not interested in looking at alternative explanations for their findings and as a result they are in great danger of over-interpreting the results.

Dr Knudsen: I was trained in a tradition that said, "Give me an interesting empirical finding and give three alternative explanations for it" or "Give me an interesting theoretical statement and give me three different empirical implications of it". In other words, it is important to have competing points of view and supplementary types of data. This is why diversity in theories and methods is needed.

This is precisely what you and Gudmund Hernes did in your well-known report on education and inequality from 1976?

Dr Knudsen: We used survey data, occupational history data and the first survey of living conditions together with

simulations, and we looked for competing strategies for explaining phenomena. To explain differences in education, we examined socialisation in the home, but we also took the significance of abilities and heredity seriously. I hope that someone today will attempt a similarly comprehensive analysis of the way in which the educational system functions, but using all the data sources and analytical techniques that are now available.

Dr Salvanes: Well, this is the tradition I was trained in during my studies in sociology, and it is the main focus of the methodological revolution in the use of microdata in economics taking place today. Taking into account, and possibly distinguishing between, explanations is not merely an academic discussion. It is crucial when interpreting and examining the implications of the results. An example of this is education and educational policy. Many researchers from various social sciences, including myself, have found that family background, such as the parents' education, can explain a lot about children's behaviour and how they turn out as adults. Prominent social researchers in Norway interpret this to mean that socialisation in the home is the reason that children of highly-educated parents tend to pursue a higher education.

Can't this be the case?

Dr Salvanes: Yes, but there are other explanations as well. By utilising reliable data sources and considering whether to use experiments, it is possible to come far in distinguishing the effect of the parents' education from other characteristics of the parents. Registry data can be used in a clever way to draw

on information about twins, adopted children and real-life experiments such as educational reforms and the like. Social researchers have had a strong tendency to interpret everything as a social construct. They leap directly to one explanation of a phenomenon, but they need to explore alternative explanations as well!

Dr Knudsen: Our research clearly showed that not only socialisation but also a person's abilities are important for explaining inequalities in education. At the time, the evidence that education plays a role in reproducing inequalities caused strong reactions. A lot of bureaucrats and educational policymakers were sceptical. Many were convinced that social democracy would necessarily result in more equal opportunities. Then our report came out showing that this was not the case. It caused an uproar, but we couldn't change what the empirical data revealed. That is the nature of research."

Dr Salvanes: What will it mean if everything can be attributed to socialisation? It means that everything can be fixed. Then we can apply policy in all areas. If parents are given an education, then they will socialise their children themselves. But obviously some explanations are more difficult to do something about. The problem is that some social scientists tend to interpret society in the direction they would like to see it develop.

Dr Knudsen: It is essential that others are given insight into the research process and that the findings are reviewed with a critical eye. Alternative hypotheses must be given the greatest chance possible! We need to revitalise this attitude. Today we

are seeing a movement in the direction of more empiricism and quantitatively oriented social research. This is an interesting pattern which is emerging in the UTDANNING2020 programme, where I sit on the programme board. Here it appears that most of the larger projects have incorporated relevant quantitative data. The portfolio contains well-grounded qualitative studies as well.

Why is this movement occurring now?

Dr Knudsen: I think it is driven mainly by the researchers themselves. Dr Salvanes and other economists have made some interesting analyses based on information from the data registries. Others have followed in their footsteps. Both researchers and research administrators have come to understand that it is possible to extract rather precise information from relevant registry data.

Dr Salvanes: In the mid-1990s I was at Statistics Norway (SSB) with a colleague and good friend of mine, the late Tor Jakob Klette. That was when we first learned about SSB's large datasets. Nobody in the economics community had ever used these in a systematic way. Both of us used these data, first as the basis of a joint project and later for larger projects when he was employed at SSB and I worked at the Norwegian School of Economics and Business Administration. We were motivated by curiosity and our interest in the field. There was also a movement taking place abroad in this field, and there is no doubt that our research stays in other countries had an impact on us. People realised that it is possible to study a great deal by linking together various datasets.

Norway has the world's best data registries, but do researchers have sufficient access to them?

Dr Knudsen: Let's create the world's best system for access to data! For many years Norwegian social researchers have had straightforward access to ordinary survey data via the Norwegian Social Science Data Services (NSD). This access does not depend on a person's position level or resources: a master's student, a doctoral student and a professor are on equal footing. Everyone gets a quick response and efficient service. What is important in the future is to organise something similar for registry data. There are a variety of good reasons why this has taken time. Complexity is one thing; considerations relating to personal privacy are another. But both SSB and NSD are now working constructively together on this, and I'm certain that their efforts will result in effective schemes.

Dr Salvanes: Many research groups have now had experience working with these data. We know the data are reliable, and we also know their limitations. Many groups today have specialised knowledge about these data. Now is the time to make them much more accessible.

Dr Knudsen: The Swedes are facing the same challenges as us. They have excellent registries and biobanks that researchers do not fully utilise. The Swedish Research Council has established a large-scale programme focusing on registry data. Under the Swedish Initiative for Research on Microdata in the Social and Medical Sciences (SIMSAM), researchers from a variety of subject areas consolidate their expertise and knowledge about registry-based research and disseminate this through consultancy services, courses and conferences. In order for

Norway to boost its use of registry data for research purposes, a similar large-scale initiative to raise the level of expertise in needed, both with regard to advanced methodologies and the adaptation of data.

Dr Salvanes: Previously the research groups had to be of a certain magnitude in order to be permitted to access and use these data. At the Norwegian School of Economics and Business Administration we have used several person-years to learn how to use registry data. The costs have been high, but the investment has also yielded high returns in the form of increased knowledge about how working life and community life function. To increase the overall level of competence in handling registry data, the data must be made more accessible and at the same time there must be a plan for how the educational institutions can upgrade their knowledge about methodology.

Dr Knudsen: Norwegian researchers still make too little use of the large international, comparative databases that they have access to through NSD. Political scientists and sociologists probably assume that these data are more difficult to deal with than they actually are. Norwegian researchers can learn a lot from cross-national analyses. For example, the Research Council is investing substantial sums in the fantastic data from the European Social Survey, but in my opinion Norwegian social researchers could use these data more frequently. This is a challenge that the research communities must address.

Translated by Connie Stultz and Carol B. Eckmann.

Sven Erik Nordenbo, Danish Clearinghouse for Educational Research, Aarhus University, Denmark

Evidence and Synthesis: [1]
A New Paradigm in Educational Research

Introduction

A paradigm shift is haunting educational research: the spectre of evidence-based practice and policy. As with most paradigm shifts it causes uneasiness and resistance, and as with other changes of paradigm we see that at *first* it is received with a shake of the head – because it is believed that this paradigm makes no sense. Then *later it* is rejected – because it is believed that although it does make sense, it is flatly wrong. And *finally* – in the third and last phase – all and sundry say: "Is this perhaps a novelty?"

In the area of educational research in Denmark we have chosen to name the unit set up to work on the basis of this new paradigm the *Danish Clearinghouse for Educational Research.*
• I will start by commenting on the concept of evidence and give a few examples of what we have been working on during the last three years at the Danish Clearinghouse for Educational Research.
• Having dealt with this, I will look at important concerns and critical comments that have been put forth about this new paradigm.
• Finally, I will conclude by mentioning some of the profes sional challenges that we face when we attempt to exploit research of the type that the Danish Clearinghouse for Edu-cational Research is working to produce.

What does a clearinghouse for educational research do?

In his best-selling book *Dr Spock's Baby and Child Care*, Dr Benjamin Spock wrote, "I think it is preferable to accustom a baby to sleeping on his stomach from the beginning if he is willing". This statement was included in most editions of the book and in most of the 50 million copies sold from the 1950s to the 1990s.

During this same period, more than 100,000 babies died of sudden infant death syndrome (SIDS), also called *crib death* in the United States and *cot death* in the United Kingdom, where a seemingly healthy baby goes to sleep and never wakes up. In the early 1990s, researchers became aware that the risk of SIDS decreased by at least 50 per cent when babies were put to sleep on their backs rather than face down. Governments in various countries launched educational initiatives, which led to an immediate and dramatic drop in the number of SIDS deaths.

While the loss of more than 100,000 children would be un-speakably sad under any circumstances, the real tragedy lies in the fact that many of these deaths could have been prevented. Gilbert, Salanti, Harden & See (2005) write:

> Advice to put infants to sleep on the front for nearly half a century was contrary to evidence available from 1970 that this was likely to be harmful. Systematic review of preventa-ble risk factors for SIDS from 1970 would have led to earlier recognition of the risks of sleeping on the front and might have prevented over 10,000 infant deaths in the UK and at least 50,000 in Europe, the USA and Australasia.

This example is one of several cited by Sir Iain Chalmers in a talk entitled "The scandalous failure of scientists to accumu-late scientifically" (Chalmers, 2006). The theme of this talk was that we live in a world where the utility of almost any intervention will be tested repeatedly and that rather than

---

[1] Keynote speech in Oslo on 4 March 2010 by invitation from the Research Council of Norway (revised version 31 October 2010).

looking at any study in isolation, we need to look at the body of evidence (cf. Borenstein et al., 2009: xxi).

I repeat the last sentence:

Rather than looking at any study in isolation, we need to look at the body of evidence.

That is what we at the Danish Clearinghouse have attempted to do in our research from our establishment in 2006 until today.[2] The procedure is in essence very simple. We distinguish between primary empirical research and secondary research. Primary, empirical research examines "reality" – the real world – and attempts to use conventional research methods to achieve insight into this reality. A decisive factor for the quality of this research is the relationship between the type of question that is being posed and the research designs employed to address the question. Petticrew & Roberts (2003, 2006) have given a good account of the relationship between research question and research design; see Table 1.

Research question	Qualitative research	Survey	Casecontrolstudies	Cohort studies	RCTs	Quasiexperimental studies	Non experimental evaluations	Systematic reviews
Effectiviness Does this work? Does doing this work better than doing that?				•	••	•		•••
Process of service delivery How does it work?	••	•					•	•••
Salience Does it matter?	••	••						•••
Safety Will it do more good than harm?	•		•	•	••	•	•	•••
Acceptability Will children/parents be willing to or want to take up the service offerd?	••	•			•	•	•	•••
Cost effectiveness Is it worth buying this service?					••			•••
Apprpriateness Is this the right service for these children?	••	••						••
Satisfaction with the service Are users, providers, and other stakeholders satisfied with the service?	••	••	•	•				•

Table 1: Typology of Evidence

---

[2] The Danish Clearinghouse for Educational Research was established following recommendations by the OECD, cf. Ekholm (2003) and Mortimer (2004).

As this survey shows, no research design has a higher intrinsic status than another. On the other hand, it is critically important to understand what type of question one wishes to answer.

Primary empirical research establishes its data by looking at observable relationships in "reality". What Sir Iain Chalmers is referring to, however, can be defined by a number of labels: "accumulated research", "meta-research", "second-order research" or "systematic review". In all of these, one assumes that what is accumulated is the data and results of first-order research.

How is this "second-order" data presented in "reality"? In the form of published research.

Just as primary empirical research starts out with a research problem, meta-research also starts out with a problem. When we collect data from primary research, this is done by "reviewing" the research reports from primary research. So the problem formulation for a meta-research task is usually called the "review question".

The first task after formulating a review question is therefore to obtain the data for the study - the reports or documents from the primary research. I will not discuss the techniques developed for doing this here.

