CONTENT

Editorial

THEMATIC SECTION

Christina Segerholm The Quality Turn: Political and Methodological Challenges in Contemporary Educational Evaluation and Assessment

Xavier Pons The Turn and the Paths: School External Evaluation in England, France and Switzerland: A Sociological Approach

Laura S. Hamilton, Brian M. Stecher & Kun Yuan Standards-Based Accountability in the United States: Lessons Learned and Future Directions

Peter Dahler-Larsen Constitutive effects as a social accomplishment: A qualitative study of the political in testing

Barbara Crossouard Classroom assessment and education: Challenging the assumptions of socialisation and instrumentality

D. Royce Sadler Assessment, evaluation and quality assurance: Implications for integrity in reporting academic achievement in higher education

Thomas A. Schwandt Quality, Standards and Accountability: An Uneasy Alliance

OPEN SECTION

Maria Hedlin Admission policy of Swedish teacher education favouring men: Discussion in Parliament in 1962

Stig-Börje Asplund Being a skilled reader: Reception patterns in vehicle engineering students’ literature discussion

Anna-Carin Jonsson & Dennis Beach Predicting the use of praise among pre-service teachers: The influence of implicit theories of intelligence, social comparison and stereotype acceptance

Kerstin Bergqvist “Own work” in primary school – A teaching and learning practice in the context of administration and control
EDUCATION INQUIRY

Education Inquiry is an international on-line, peer-reviewed journal with free access in the field of Educational Sciences and Teacher Education. It publishes original empirical and theoretical studies from a wide variety of academic disciplines. As the name of the journal suggests, one of its aims is to challenge established conventions and taken-for-granted perceptions within these fields.

Education Inquiry is looking for lucid and significant contributions to the understanding of contextual, social, organizational and individual factors affecting teaching and learning, the links between these aspects, the nature and processes of education and training as well as research in and on Teacher Education and Teacher Education policy. This includes research ranging from pre-school education to higher education, and research on formal and informal settings. Education Inquiry welcomes cross-disciplinary contributions and innovative perspectives. Of particularly interest are studies that take as their starting point, education practice and subject teaching or didactics.

Education Inquiry welcomes research from a variety of methodological and theoretical approaches, and invites studies that make the nature and use of educational research the subject of inquiry. Comparative and country-specific studies are also welcome.

Education Inquiry readers include educators, researchers, teachers and policy makers in various cultural contexts.

Every issue of Education Inquiry publishes peer-reviewed articles in one, two or three different sections. Open section: Articles sent in by authors as part of regular journal submissions and published after a blind review process. Thematic section: Articles reflecting the theme of a conference or workshop and published after a blind review process. Invited section: Articles by researchers invited by Education Inquiry to shed light on a specific theme or for a specific purpose and published after a review process.

Education Inquiry is a continuation of the Journal of Research in Teacher Education, which is available in printed copies as well as electronic versions and free access at http://www.use.umu.se/forskning/publikationer/lof/

Editors
Associate Professor Nafsika Alexiadou, Umeå University, Sweden
Assistant Professor Linda Rönnberg, Umeå University, Sweden

The editorial board
Professor Marie Brennan, Victoria University, Australia
Professor Bernard Cornu, Directeur de la Formation – CNED, Directeur de CNED-EFAD, France
Professor Per-Olof Enson, Umeå University
Professor David Hamilton, Umeå University, Sweden
Professor Brian Hudson
Professor Gloria Ladson-Billings, University of Wisconsin-Madison, USA
Professor Martin Lawn, University of Edinburgh, UK
Assistant Professor Eva Lindgren, Umeå University, Sweden
Professor Lisbeth Lundahl, Umeå University, Sweden
Assistant Professor Linda Rönnberg, Umeå University, Sweden
Professor Kirk Sullivan, Umeå University, Sweden
Associate Professor Manya Sundström, Umeå University, Sweden
Professor Gaby Weiner, University of Edinburgh, UK
Professor Pavel Zgaga, University of Ljubliana, Slovenia

Language Editor
Murray Bales, Ljubljana, Slovenia

Guidelines for Submitting Articles
See Education Inquiry’s homepage: http://www.use.umu.se/english/research/educationinquiry
Send Manuscripts to: EducationInquiry.Editor@adm.umu.se

©2012 The Authors. ISSN online 2000-4508
The Quality Turn
Political and Methodological Challenges in Contemporary Educational Evaluation and Assessment

Christina Segerholm*

In December 1972 a group of scholars engaged in educational evaluation and curriculum reform assembled, in what was later to be called the Cambridge conference, to challenge what they viewed as the then “traditional style of evaluation” (MacDonald & Parlett 1973:74). Their aim was to develop guidelines for the future of educational evaluation, and to challenge the traditional evaluation style. In the present policy context, it is instructive to revisit the manifesto that summarised this conference (MacDonald & Parlett 1973) to see what these scholars wanted to challenge, what they viewed as problematic and what they wanted to promote.