After obtaining the documents, they are subjected to a systematic procedure that has two basic phases: (a) a professional evaluation of the registered documents with respect to their *relevance* and *quality* for the "review question" and (b) a *synthesis of the evidence* that the qualified documents present. As mentioned, the result of this procedure is a "systematic review". Frequently the answers to the review question posed are then formulated as statements, that there is now evidence for this or that.

The *concept of evidence* has sometimes given rise to misunderstanding, so I will devote a few words to discussing the various ways this concept can be interpreted in (1) primary research, (2) the philosophy of science and (3) meta-research.

1. Scientists frequently disagree about whether, or to what extent, a given set of data or observational results constitute evidence for a scientific hypothesis. Disagreements may be over empirical matters, such as whether the data or observational results are correct, or whether other relevant empirical information is being ignored.

In this context, the use of the concept "evidence" is not particularly controversial, and its meaning remains a concrete problem within the boundaries of the research community in question. It is of course vital to agree as to which data can be accepted as evidence in a specific research project, but the use of the term itself is not controversial.

2. However, conflicts also arise about the concept of evidence because philosophers of science employ incompatible concepts of evidence in their theory-making. I will give just a few examples of such central theories: (1) the causal-inductive theory of evidence (2) the theory of "inference to the best explanation" (3) the theory of falsificationism (4) probabilism and (5) evidence nihilism (Achinstein, 2005: 1-5).

The principal elements in these theories of evidence are shown in the following textboxes:

---

(1) The causal-inductive theory of evidence:
- Expounded by Sir Isaac Newton and John Stuart Mill, among others.
- Only deduces the sufficient number of causes necessary for explaining the phenomenon.
- The same effect arises from the same cause.
- If all observed phenomena have the same characteristics, then all such phenomena (in the universe) have them.
- All conclusions about them are true until proven otherwise.

---

(2) The theory of "inference to the best explanation":
- Propounded by William Whewell, W.V. Quinne and others.
- Scientists generate hypotheses from observations and guesswork.
- Three requirements:
  - Explain the hypothesis — and predict?
  - Does it predict new phenomena?
  - Does the system become simpler, more coherent or unified?

---

(3) The theory of falsificationism:
- Propounded by Karl Popper in particular.
- Yes, scientists guess at hypotheses.
- Observations do not prove hypotheses.
- We can prove that hypotheses are false, never that they are true.

---

(4) Probabilism:
• Developed during the 20th century. "Scientific evidence is to be understood completely by reference to mathematical probability." Also known as Bayesianism.
• Scientific evidence refers to mathematical probability.
• A set of data constitutes evidence if it increases the probability of a hypothesis.
• Explanations do not need to be causal or explanatory.
• Adherents: many statisticians and economists.

(5) Evidence nihilism:
• Propounded by Paul Feyerabend in particular.
• Any rule of evidence must be broken in the search for new theories.

This is a yet unresolved conflict in the philosophy of science, but it is perhaps worth noting that nowadays we increasingly observe that probabilism is making advances as the dominant frame of reference for evidence within research into education and educational policy.

3. In meta-research the concept of evidence has a more definite meaning. The available primary research is considered to be the source for establishing evidence since the requirement is (in a slightly modified version of the words of the Canadian researcher David Sackett):

> ... to carry out a systematic and considered utilisation of the best available evidence from educational research so as to make practice and policy evidence-based.

In meta-research the task therefore becomes to show that existing primary research results contain arguments for shaping daily practice and policies based on insights that have already been achieved – if one goes to the trouble of finding them, as in the case of sudden infant death syndrome (SIDS).

The Danish Clearinghouse for Educational Research has participated in this international task from a Scandinavian perspective, which manifests itself in two particular ways: the review questions that have been established are formulated on the basis of specifically Scandinavian requirements, and whenever possible we have included research published by Scandinavian researchers who work on problems in the Scandinavian countries.

I will conclude this section about what a clearinghouse in educational research attempts to do by mentioning four of our completed projects:

a) Can we identify which teacher qualifications improve pupils' learning? Yes, there are three central groups of competencies: competence in relationships, competence in rule-based leadership, and didactic competence.

b) Can centrally organised tests, referred to as "national tests", be useful to teachers? We do not know much about this in terms of research results! But it is worth discussing whether the activity of testing harms more than it helps – especially if the perspective is focused on the more poorly performing pupils.

c) Can we identify those school factors that assist learning (in a broad sense) at the school? Yes, 11 factors are particularly important.
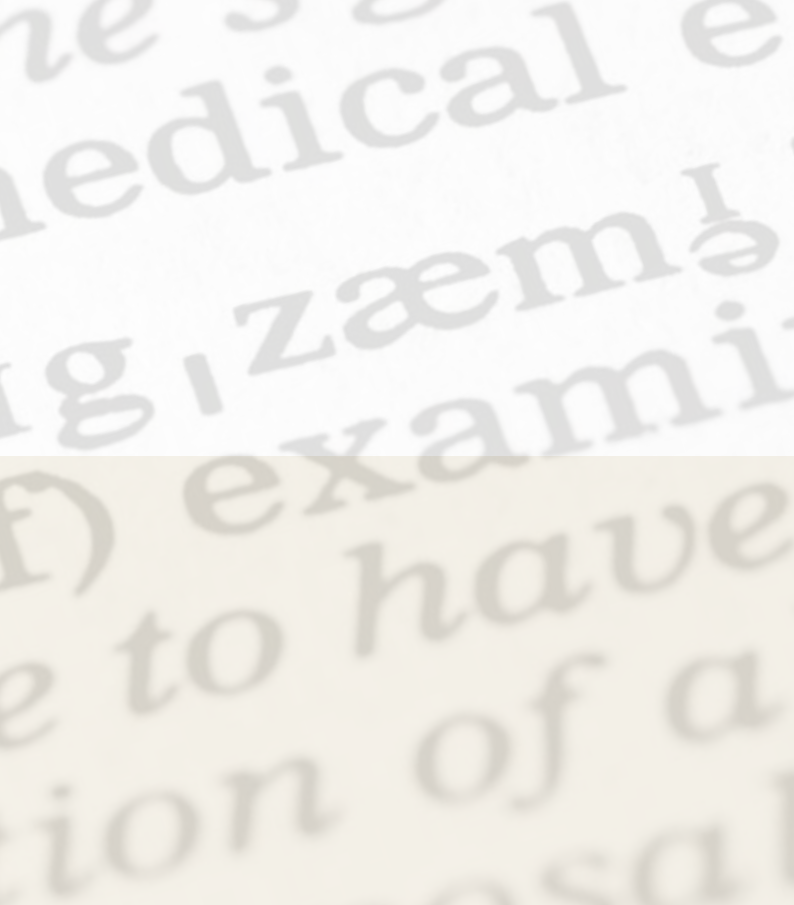
d) What do we know about Scandinavian research into institutions for children from 0 to 6 years old? We know a lot - for example, the amount of research published annually, the problems this research focuses on, and the relationship between quantitative and qualitative research approaches. This is knowledge we have not previously had. [3]

I will now leave the question of what a clearinghouse for educational research actually does and look at an important analysis of how a clearinghouse may be viewed as a new and subtle tool that cleverly overcomes the resistance of the practitioners and researchers towards control by a centralised authority.

**Critical arguments against the new paradigm
– another analysis**
In a recently published monograph, a young Danish researcher, Justine G. Pors, presents an analysis of the way in which a modern administrative institution such as the Danish Ministry of Education carries out its managerial functions with respect to the Danish school system (Pors, 2009).

The starting point for Pors' analysis is two OECD reports (Ekholm, 2003; Mortimer, 2004) on Danish basic education and an evaluation of Danish educational research, respectively. A similar report was issued for Norway in 2005. The problem that the OECD reports sought to explain was why Danish pupils scored relatively poorly in international comparisons such as the PISA study of 2000. The explanation was that Danish schools lack an *evaluation* culture and that Danish educa-

tional research has the wrong focus. The OECD reports were followed by a wave of critical attention directed at the Danish basic school and Danish educational research, and in its wake came legislative changes, evaluation conferences and a string of additional initiatives intended to assist - in particular - Danish school teachers in contributing to an improved evaluation culture and to motivate researchers to change the focus of their research. One of these initiatives was the creation of the Danish Clearinghouse for Educational Research.

These initiatives were met with resistance from representatives for the Danish school system in general, from teachers' representatives and from the research community in particular. In debates in Danish professional forums and within the Danish educational research community, the national tests, the notorious pupil plans and the Danish Clearinghouse were attacked as symbols of control, bureaucracy and a narrow professional goal orientation.

Representatives for the school system and the educational research community described how political control and excessive management were corroding the inner values of the school system and the fine tradition of freedom within educational research. This creates a dichotomy between internal school evaluation and free research on the one hand, and external evaluation and directed research on the other. Politicians and their civil servants represent the external power that is breaking down the values that could only be protected and developed fittingly within the confines of the school and the educational research community.

---

[3] The reports mentioned are available at http://www.dpu.dk/clearinghouse.

Pors advances the thesis that:
> The strategies of the Ministry for managing the school system are more varied and more focused towards the culture and identity of the school teachers than school representatives wish us to believe with their diagnoses of political control and compulsion (Pors, 2009: 7).

And similarly this thesis can be expanded to claim that:
> The strategies of politicians for controlling educational research are more varied and more focused towards the culture and identity of the research community than the critics wish us to believe with their diagnoses of political control and compulsion.

Thus Pors' claim is that setting up this dichotomy between internal values and external control or compulsion neglects the crucial mechanisms through which management or control takes place.

As far back as 1978, the American organisation theorist James March noted that:
> [A]ctivities in schools are not easily or precisely controlled through hierarchical directives or managerial incentives (March, 1978: 238).

The solution to this, claimed March, was to be found elsewhere:
> Rather, educational management is controlled by diffusion of ideas and the development of social and professional norms. Good administration is encouraged by good ideas (ibid.).

In fact, both the ministry and politicians find themselves in a managerial dilemma: the more one consciously attempts to influence the teaching community, the more resistance one encounters. The more one consciously attempts to direct educational researchers, the more one is criticised for promoting what is often referred to as "invoiced research".

The sword that cuts this Gordian knot is referred to as "self-management" or "self-management competence". The freedom of method and freedom of research enjoyed by teachers and researchers should not be perceived as a problem, but as an opportunity. If the true power-holders – the ministry and politicians – can persuade teachers and researchers to manage themselves based on precisely the same desired goals, then their freedom of method and research will assist in exercising that power. The key, then, is to create an evaluation culture and – said in all modesty – a research clearinghouse.

The aim is to influence the attitudes and behaviours of teachers and researchers. The method is to offer them ways of thinking that they themselves are willing to adopt. The task is thus to develop a management strategy that leads teachers and researchers to act on the basis of their own free will

and in their own interest to strengthen the evaluation culture and a research culture that develops useful knowledge for practitioners and guidance for policymakers. The ministry and politicians achieve this by offering a range of evaluation tools, practical recommendations, definitions of best practice, etc. - in other words by designing a campaign programme focused on selected aspects of problems, goals, and methods with a view to making these the basis for future dialogue, negotiation, etc. within the school system and the research community. The purpose of the campaign technique is to transform political struggles into institutional change.

In the final analysis what we are influencing is the identity of the teachers and researchers. A campaign invites teachers and researchers to transform themselves into goal-oriented, reflexive and opposition-responsive individuals who in a continual, ongoing process take responsibility for managing an ever-larger proportion of their tasks, themselves and their interactions with other teachers and researchers - in short, to become self-regulating, reflexive and autonomous. By offering tools with which to reflect on all aspects of their professional work, we encourage the teacher and the researcher not only to evaluate their daily practice in the classroom and the research community, but also to reflect about their ability to reflect about their reflection, and in the third-order case to reflect about whether their reflection about their reflection is leading to autonomous self-regulation. In this light, the difference between internal and external control disappears. The control has moved inside the heads of the teachers and researchers.

One particular perspective concerning this control strategy can be obtained when one also considers what it renders invisible. To become a "good teacher and researcher" in this framework implies that one assigns a special significance to certain concepts. The concept of reflexion is particularly vital. To be a reflexive teacher or researcher acquires the meaning of being a competent, well-documented and theoretically well-grounded teacher (or researcher). But at the same time this concept has another interpretation which is not captured here. To use one's own personal experience or one's intuition - according to the ministry's evaluation campaign or to the Danish Clearinghouse - is absolutely not the same as being reflexive and evidence-based. In this way new concepts and interpretations arise, defining how meaningful communication can take place. Concepts acquire the status of being self-evident.

I will now leave this constructivist discourse which has analysed the governmentality of educational practice and research and in this final short section take a look at possible answers and future challenges in educational practice and research.

Answers to the critics and future challenges of the Danish Clearinghouse

Pors' analysis is important and illuminating. It refrains from normative statements and keeps its focus on the primary problem: to reveal the mechanisms through which politicians and ministries try to produce the teacher and educational researcher best suited to fulfil the policies of those same politicians and policymakers.

But what in fact is the substance of the criticism? That the practitioners base too much of their practice on personal experience and intuition and that the researchers do not produce research of relevance for practice and policymaking.
If this is the charge, then I find it easy to live with – with a few supplementary comments:

Practitioners should not *merely* base their practice on personal experience and intuition. They ought also to make use of evidence-based knowledge from educational research in their reflections about their practice. Nobody wishes - in a European situation – to see teachers deprived of their professionalism by basing practice solely on top-down educational programme packages. But where is the difficulty in being well-informed about research results within the field in which one is practicing?

Educational researchers should not abandon the critical research potential outside the research institutions but rather incorporate it into their research processes. And if research is to have any effect on practice and policy, it must be designed in ways that permit professional criticism and professional agreement. This demands a certain professional discipline of the research within this field so that mutual – constructive - criticism of research becomes possible and the accumulation of research results can be promoted.

During its limited lifetime, the Danish Clearinghouse for Educational Research has discovered that existing educational research environments can be very difficult to cope with.

If we look especially at Scandinavian research, we see from our overview of Scandinavian research on institutions for children from 0 to 6 years old that for the very first time we have produced a picture of an entire research area which shows what sort of research is being produced – problem formulations, research design and research approach, and how much is being produced. This field is producing some 50-60 research reports of high quality every year. Of this, qualitative research accounts for slightly more than half and the descriptive or quantitative research for about 40 per cent, while almost 10 per cent are intervention studies and/or programme studies. If we set aside those problems related to synthesising qualitative research, which we are also trying to solve, then the principal obstacle is the basic "impressionism" that still flourishes in quantitative educational research, which can make it very

difficult to assess whether the variables that are studied in one research project are actually similar to those studied in another. There is a conceptual chaos in educational research, which might be considered constructive by some, but is actually rather destructive. Are we in fact examining the same educational reality in the various primary studies?

One depressing consequence is that this situation can in fact be transformed into a criticism of the educational research that is being carried out: perhaps teachers are forced to fall back on their personal experience and intuition, for where can they find the evidence on which to develop an evidence-based practice?

References

Achinstein, P. (ed.) (2005). *Scientific Evidence - Philosophical Theories & Applications*. Baltimore & London: The Johns Hopkins University Press.

Chalmers, I. (2006). The scandalous failure of scientists to cumulate scientifically. Abstract to paper presented at: Ninth World Congress on Health Information and Libraries; 2005 Sep. 20-23; Salvador, Brazil. Retrieved from http://www.icm19.org/program/activity.php?lang=pt&id21. Accessed on 3 February 2009.

Ekholm, M. (2003). *OECD Nation Educational Policy Reviews - Pilot Review of the quality and equity of schooling outcomes in Danish basic education. Background report:* EDU/EC.

Gilbert, R., Salanti, G., Harden, M., & See, S. (2005). Infant sleeping position and the sudden infant death syndrome: systemtic review of observational studies and historical review of recommendations from 1940 to 2002. *International Journal of Epidemiology, 34*, 874-887.

March, J. G. (1978). American Public School Administration: A short analysis. The School Review, 86(2), 217-250.

Mortimer, P. (2004). *Special Session of the Education Committee: Pilot review of the Quality and Equity of Schooling Outcomes in Denmark. Examiner's report:* EDU/EC.

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health, 57*, 527-529.

Petticrew, M., & Roberts, H. (2006). *Systematic Reviews in the Social Sciences. A Practical Guide.* Malden: Blackwell Publishing.

Pors, J. (2009). *Evaluering indefra - Politisk ledelse af folkeskolens evalueringskultur.* Frederiksberg C: Nyt fra samfundsvidenskaberne.

Spock, B. (1946, 1998). *Dr Spock's Baby and Child Care.* New York: First Pocket Books.

Olaf Gjerløw Aasland, University of Oslo

Evidence-based Medicine: Experience and Advice from Medical Research

The concepts of evidence-based research and evidence-based practice are said to have originated in medical research. Yet today they are applied in research generally, including in educational research. In this commentary on Sven Erik Nordenbo's paper, I offer some advice to the field of education based on my personal experience with evidence-based medicine.

A question that comes to mind while reading Nordenbo's excellent manuscript is: Why did it all start with medicine? The story usually begins with Archie Cochrane's experience as a doctor and prisoner of war during WWII:

> I had considerable freedom of clinical choice of therapy: my trouble was that I did not know which to use and when. I would gladly have sacrificed my freedom for a little knowledge. I had never heard then of "randomised controlled trials", but I knew there was no real evidence that anything we had to offer had any effect on tuberculosis, and I was afraid that I shortened the lives of some of my friends by unnecessary intervention (Cochrane 1971).

While theory and practice tend to live separate lives in most other professions, both in training and in everyday activities, medicine is always a combination of the two. The "evidence filter" – convincing documentation – is a powerful political tool that can be used or abused, and the examples of this are abundant. Medicine makes a big difference, both in private and in public life.

What might be the reason for this new variant of medical imperialism? Why does one body of knowledge after another dress in these new clothes? One explanation could simply be

that it tastes good and looks nice, and has become a fashion. Some people might call it new wine in an old bottle. Another possibility could be that when the medical knowledge monopoly, regarded by our forefathers as one of the strongest and most precious components of professional autonomy, was dissolved, which actually started before the Internet era with a number of pamphlets and do-it-yourself books, it paved the way for a deluge of de-professionalisation and perhaps the emergence of a new professionalism. In this new professionalism, the professionals no longer individually possess specific knowledge, but they know how to extract and synthesise it from a common and almost infinite knowledge base. The modern heroes are no longer the professors, but young "knowledge athletes" - those who master these extraction and processing skills.

These changes could probably not have taken place if we were still in the era of paper books and journals. The existence and availability of extremely large electronic databases are necessary in order to make modern, systematic reviews. So it could really all be technology driven.

Finally, the reference to medicine as the point of origin for the evidence movement may simply be artificial and an act of convenience, the same way we define a gunshot in Sarajevo as the beginning of WWI and ignore the rest of history and geography.

Nordenbo concludes his paper like this:

> If we set aside those problems related to synthesising qualitative research, which we are also trying to solve, then

the principal obstacle is the basic "impressionism" that still flourishes in quantitative educational research, which can make it very difficult to assess whether the variables that are studied in one research project are actually similar to those studied in another. There is a conceptual chaos in educational research, which might be considered constructive by some, but is actually rather destructive. Are we in fact examining the same educational reality in various primary studies?

One depressing consequence is that this situation can in fact be transformed into criticism of the educational research that is being carried out: perhaps teachers are forced to fall back on their personal experience and intuition, for where can they find the evidence on which to develop an evidence-based practice?

I read this as a message from a slightly frustrated professor, almost on the brink of resignation. He wants a modern, streamlined research assessment system, while the current situation is rather chaotic and non-linear.

I regard Professor Nordenbo as a pioneer in the field of systematic educational research, and I therefore hope that the following advice and examples will inspire him and his colleagues to not abandon their ambitious goals of building a modern clearinghouse for educational research.

A pyramid of evidence
In order to deal with the large variety in quality, the medical evidence-based community uses hierarchical models. One of
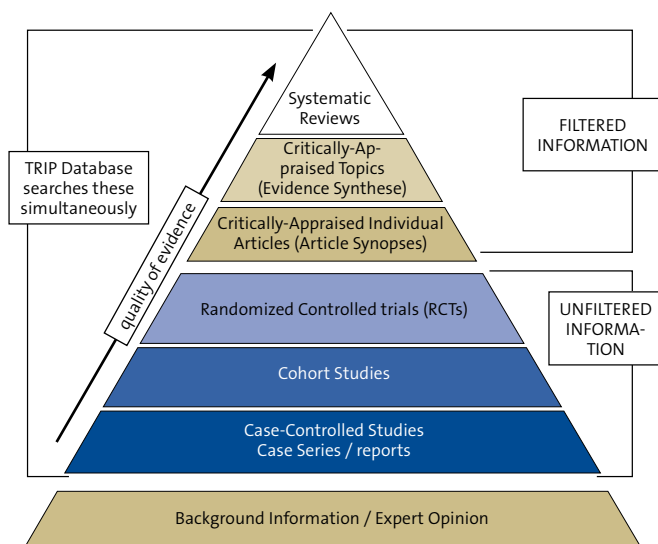
several variants is proposed by Sackett et al. (Sackett, Straus, Richardson et al. 2000).

It is clearly necessary to have such a "cleaning tool"; in medicine this has echoed the role of a surgeon's knife, cutting away a rather large body of unreliable research - unreliable because the methodology was not good enough. Such operations will probably be necessary in all areas where evidence-based systems are implemented, including in educational research.

Five pieces of advice based on personal experience
In the following I will offer five pieces of advice, which are illustrated with some examples.

1. Don't bother with the obvious.
Mostly for fun, Smith and Pell published a paper in *British Medical Journal* entitled "Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials" (Smith & Pell, 2003: 327, 1459-61). Their conclusion was that no randomised controlled trial of parachute use has been undertaken. The basis for parachute use is purely observational, and its apparent efficacy could potentially be explained by a "healthy cohort" effect. *Individuals who insist that all interventions need to be validated by a randomised controlled trial need to come down to earth with a bump.*

2. Don't forget that our study objects usually are merely ordinary human beings.
One of the early randomised controlled trials which appeared in 1989 in the prestigious journal The Lancet dealt with the registration and counting of fetal movements by pregnant women. The conclusion, based on data from 68,000 women, was that even though the mothers counted fetal movements, most of the dead fetuses were dead by the time the mothers received medical attention. "The study does not rule out a beneficial effect, but at best, the policy would have to be used by about 1,250 women to prevent 1 unexplained antepartum late fetal death, and an adverse effect is just as likely. In addition, formal routine counting would use considerable extra resources" (Grant & al., 1989:2(8659): 345-9).

This sensational publication caused a lot of discussion, but also put an end to further research on this particular topic. Due to the seemingly irrefutable methodology, the entire scientific community changed its focus.

But there was a critical flaw here – the pregnant women who were not advised to count the fetal movements, and who more or less attended the same clinics as those in the intervention group, talked with each other and of course followed the good advice regarding counting. In other words, both groups became intervention groups, a phenomenon termed "contamination". So the original evidence-based finding that counting did not matter was discarded, and today we are back to counting, as has been the common practice throughout the centuries.

3. New crazy ideas? Check them out!
Peptic ulcer, a disease thought to be mainly stress-related, used to be treated with extensive surgical removal of large parts of the stomach, an operation pioneered by the Austrian surgeon Theodor Billroth (1829-1894).

An alternative theory of infectious aetiology had been circulating for some time, and potential agents were identified. One of the central researchers, Barry Marshall of Australia, even drank a bacterial culture with helicobacter pylori to prove that it could survive in the highly acidic environment of the ventricle (stomach). This was the main reason why the sceptics were not willing to accept the hypothesis of infection.

Another *Lancet* publication, one year prior to the fetal movement publication, presented definite proof that the helicobacter theory was correct (Marshall et al. 1988: 1437-42). Within 15 years Marshall and his colleague Robin Warren had won the Nobel Prize in Physiology or Medicine.

4. Don't ignore individual differences and rare events.
In 2002 Ames and his colleagues published a paper on how high doses of vitamins could have dramatic healing effects on patients with rare genetic diseases. (Ames, Elson-Schwab & Silver, 2002: 616-658). This extensive article consists of 43 pages with 377 references. Having reviewed all the material in detail, the authors suggest that all anecdotal reports on megadose vitamin effects be published on a special website.

This is a new variant, almost antithetical to the randomised controlled trials of the EBM movement. It is intuitively appealing and reminds us of the importance of anecdotal information about unusual occurrences that will never be reported as "evidence". This approach could be used in a number of other situations as well, such as for collecting data on other unorthodox therapies.

5. Don't forget that research is mainly about constructing and fitting models. Real life is not even close to being monocausal or linear.
Here I would like to use an example from my own non-medical pastime.

The bandoneon – the Argentinean tango accordion made in Germany – is probably the only non-linear musical instrument still in use. It has 71 buttons which play mostly different tones on the in and out movements (although some play the same tone) and which are seemingly placed completely at random. The bandoneon is really four instruments in one, spanning five octaves with one and a half octaves overlapping: in and out with the left hand and in and out with the right hand.

The first thing a good player must do is to discard the thought of linearity and system; he must rely completely on the brain's ability to build new paths and make new reflex systems. I call this a "deep biological" logic.

The following illustrations show how the instrument's non-linearity compares with the linearity of an ordinary accordion when the chromatic trajectory is drawn.

ow·ing···
ort her ~ly. (b) i
ot somebody.
knowl·edge l'
ing: A baby has
liarity gained i
ion: My ~ o
ha

Button accordion

Piano accordion

Opening
Left hand
Closing

142 tone bandoneon non-line-
ar, four instruments in one

Opening
Right hand Bandoneon
Closing

Bandoneon

My final point here is simply that a non-linear, seemingly "chaotic" instrument such as the bandoneon can produce beautiful music in the hands of a skilled performer.

References

Ames, B., Elson-Schwab, I. & Silver, E. A. (2002). High-dose vitamin therapy stimulates variant enzymes with decreased coenzyme binding affinity (increased Km): relevance to genetic disease and polymorphism. *American Journal of Clinical Nutrition*, vol. 75, no 4, pp. 616-658.

Cochrane, A. (1971). Effectiveness and efficiency—random reflections on health services. London: Royal Society of Medicine Press Limited.

Grant, A., Elbourne, D., Valentin, L. & Alexander, S. (1989). Routine formal fetal movement counting and risk of antepartum late death in normally formed singletons. *The Lancet*, vol. 334, issue 8659, pp. 345-9.

Marshall, B. J., Goodwin, C. S., Warren, J. R., Murray, R., Blincow, E. D., Blackbourn, S. J., Phillips, M., Waters, T. E. & Sanderson, C. R. (1988). Prospective double-blind trial of duodenal ulcer relapse after eradication of Campylobacter pylori. *The Lancet*, vol. 332, issues 8626-8627, pp. 1437-42.

Sackett, D. L., Straus, S. E., Richardson, W. S. et al. (2000). *Evidence-based medicine: how to practice and teach EBM*. 2nd ed. Edinburgh: Churchill Livingstone.

Smith, G. & Pell, J. (2003). Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. British *Medical Journal*, vol. 327, pp. 1459-61.

Dominique Foray, Ecole Polytechnique Fédérale de Lausanne

Educational Innovation: An Economist's Perspective

This paper discusses innovation in education. Educational innovation is the act of creating and then diffusing new educational tools as well as new instructional practices, organisations and technologies.

An initial look at innovation in education

Innovation is not research. It is (often) based on research and the advancement of knowledge, and consists of changing processes and practices in order to improve the quality and productivity of the service that is delivered. Creating an educational sector in which valuable innovations are constantly generated and efficiently used and managed is a major challenge to "re-inventing" public education and finding solutions to what has been termed "Baumol's disease".[1]

Some changes are proposed from outside the schools and then are disseminated to them by "reformers". The source of these changes is not innovation but reform! A reform ("outside-in") logic creates little chance for the successful adoption, implementation and institutionalisation of new practices. Moreover, policymakers will be frustrated by the failure of many reforms to endure and to displace poor practices. Innovation involves a decentralised way of using new knowledge and information (both from research and current practice) in order to identify problems and generate solutions. Also, people are motivated to disseminate knowledge and solutions that they themselves have created, and there are natural but under-used channels for easy dissemination (Foray and Hargreaves, 2003).

Last but not least, it is useful to stress that one of the major challenges associated with the study of educational innovation is the lack of data. Studies of technological innovations traditionally focus on R&D spending and patenting. These measures are unlikely to be satisfactory in this context (although I will use patent data to a certain extent below in this paper).

A difficult science and a poor link to practice

The educational sector is often characterised by experts as a sector suffering from an innovation deficit and a structural inability to advance instructional technologies and practical knowledge and know-how about pedagogy at the same rate as some other sectors.[2]

> Consider the efforts to develop more effective educational practices in schools: even if we do know more about educational practices that we did previously, knowledge creation in this domain has been slow and there have been severe difficulties in diffusing "new and superior" knowledge. (Nelson, 2003)

The main problem points to the difficulty of developing a science that can illuminate practices and provide guidance for their systematic improvement (Foray, 2001, 2006). Formal R&D is of secondary importance, both for the training of professionals and for the generation of useful innovation. What Nelson and Murnane wrote more than 20 years ago about education is still by and large true: educational R&D is very weak

---

[1] Many years ago, Baumol introduced an interesting distinction between progressive and non-progressive sectors. Non-progressive sectors are the sectors in which productivity growth is limited, very sporadic and far smaller in magnitude than what is taking place in the progressive part of the economy (Baumol & Bowen, 1965, Baumol, 1967). Such a productivity gap between two sectors gives rise to Baumol's (or cost's) disease. Education has always been considered by experts to be a paradigmatic example of a non-productive sector (Roza, 2008, Hill & Roza, 2010).

[2] In this case, technical knowledge involves the broad set of both embodied and disembodied knowledge that enables the development of pedagogical practices and instructional technologies.

in producing practical solutions. "R&D should not be viewed as creating 'programmes that work'; it only provides tidy new technologies to schools and teachers. It is thus a mistake to think of educational R&D in the same way as industrial R&D" (Murnane and Nelson, 1984). Too rarely does educational R&D generate knowledge of immediate value for solving problems and developing applications. Of course, social science will continue to contribute its theory to the field of education. However, the goal of this kind of research is not to provide and develop a repertoire of reliable practices and tools for solving immediate problems that teachers meet daily in their professional life: "For novice teachers, practical problems in classrooms are not usually perceived to be solvable by drawing upon the psychology of education or child development that have been studied in universities" (Foray and Hargreaves, 2003).

The problem of a very weak linkage between science and the improvement of practice is crucial since it influences both the supply of and the demand for research. This in turn creates a fundamental inertia in the system caused by the negative externalities that exist between weak supply and insufficient demand.

There are three factors which explain why the role of science in illuminating practices in education is limited:

- On the supply side, the educational sciences are a difficult field of inquiry. Berliner (2007) wrote that educational research is the most difficult science of all. "We do our science under conditions that physical scientists would find intol-

erable," he stated. Compared with designing a bridge, the science of helping schools and classrooms to change is harder because the context cannot be controlled and the inherent lack of generalisability across contexts reduces the value of any research method in illuminating a body of practices.[3] Educational science is a genuine field of inquiry, but it is nothing like an applied science or engineering discipline with regard to developing a body of knowledge and techniques that can shed light on educational practices.[4]

- On the demand side, most practitioners who are (or should be) involved in the improvement of practice do not believe that the educational problems they face in their professional life can be solved by inquiry, evidence and science (Elmore, 2002). For example, they do not believe it is necessary to have a developmental theory of how students learn the content of a subject area and how the pedagogy used relates to the development of knowledge and content. Weak incentives for teachers to use research are rooted in a deep-seated culturally-based belief that the ability to teach is an individual trait and that the foundation of a teacher's performance involves natural aptitude, inspiration and talent rather than a set of competencies acquired over the course of a career (Elmore, 2002). Because of this belief, it is very difficult to make a case for knowledge management, specifically for building databases that contain evidence about what works and encouraging teachers to act like engineers by searching for solutions to problems in casebooks. "Teachers are primarily artisans, working alone in a personally designed environment where they develop most of their skills by trial-and-error tinkering ... In short, they

---

[3] See the special issue of EINT (Foray, Murnane & Nelson, 2007) about the comparison between educational research and research in the biomedical area. Cooke and Foray (2007) describe the US policy experience with developing education science through the development of experimental research capacity.

[4] See the chapter by Richard Shavelson in this volume which addresses the issue of rigorous and relevant research in education in a very insightful way.
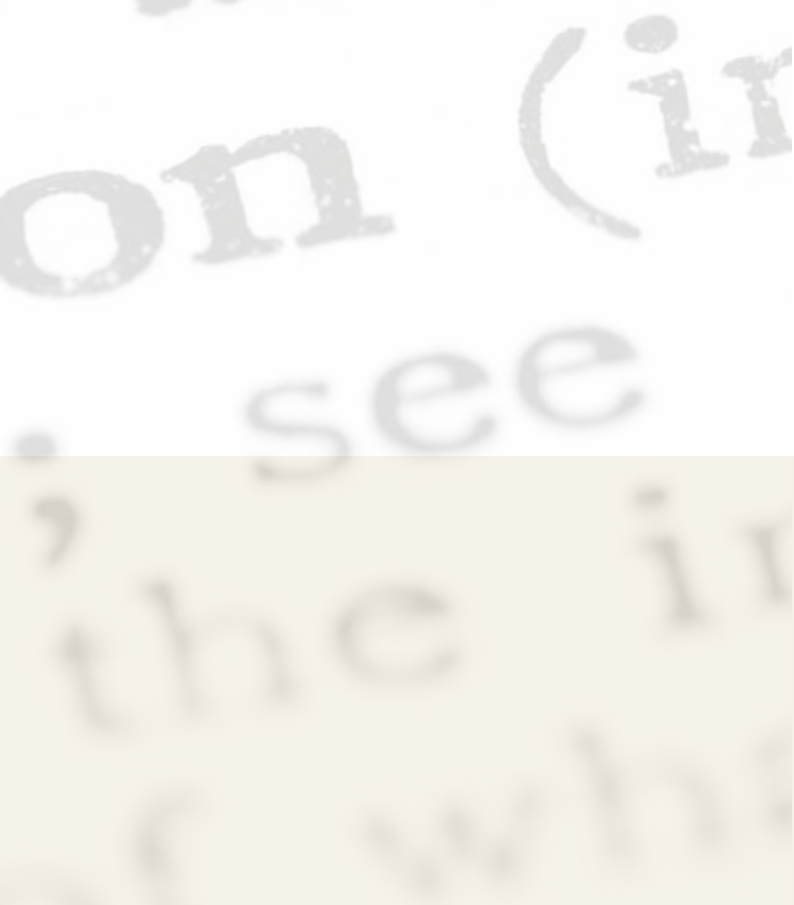
learn to tinker, searching pragmatically for acceptable solutions to problems their 'clients' present" (Foray and Hargreaves, 2003).

- Finally, incentives to codify technical knowledge and know-how are generally lacking, and the resources allocated to codification are weak. Numerous practices remain tacit. These are not explicated or articulated, and they are invisible and difficult to transfer. "There is no more in education than a weak equivalent in the field of pedagogical knowledge to the systematic recording and widespread use of cases found in surgery or law and the physical models in engineering and architectural practice. Such records coupled with comments and critiques of experts allow the new generation to pick up where earlier ones left off" (Foray and Hargreaves, 2003). Some important mechanisms to support the cumulative nature of knowledge and its progressivity and to realise the potential for spillovers are simply missing. "The beginner in teaching must start afresh, uninformed about prior solutions and alternative approaches to recurring practical problems. What student teachers learn about teaching is intuitive and imitative rather than explicit and analytical" (ibid.). When excessive stocks of knowledge are left in tacit form, this makes them more costly to locate, appraise and transfer. This may lead to excessive insularity and wasted resources, resulting in the underuse of the existing stock of knowledge. This may therefore create private and social inefficiencies.

Translating increasing pressure to perform into innovation
To paraphrase Nelson, the key to successfully advancing technical knowledge has been to design practice around what is known scientifically. For various reasons, this key is not operating well in education.

As a result, policymakers, industries and society as a whole are asking schools to make improvements in the presence of an extremely weak technical core. "Consider what would happen if you were on an airplane and the pilot came on the intercom as you were starting your descent and said, 'I've always wanted to try this without the flaps'. Or if your surgeon said to you in your pre-surgical conference, 'You know, I'd really like to do this the way I originally learned how to do it in 1978'. Would you be a willing participant in this? People get sued for doing that in the 'real' professions, where the absence of a strong technical core of knowledge and discourse about what effective practice is carries a high price" (Elmore, 2002).

The problem is not so much the lack of incentives for schools and managers to improve educational practices and technologies. These incentives are there, probably to a lesser degree than in other sectors, but pressure on the schools to perform, which is generated through higher standards and greater accountability, is increasing and thus creates such incentives. However, the problem lies in the way practitioners, teachers and administrators try to respond to these incentives and pressures. The problem lies in the failure to translate such pressures into innovation, improved practices and the development of instructional know-how and technologies. Practitioners do not try to improve practice by relying on a strong technical core of knowledge which should be available in casebooks and databases. Instead, they respond to increased accountability by changing the structure, even though changing the structure does not change practice. As Elmore (2002) forcefully argues, people and schools put an enormous amount of energy into changing structures and usually leave instructional practice (innovation) untouched.
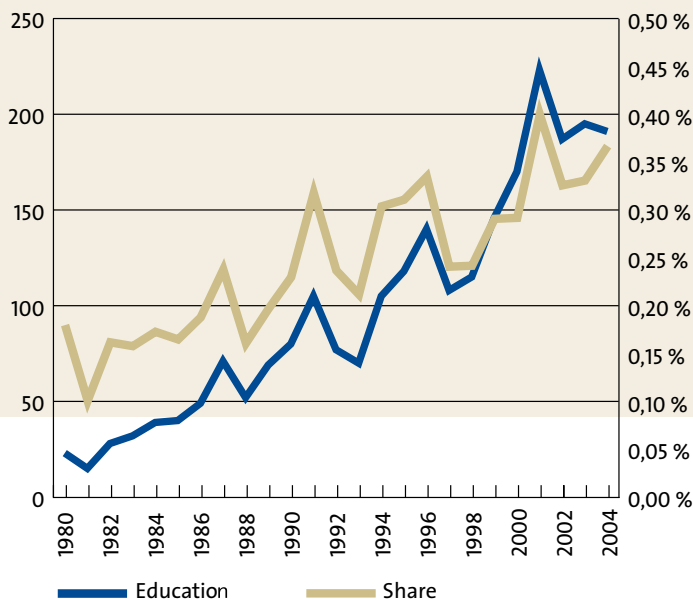
Figure 1. Evolution of education-related technologies
Source: PATSTAT (September 2008). Figures express triadic patent families declaring a G09B IPC subclass.

A small (innovation) explosion?

A brief look here at patent data provides us with a slightly different view of innovation in this sector. Looking at the IPC subclass G09B in PATSTAT, it becomes clear that patent applications in the area of educational and teaching technologies have increased dramatically since the early 1990s (figure 1).[5]

Also, a positive trend is found for these technologies as a share of the total production of technologies, which demonstrates that this traditional sector is growing faster in technological terms than the average.

However, this growth cannot be explained by the strategic behaviour of large firms alone. We can also observe the formation of a population of small firms that specialise in the development of technological solutions to educational problems and issues. This is apparent from the entrance of new firms onto the market (cf. figure 2a) as well as the declin-

ing (technological) concentration as evidenced by various indicators. In figure 2b it can be observed that the concentration – expressed by the technological shares held by both the top four and the top 10 firms – has been declining steeply over the past two decades. The inverse Herfindahl-Hirschman Index (HHI) provides a similar picture, showing that the technological concentration has been reduced from about 30 to 60 "ideal" firms. Furthermore, all three indicators suggest that this evidenced de-concentration might be slowing down or, if we consider the HHI, even regressing. In any case, these preliminary results suggest the emergence and consolidation of an industry specialising in the production of educational and instructional tools and knowledge with strong roots in new information technologies. A large segment of this industry is comprised of small, specialised firms.

---

[5] In our study (Foray & Raffo, 2009) we consider educational and teaching-related technologies to be any patent filed under the G09B IPC subclass. This subclass is defined as *educational or demonstration appliances; appliances for teaching, or communicating with, the blind, deaf or mute; models; planetaria; globes; maps; diagrams.* This subclass covers simulators regarded as teaching or training devices, which is the case if they give perceptible sensations similar to the sensations a student would experience in real life in response to his or her actions; models of buildings, installations, and the like. It does not include simulators that merely demonstrate or illustrate the function of an apparatus or system by means of computing and as such cannot be regarded as teaching or training devices; or components of simulators, if identical to real devices or machines.

(a) Firms' entry.

(b) Technological concentration

Education Entry (all firms)
Education Entry (new firms)   Share

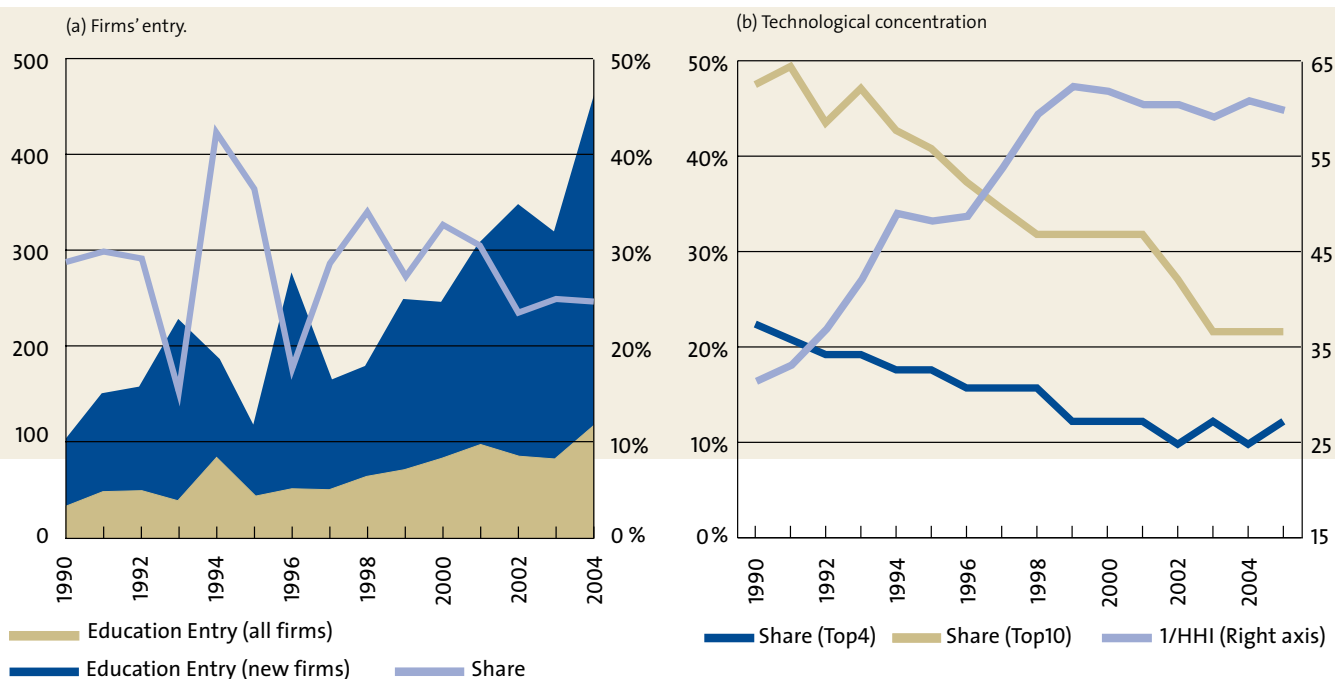Share (Top4)   Share (Top10)   1/HHI (Right axis)

Figure 2. Firms producing education-related technologies (entry and technological concentration)
Source: PATSTAT (September 2008). (a) Firm figures have been retrieved from their triadic patent families.
(b) Figures are compiled from the firms' triadic patent families portfolios.

The development of instructional technologies in the wake of a great general purpose technology

The new information and communication technologies (ICTs) are clearly a source of innovation within the educational system. ICTs potentially offer a wide range of new tools and instruments that can profoundly change the technological, organisational and institutional foundations of the sector in question. In this case, the development of ICTs provides opportunities to enlarge the repertoire of instructional technologies. The process referred to as the co-invention of applications is not a minor matter since it is the process by which the technology is diffused across a wide range of sectors and specific applications are generated.

In fact, the characteristics of a general purpose technology (GPT) are horizontal propagation throughout the economy and complementarity between invention and application development. Expressed in the economist's jargon, the invention of a GPT extends the frontier of invention possibilities for the whole economy, while application development changes the production function of one particular sector. The basic inventions generate new opportunities for developing applications in particular sectors. Reciprocally, application co-invention increases the size of the general technology market and im-

proves the economic return on invention activities related to it. Therefore, dynamic feedback loops are created depending on which inventions give rise to the co-invention of applications, which in their turn increase the return on subsequent inventions. When the situation evolves favourably, a long term dynamic develops, consisting of large-scale investments in R&D whose social and private marginal rates of return attain high levels.

It seems that this sort of renaissance of innovation in practices and methods of pedagogy and instruction is strongly associated with the dynamics of ICTs. The application of ICT in education is not a single innovation, but an array of technologies that can be applied in a variety of ways. ICTs are also viewed as an enabler of change: schools engage in a series of activities which would not have been possible without ICT. It is premature, however, to claim that the education sector has reached a position today as a central user of these technologies with the potential to significantly boost the dynamics of ICTs.

Discussion: An emerging educational tool industry

Intensive innovation activity related to the development of new instructional tools and technologies can be observed.

However, the locus of this activity is not actually inside the sector but rather on the supply side. We can observe the formation of a tool industry: a population of specialised firms that invent, design and commercialise educational tools. Such a process, as in any historical case of a tool industry emergence, involves the delocalisation of knowledge – at least in part from the delivery of the educational service. There is a kind of shift in knowledge "holding" which involves the emergence of a new site of knowledge accumulation: the tool producer. Historically, one important reason for the emergence of a tool industry (beyond the classical reason of market size increase) is the rise of the systemic approach to the problem of increasing the productivity of industrial or service outputs. The process of relocating the specialised knowledge about tools outside the institution that delivers the final service (the school in our case) allows for the production of generic, multi-purpose machines and tools which replace the specialised tools developed previously within each specific organisation that delivers the service.

Historically, the formation, emergence and development of tool industries have often generated efficiency gains and economic growth through greater specialisation, intra-segment competition between the tool producers and an effective co-ordination between the tool companies and the downstream organisations.

Given our observation and discussion of the innovation deficit in the "core" of the system (the classroom), it is good news that a population of entrepreneurs enters and grows in the market for new educational tools. Companies competing to invent and commercialise tools are expected to play a sig-

nificant role in enhancing innovation and productivity in the downstream sector.

However, there is a need to qualify this trend. One important concern is related to the ability of the public sector to exploit the opportunities offered by the emerging tool industry. Another concern is related to the increasing activity of patenting which is needed for small, specialised firms to enter and compete in the market, but which is likely to adversely affect static efficiency through the pricing of ideas and knowledge which used to be freely accessible.

Patent problems with the new structure
The development of a market for instructional tools implies that potential users must pay now for access to methods and knowledge of the kind that used to be obtained at no charge but is now explicitly priced in the form of licensing agreements. In educational communities, some of the new patents are likely to generate great anxiety as practitioners realise they are infringing on patents and violating the law just by applying methods and practices that they have applied freely since the beginning of their professional careers. We know that researchers in the biomedical sciences are quite good at simply "ignoring" (in the sense of failing to comply with) the patents on research tools, and the firms that have been granted these patents either anticipate weak appropriability of their knowledge by granting licenses on a large scale or they simply tolerate infractions, especially by academic researchers. This set of norms and practices on both sides quite effectively minimises the social inefficiencies which are potentially generated by the "anti-commons problem" in biomedical research. It is not clear whether school managers and

teachers are in a position to behave in a similar manner and what the strategic responses of the small, specialised firms holding the patents would be.

For example, Blackboard Inc.'s US patents on "technology used for Internet-based education support systems and methods" encompass 44 different features which comprise a learning management system. F. Lowney, director of the IT management system at the Georgia College and State University Library, wrote, "Much of what Blackboard claims to have invented really came from and was freely given by the education community. Now the community is being punished through a gross lessening of competition in this market" (Networkworld, 2008). For one associate professor of medical education, the real question is, "What are they going to do next, try to patent word processing and charge you royalties if you are using it in a classroom? If obvious uses of technology to facilitate teaching based on standard software applications are allowed to be patented just because they are used to support education, we are in real trouble" (Inside Higher Ed, 2006). The problem with Blackboard's patents, and we suspect with hundred of patents for educational technologies, clearly involves the usual conflict between open source communities, which are proliferating in the educational sphere, and for-profit businesses, which are attempting to enforce their claims on some (software) patents. Consequently, a new problem has arisen regarding patents in an area where traditionally the norms of the public good and free access have been paramount.

Another problem with the vertically disintegrated structure relates to the ability of the small, specialised companies to reap the benefits of their innovation. Transaction and bargain-ing costs in these markets for pedagogical methods are likely to be very high, and patents as a means of capturing the value of the innovation might not be so effective (depending partly on how the first problem will be solved). The problems of the firms considered here are rather similar to what has been described by Cockburn (2003) with regard to the tool companies in the biotechnology sector.

There is now a tool industry, but for what market?
Innovation needs entrepreneurship, or at least it needs a complex distribution of the size and age of the firms, including a strong population of entrepreneurs at one end of the continuum. Baumol has written extensively and convincingly on the role and crucial position of the entrepreneur or young innovative firms as a mechanism for fuelling innovation and as an organisational form which is needed to complement large companies' modes of operation. But the educational sector has built severe barriers to entry so that entrepreneurial activities in the sector do not seem particularly attractive. The reward structure in this sector does not favour the competitive entry of new firms and radical innovators who are willing to take risks and be creative in the prospect of huge private returns on R&D and other innovation activities. Among these barriers are (Berger and Stevenson, 2007):
• The education sector does not invest in innovation.
• In many countries, there is a "big edu" – an oligopoly of a few very large suppliers of educational resources which solve the problem of a highly atomised demand by building enormous sales forces; entrepreneurs cannot afford to play this game.
• Slow sales cycles, involving too many people "in charge" at different levels (state agencies, districts, schools).

- The constraints placed on pilot programmes regarding the testing of innovative tools mean that start-ups cannot sell these tools at a scale which is economically viable.
- There is no business culture to manage innovation in the school system. Administrators usually choose to solve problems by making greater use of in-house staff rather than buying new tools and systems, as the staff are already paid for and thus additional costs are not incurred. Few school administrators have formal training in business decision-making or in calculating return on investment.
- Teaching time is a sunk cost; there are no benefits to saving this time.
- Public authorities often recommend that administrators not meet with entrepreneurs and vendors to avoid any unfair advantages, but in the face of such a "vendor wall", how do they become informed about anything?
- Because the various barriers described here constrain the size of potential returns, and educational companies take too long to obtain a meaningful return, no venture capitalists (VC) are interested and most innovative start-ups in this field fail to convince professional VCs to fund them. Angel investors can be a substitute to a certain extent.
- It is common in the field of education for foundations and charities to give away at no charge the very things that entrepreneurs are trying to turn into a business! This unintended consequence of a strategy of building a commons kills entrepreneurial spirit, a phenomenon we also know well from development economics.

In addition to all the problems identified above, the public education sector is also a special market in the sense that the "consumers" do not necessarily want to buy a new and/or better product every year. In order to develop and offer new products and commercialise these, restless innovative firms require a more vibrant market that is more open towards new developments.

As a result, we are facing a disturbing dilemma: We observe some intensive innovation activities, but the market seems very hostile. So what are these entrepreneurs doing? The answer is straightforward: The companies are targeting markets other than the educational system from primary to upper-secondary level, such as corporate education, education during leisure time, and tertiary and vocational education, which are perhaps smaller markets but far more "entrepreneur friendly". In fact, an in-depth analysis of the top 50 specialised companies in patenting educational tools shows that very few are successful in commercialising their inventions in the school systems from primary to upper-secondary. Most companies target the other markets.

Conclusion

The good news for education – a sector which displays notorious difficulties in generating and exploiting innovations to improve practice – is that an educational tool industry has emerged - that is to say, a population of small firms that invent and commercialise instruction (mainly ICT-based) technologies. New sites of knowledge generation and accumulation have emerged: the tool producers.

However, the main commercial target of these companies is not the huge public school system from primary to upper-secondary. This market does not satisfy most conditions for attracting and sustaining strong entrepreneurial activity in the tool business. It appears that the public school system is not (yet?) able to exploit the opportunities offered by the development of a tool industry, and there are still too many obstacles to innovation in the public sector, such as poor management practices, governance and culture, as well as funding and resource allocation logics.

However, other "smaller" markets seem to be attractive enough for entrepreneurs, and this connection explains in part why we have observed the patent explosion and some increase in the number of firms specialising in the tool business.

An important question for further research is whether the invention of tools for corporate education and other "smaller" markets" has *spillover* effects in the sense of building user capabilities (in a very broad sense) in the large primary and upper-secondary education sector so that this sector can progress in learning how to exploit the opportunities offered by the growing educational tool industry.

References

Baumol, W. (1967). Macroeconomics of unbalanced growth: the anatomy of urban crisis. *The American Economic Review, vol. 57,* no. 3.

Baumol, W. & Bowen, W. (1965). On the performing arts: the anatomy of their economic problems. *American Economic Review,* vol. 55, no. 1/2.

Berger, L. & Stevenson, D. (2007). K-12 entrepreneurship: slow entry, distant exit. *American Enterprise Institute Conference on the Future of Educational Entrepreneurship.* Washington D.C., October.

Berliner, D. (2007). Educational research: the hardest science of all. Educational Researcher, vol. 31, 8.

Cockburn I. (2003). O brave new industry that has such patents in it! Reflections on the economics of genome patenting (draft). Boston University.

Cooke T. & Foray D. (2007). Building the capacity to experiment in schools: a case study of the Institute of Educational Sciences in the US Department of Education. *Economics of Innovation and New Technology,* vol. 16(5).

David P.A. & Foray D. (2002). On the fundamentals of the knowledge society. *International Social Science Journal,* no. 171.

Elmore R. (2002). The limits of change. *Harvard Education Letter.* January-February.

Foray D. (2001). Facing the problem of unbalanced development of knowledge across sectors and fields: the case of the knowledge base in primary education. *Research Policy,* 30.

Foray D. (2006). *The economics of knowledge* (Chapter 9). Cambridge: MIT Press.

Foray D. & Hargreaves D. (2003). The production of knowledge in different sectors: a model and some hypotheses. *London Review of Education,* vol. 1(1).

Foray D. & Raffo, J. (2009). A small explosion: patent in educational and instructional technologies and methods; what do they tell us? Fourth EPIP Conference, Bologna.

Foray D., Murnane R. & Nelson R. (2007). Randomized trials of education and medical practices: strengths and limita

tions. *Economics of Innovation and New Technology,* special issue, vol. 16(5).

Hill P. & Roza M (2010). Curing Baumol's disease: in search of productivity gains in K-12 schooling. *CRPE white paper, no. 2010_1.* University of Washington.

Inside Higher Ed (2006). *Blackboard Patents Challenged.* www.insidehighered.com/news/2006/1201/patent

Murnane, R. & Nelson R. (1984). "Production and innovation when techniques are tacit: the case of education. *Journal of Economic Behavior and Organization,* 5.

Nelson, R. (2003). On the uneven evolution of human know how. Research Policy, 32.

Networkworld (2006). *Software patent ignites firestorm in education.* www.networkworld.com/news/2006/11106-software-patent-ignites-firestorm-in-higher-education.html

Roza, M. (2008). Must public education suffer from Baumol's disease? *The Denver Post,* 3 August 2008.

Shavelson, R. (2010). Issues in conducting rigorous and relevant research in education. Chapter in this volume.

Ingrid Carlgren, Stockholm University

The Missing Link?
The Absence of and Need for "Clinical" Research Traditions in Education

According to Foray (this volume), education is a structurally non-progressive sector. A non-progressive sector is one in which productivity does not rise. In contrast to intrinsically non-progressive sectors (Foray uses musical concerts as an example of this), a structurally non-progressive sector is characterised by structural obstacles to progress. Although I believe that several aspects of education must be considered as intrinsically non-progressive,[1] my point of departure is that certain aspects of educational activities could be more progressive. The relationship between the ways teachers design and carry out their teaching activities and the anticipated learning outcomes is often vague and arbitrary. This perceived relationship is based on common sense and strong beliefs rather than on research and reflection. In order to establish an orientation towards innovation and knowledge generation, the common school culture of certainties and strong beliefs must be transformed into a culture of uncertainty. In addition, there is a need for epistemic practices (Knorr Cetina, 2001) in relation to the development of teaching. The amount of research and development work carried out must be multiplied in order to develop the teaching-learning relation into conscious and reflective means-ends relations.

Two different lines of reasoning concerning how to change the situation in education can be discerned. One is based on the fact that teachers do not use existing research results and therefore should be trained to be research consumers. The other argues for re-establishing a tradition of teacher-led research and development in schools. These two lines of reasoning can be linked to two different interpretations of how teaching in schools can become evidence-based – either as recommendations for actions based on syntheses of existing research or as school-based research in which teaching practices are developed in close cooperation with teachers and based on inquiries into teaching-learning relations. The purpose of this paper is to argue for the second line of reasoning, i.e. the need to develop epistemic practices in schools through teacher-driven research and development projects. The main reason for this is that there is not a direct relationship between research results from different fields and actions in the classroom. The use of such results presupposes an "epistemification" (i.e. an establishment of objects of knowledge) of the teaching-learning relation. One of the reasons for the lack of such an epistemification is that teachers are seldom involved in research and development work.

Stigler and Hiebert (1999) identified two different strategies for school development. In *reform strategies*, school development is expected to take place through the implementation of reforms (which most often are rejected or distorted by the teachers). An example of the other strategy, *teacher-owned development work*, is the Japanese lesson study tradition (Hiebert & Stigler, 2002, Lewis, 2000). Whereas the teachers are excluded from research and mainly perceived as research consumers and implementers in the former strategy, the latter entails greater teacher involvement in research and development.

Sweden is a "reform strategy country". Although not always the case, this is what emerged in the wake of the large-scale,

---

[1] Growth and cultivation take time. This presupposes genuine human encounters and depends on the ability of teachers to "capture the moment" and create time and space for meaning-making activities that would be ruined by the aim to make them shorter or more effective.

comprehensive school reforms that were first implemented in 1950 and continue to the present day. In the following section, I will describe how the relation between teachers and research in Sweden has changed over different periods as a way of illustrating how some of the structural aspects that appear as an obstacle to progress today have come about. I will then compare the situation of research in education with that in medicine – a field which is often held up as an example of a progressive sector – in order to draw some conclusions about what is needed for education to develop into a progressive, research-based sector.

Teachers and educational research – the case of Sweden

From an historical perspective, the relation between teachers, educational research and school development has changed in important ways. I have identified three distinct periods in which the relation between teachers and educational research has been perceived in different ways (Carlgren, 1986, 2006, 2009).

Before 1950 – teachers as subjects in research

During the first decades of the last century teachers were actively involved in discussions about educational change and school development. Every fifth year "Nordic school meetings" were held in which teachers met (together with other educational actors) for a week of lectures and seminars in one of the Nordic countries. The primary teachers' unions were active in public debates on the need for a comprehensive school system as well as for the introduction of progressive methods, further education for teachers and research. During the 1940s a state commission outlined a comprehensive school sys-

tem, including a programme for teacher-based research and development activities as a basis for school development. It also included a proposal for 10 state experimental schools (in connection with the university hospitals) (SOU 1948:27). However, these plans were not realised. One state experimental school was established, but it was closed down after 10 years. In 1950 a parliamentary decision on the establishment of a comprehensive school system was made. The decision, however, had a condition attached: the new school system had to be followed up by research. Therefore, educational research at a systemic level was needed - rather than classroom-based research in connection with teachers' professional tasks within the framework of a comprehensive system. The year 1950 marked the starting point for a centralisation process and a gradual exclusion of teachers from research and development as well as from their control of professional further education (Hermansson, 1974).

1950-1980 – teachers as objects of research and change

During this period a comprehensive school system for primary and lower secondary school[2] as well as for upper secondary school[3] was established. The period was characterised by the centralisation of the school administration and a large-scale expansion of educational research. Although at the beginning of this period teachers were involved in developing practical solutions to problems that arose in the new comprehensive schools, they were gradually marginalised from the development aspect and turned into implementers of the results of educational research. Educational research was carried out in relation to reform and policy issues.

---

[2] The first National Curriculum for the nine-year comprehensive school was adopted in 1962.
[3] With its first National Curriculum from 1971.
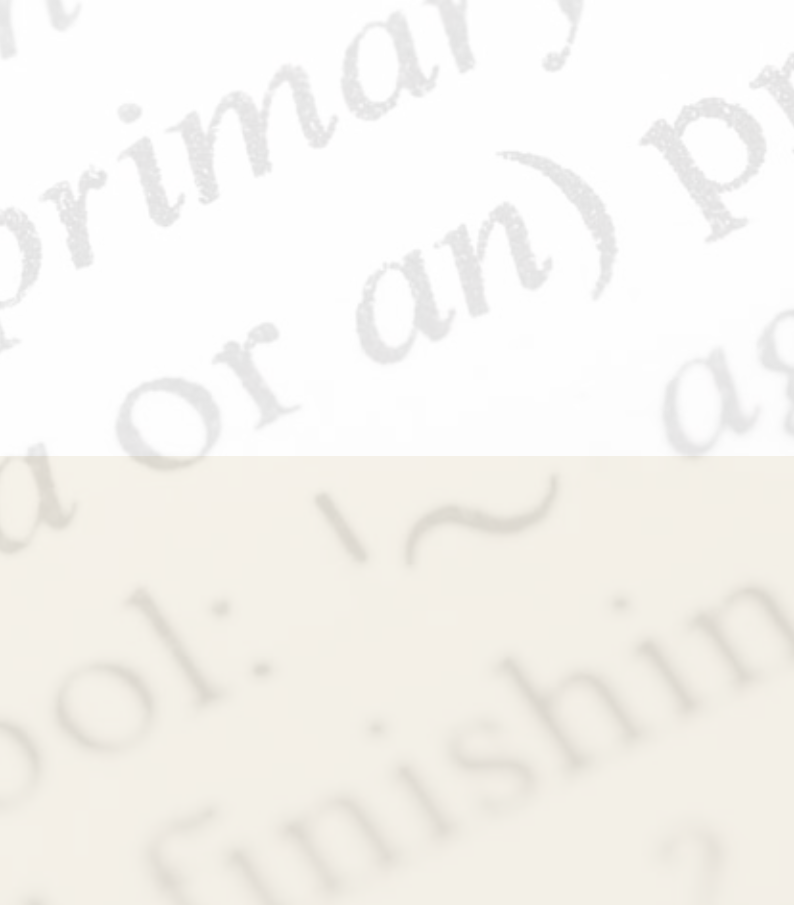
Gradually the gap widened between teachers and educational researchers (who at this time came largely from the field of psychology). Teachers changed from a position as subjects involved in designing and developing solutions to one in which they became part of the problem and an object of research. Phenomena such as "teacher resistance" and "unwillingness to change" became key research areas. Implementation research emerged as an important branch of research with a special focus on how teachers would not adopt new ideas or how they distorted them when they were implemented (Lindblad, 1980, Sandström & Ekholm, 1984).

After 1980 – teachers back as subjects?
Several trends emerged during the 1980s (e.g. decentralisation, the introduction of choices for students, profiled schools and marketisation). These were incorporated into the restructuring reforms of the 1990s (Lindblad & Popkewitz, 1999, Carlgren & Klette, 2008) which were similar in many ways to the restructuring changes being implemented in other countries. The reforms were characterised by a mixture of professional, bureaucratic and market-steering mechanisms.

Rhetorically the reforms were framed by ideas of decentralisation and professionalisation. The National Curriculum of 1994 was a minimal curriculum, giving teachers a great deal of latitude for decision-making. However, after four decades of centralisation teachers were used to receiving instructions from above, not to finding solutions to the problems, as they were now being asked to do. There were no resources for research and development activities connected with the new responsibilities assigned to the teachers. Research resources were instead given to "traditional" academic research in order

to inform decision-makers about what was going on in the schools – i.e. research from "above" rather than from "below". Gradually the language of decentralisation and the professionalisation of teachers were replaced by a language of centralisation, marketisation and control.

However, during the past decade the interest in encouraging teachers to carry out research has been rising. Graduate programmes for teachers have been developed and research schools focusing on issues of domain-specific teaching and learning have been established. Yet the dominant educational discourse in Sweden still looks upon teachers as consumers rather than producers of research results. The tension between traditional educational academic (social science) research and the school sector is growing. From an academic perspective the problem is that schools do not use existing research results, while from a school perspective educational research lacks relevance.

Professionals without tools and resources – a structural obstacle to progress
Prior to the comprehensive school reforms, teachers were important actors in school development and advocates of educational research. Parallel to the implementation of the comprehensive school reforms the situation changed into what appears today to be a structurally non-progressive situation. Although the restructuring reforms of recent decades have called for more professional steering, the conditions for this are lacking. Given the experiences of the last half century, it is not surprising that teachers today lack the tools and infrastructure needed to be more actively involved in innovation and knowledge formation.

Educational versus medical research

In light of the commonly held view that priority should be given to establishing research-based development of school activities, the medical sector is an interesting case for comparison. In contrast to the situation in education, clinical medical practices are generally perceived to be based on research and scientific knowledge. What are the critical differences when it comes to the role and position of research in these two fields?

The amount of research

In Sweden the amount of resources allocated to medical research is about 20 times the amount invested in educational research.[4] If we take the size of the two fields into consideration, the difference would be even greater. One implication of this concerns the kind of questions that can be the topic of inquiry. More resources open up avenues for inquiries into many more aspects of an issue and for performing more detailed analyses of more precise, concrete questions. A lack of resources entails a risk of asking too many questions at the same time, resulting in research that lacks precision and depth. In this case, the research results will not be particularly relevant for classroom practice.

One can ponder whether it is the situation in medicine or in education that is unique. I will argue that it is education. In knowledge intensive firms it is normal practice to allocate at least 10 per cent of the budget to research and development activity. In education this figure is less than 0.1 per cent (Alexandersson, 2006). The absence of resources al-located to research and development activity in the schools is noteworthy.

At the same time it is not unusual for up to 10 per cent of a municipality's school budget to be earmarked for the professional development of teachers. Perhaps a shift from investing so much in professional development, in which teachers are expected to receive information and knowledge, to investing more in research and development, in which teachers are expected to take part in the production of knowledge and information, would accomplish the crucial transformation of education into a progressive sector?

The kind of research

Another aspect of this issue is the *kind of research* that is carried out. In medicine there are different sources for research as well as different kinds of research, and these complement each other. The state allocates funds to graduate programmes and resources for basic research, research councils award funding to the best projects (in competition with other researchers) and other actors (such as county councils and pharmaceutical companies) provide resources for more clinical research. Most of this research is clinical research, i.e. inquiries into problems more directly connected to disease and the search for cures. Due to the large amount of clinical research, an epistemic culture has developed in connection with medical clinical practices, in turn spawning an orientation towards innovation and new knowledge. The ratio of non-clinical to clinical research is interesting. A large amount of clinical research is probably needed in order for epistemic

---

[4] This is not documented information but an estimate based on the research resources allocated to medicine and education by the state (universities and research councils), county councils and municipalities. It is probably not an overestimation.
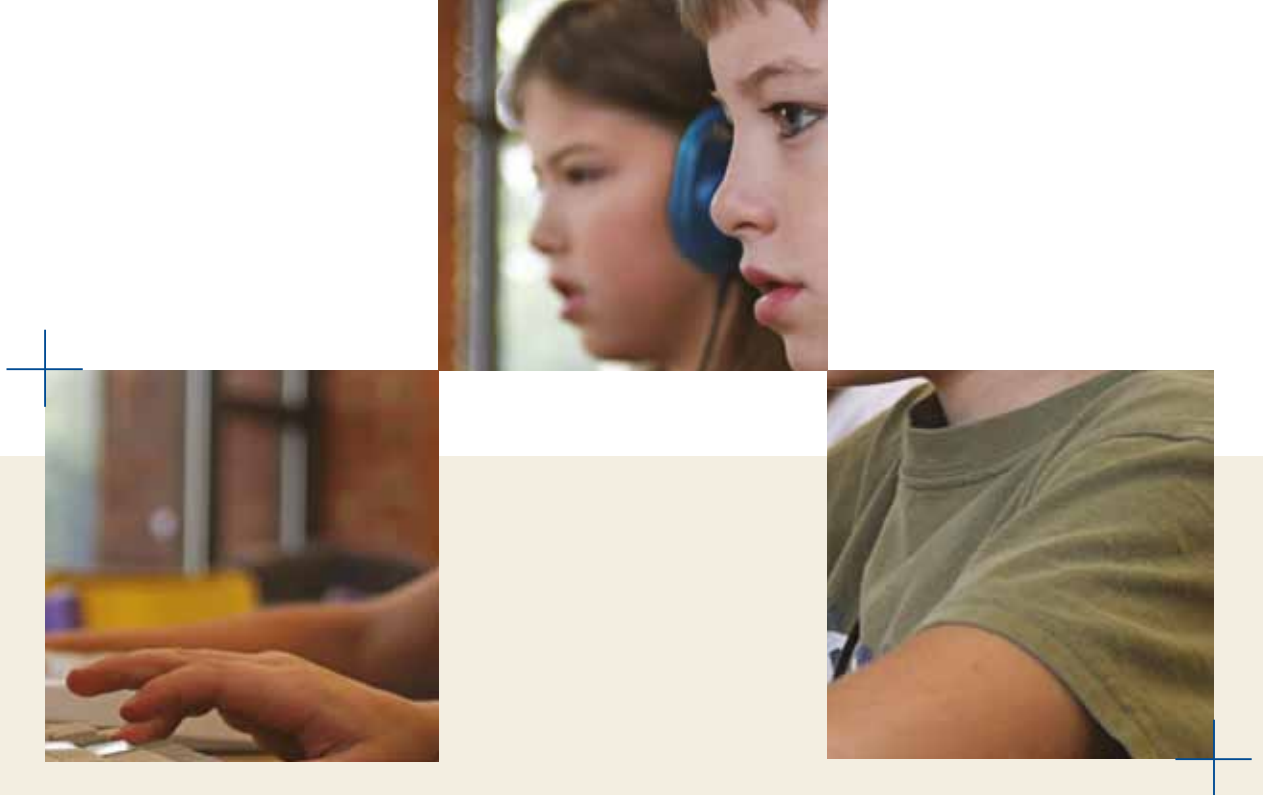
cultures to develop. Although basic research is of great importance for clinical progress, clinical research is what moves the profession forward. The main objective of clinical research is not to build theory. It focuses instead on understanding clinical problems and on developing and testing solutions. Clinical research can therefore be distinguished, not by its methods or quality criteria, but rather by its research objects and purposes (knowledge interest).

In education the situation is the opposite. Most research resources are invested in basic research that is not directly related to the problems of teaching and learning in schools.

These differences are not discussed when the field of medicine is used for comparison - neither the amount nor the kind of research. Instead the idea of evidence-based practice is offered as the solution for education. However, this is problematic against the backdrop of the insufficient research framework – above all the lack of "clinical" research. In contrast to the situation in the medical field (where evidence is based on syntheses of large numbers of clinical studies) the educational "evidence" is based on studies "on" rather than "in" education – leading to conclusions regarding the framework of teaching rather than teaching as such. Although such results are of interest to policymakers and administrators of education, they are of little use to teachers. As an analogy, it would be as if medical research focused mainly on doctors and hospital activities rather than on patients and diseases.

Conclusions
The "audit society" (Power, 1997) is characterised by decentralisation, self-autonomy and self-organisation (inner steer-ing) on the one hand and centralised control systems (outer steering) on the other. Rather than steering processes, the focus is on evaluation of the outcomes. An important aspect of self-autonomy is professional steering, i.e. professionals take action informed by their specific "bodies" of knowledge. While in the early 1990s teachers in Sweden were given greater latitude in professional decision-making (Carlgren, Klette & Simola, 2008), the previous decade has been characterised by a growth of control measures. This has resulted in an uneven balance between inner and outer steering mechanisms, thereby creating a risk of reduced quality in education and a de-professionalisation of teachers. One problem is that the professional knowledge base of teachers is too weak. Teachers are civil servants rather than professionals in the sense that they do not have access to a specific body of knowledge they can use as a basis for formulating professional tasks and finding solutions to professional problems. If their professional knowledge base is not strengthened, teachers will become what Manuel Castells (2000) calls "flexible knowledge workers" rather than professionals.

The most obvious conclusions to draw from a comparison with the field of medicine are that the following two changes must take place for education to be transformed into a research based activity:
1. a significant expansion of research activities (possibly through the transformation of resources for further education into R&D resources), and
2. a differentiation of research activities and the establishment of "clinical traditions" in educational research.
However, it is important that the field of education develops its own research traditions which are sensitive to the par-

ticularities of educational activities. Education is a human construct and the knowledge needed is knowledge about how educational activities can be formed in order to create affordances for human cultivation. Rather than imitating medical clinical research, perhaps educational "clinical" research should develop as formative or design research (Collins, Joseph, & Bielaczyc, 2004). "Design experiments" (Brown, 1992, Cobb, P., Confrey, J., diSessa, A., Lehrer, R. & Schauble, L, 2003) are one example of this; another is the Japanese lesson study tradition, which in contrast to design experiments involves teachers as designers as well as formulators of research problems.[5] The long tradition of teacher-owned development activity appears to have resulted in a culture of uncertainty and an orientation towards innovation and knowledge generation in Japan.

The introduction of lesson studies in the reform-permeated US school cultures has not, however, been without problems. This is not surprising in light of the historical developments described above, which according to Zeichner and Noffke (2001) have been very similar in the US. On the one hand, the cultures in schools are characterised by almost half a century of centralised steering and a lack of orientation towards innovation and the generation of new knowledge. On the other hand, academic research has imperialistic ambitions. Yet the interest in and spreading of lesson studies as well as other practice-developing research approaches may mean that that this will lead to the establishment of the missing link, i.e. "clinical" research traditions in education.

References

Alexandersson, M. (2006). Praxisnära forskning och läraryrkets vetenskapliga bas. In B. Sandin & R. Säljö (ed.) *Utbildningsvetenskap – ett kunskapsområde under formering*, 355-376. Stockholm: Carlsson Bokförlag.

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

Carlgren, I. (1986). Lokalt utvecklingsarbete. Göteborg studies in educational sciences. *Acta Universitatis Gothoburgensis, 56.*

Carlgren, I., Klette, K & Simola, H. (2002). Restructuring, the mission of schooling and teacher professionalism. In Klette, K., Carlgren, I., Rasmussen, J. (eds.) Restructuring Nordic Teachers: *Analyses of Interviews with Danish, Finnish, Swedish and Norwegian Teachers. Report no. 3, 2002*. University of Oslo, Institute for Educational Research.

Carlgren, I. (2006). Är utbildningsvetenskap den vetenskap utbildningen behöver? In Sandin, B. & Säljö, R. (eds.) *Utbildningsvetenskap – ett kunskapsområde under formering*. Stockholm: Carlssons förlag.

Carlgren, I. & Klette, K (2008). Reconstructions of Nordic Teachers: Reform Policies and Teachers' Work during the 1990s. *Scandinavian Journal of Educational Research*, 52:2, 117-133.

Carlgren, I. (2009) Lärarna i kunskapssamhället – flexibla kunskapsarbetare eller professionella yrkesutövare? *Forskning om undervisning och lärande, 2,s.9-27*. Stiftelsen SAF I samarbete med Lärarförbundet

Cobb, P., Confrey, J., diSessa, A., Lehrer, R. & Schauble, L. (2003). Design experiments in educational research. ER 32, 1:9-13.

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

Fernandez, C., Cannon, J. & Chokshi, S. (2003). A US-Japan lesson study collaboration reveals critical lenses for examining practice. *Teaching and Teacher Education*, 19 (2003), 171-185.

Hermansson, E. (1974). *Upplevelser och påverkan. Jämförelsematerial för pedagogiskt intresserade*. Stockholm: Sveriges Lärarförbund.

Hiebert, J., Gallimore, R & Stigler, J. (2002). A knowledge base for the teaching profession: what would it look like and how can we get one? *Educational Researcher*, 31:5, 2-15.

Knorr Cetina, K. (2001). Objectual practice. In Schatzki, Knorr Cetina, K. & von Savigny, E. (eds.) *The practice turn in contemporary theory*. London, New York: Routledge.

Lindblad, S., & Popkewitz, T. (eds.) (1999). Education governance and social integration and exclusion: International cases of educational systems and recent reforms. *Uppsala Reports on Education No 34*, Uppsala, Sweden: University of Uppsala.

Marton, F. & Pang, M.F. (2006). On Some Necessary Conditions of Learning. *Journal of the Learning Sciences*, Volume 15, Issue 2, April 2006, 193-220.

Marton, F. & Ling, L.M. (2007). Learning from "The Learning Study". *Tidskrift för lärarutbildning och forskning.*

Lewis, C. (2000). Lesson Study: *The Core of Japanese Professional Development*. Invited Address to the Special Interest Group on Research in Mathematics Education. American Educational Research Association Meetings, New Orleans April 28, 2000

Power, M. (1997). *The audit society – Rituals of verification.* New York: Oxford University Press.

Sandström, B. & Ekholm, M. (1984). *Stabilitet och förändring i skolan*. Liber, Stockholm.

Stigler, J. & Hiebart, J. (1999). *The teaching gap. Best ideas from the world's teachers for improving education in the classroom*. The Free Press.

Zeichner, K. & Noffke, S. (2001). Practitioner research. In Richardson, V. (ed.) *Handbook of research on teaching. Fourth edition*. American Educational Research Association. Washington D.C.

---

[5] As a result of a Swedish-Chinese cooperative project, a model called "learning studies" has developed as a combination of design experiment and lesson studies (Marton & Pang, 2006, Marton & Ling, 2007).