The traditional style of evaluation they wanted to move away from was described in the manifesto as a style when “the task of the evaluator was to determine the congruence of pupil performance and project objectives, i.e. to assess the extent to which pupils exposed to a new curriculum achieved its intended learning outcomes” (MacDonald & Parlett 1973:75). Further, these scholars traced the traditional evaluation style to a model, a prototype “launched by Ralph Tylor in the early 1930s” (Tyler 1949) and from a “securely rooted tradition of educational measurement on both sides of the Atlantic” (MacDonald & Parlett 1973:75). The policy context at the time of the conference (i.e. the late 1960s and early 1970s), was described as follows:

Federal policy-makers demanded of educational innovators that they both pre-specified the intended performance gains and provided subsequent proof of ‘pay-off’. Despite mounting criticisms of the engineering-type assumptions of such ‘pre-ordinate’ evaluation, and the tentative emergence of alternative approaches, the model was still serviceable enough to be exported to us, and to command the allegiance of the first wave of evaluators in Britain (MacDonald & Parlett 1973:75).

In this quotation, the authors who summarised the Cambridge event talk about two things that are strikingly similar in today’s education policy context: a) the over-reliance of a particular rationale in education reform and evaluation and assessment

*Department of Education, MidSweden University, Härmösand, Sweden. E-mail: christina.segerholm@miun.se
©Authors. ISSN 2000-4508, pp.115–122
(measurement of the relation between the pre-defined objectives and students’ achievement/performance); and b) the ‘export’, or policy transfer/learning, of this policy of education reform and evaluation (in this case from the USA to Britain).

During the conference the group concurred that this mode of evaluation had failed because of the difficulties to agree on clear goals in advance; the problems to measure pre-defined goals with validity and reliability, and because the model did not represent the complexities inherent in education practice. Simply put, they called for a variety of methodological approaches to better capture these ‘complexities’. They also discussed the role of the evaluator (at that time most often externally contracted) vis-à-vis the decision-makers, and whether or not to design the evaluation in advance pointing to the risk of not being able to attend to issues that arise during an evaluation process (MacDonald & Parlett 1973:75-78).

During the years between 1972 and 1995 another four Cambridge conferences were arranged in which evaluation issues and alternative approaches of educational evaluation were discussed and developed (Elliott & Kushner 2007). Indeed, evaluation approaches like democratic (MacDonald 1974), illuminative (Parlett & Hamilton 1972), case study (e.g. Simons 1971), and responsive evaluation (Stake 1967) gained ground in an education policy climate that was expansive with post Second World War reforms. These innovative approaches were later followed by the development of numerous other approaches (for overviews, see e.g. Shadish, Cook & Leviton 1995, Owen and Rogers 1999, Vedung 2010).

However, with economic recession and a policy drift to neoliberal agendas in the 1980s and later, the adoption of management ideas from the private sector, and a governing rationale based on goal and results steering, evaluative activities increased and entire evaluation systems were invented. These systems developed as results of political decisions like the No Child Left Behind Act in the USA (see Hamilton et al. this issue) and are based on the same logic that underpinned evaluations at the time of the first Cambridge conference. On top of that, and sometimes included in evaluation systems (as quality indicators) are quality assurance systems/activities, their rationale also imported from the private sector aimed at securing production processes leading to products of sufficient quality in a market. Concerned about this old/new situation a sixth Cambridge conference convened in 2004 to revisit the first manifesto, to “critique the prevailing assumptions that frame current educational evaluation and methodological discourses that shape the conduct of educational evaluation.” (Elliott & Kushner 2007:325), and to create a new manifesto in the light of the new policy context. The new manifesto observes the problematic role of evaluation and assessment in relation to policy- and decision making, when merit and worth (now called quality) in education first and foremost is about the relation between pre-defined learning outcomes and the performance of students, and by that diminishes the moral/political obligation to foster democratic citizens (Elliott & Kushner 2007).
The quality turn

Since then (2004), the expansion of evaluative activities has continued and escalated internationally, and definitely in Sweden. The term “quality” has been introduced in education policy (Bergh 2010, Dahler-Larsen 2008) and is frequently used, attached to, or as a prefix to: education, evaluation, assurance, assessment, control, incitement etc. As an example, the term is used 263 times in 33 pages (mean: eight times per page) in the Swedish government’s proposal for the new quality (!) evaluation system in higher education now being implemented (Prop. 2009/10:139).

Forty years after the first Cambridge conference and ten years after the sixth, great efforts are now being taken through quality assurance systems to secure (good) outcomes of education; through evaluation to check, oversee, inspect and improve quality in education. We have called this present policy focus on quality in education the Quality Turn and, by doing so, aspire to draw attention to the need to critically scrutinise the language of education quality, the policy of education quality and, not least, the practices of measuring quality in education and its consequences. Evaluators, education researchers and scholars are implicated in these activities in different ways. This fact and the need to continue to discuss the questions raised in the Cambridge conferences, and open conversations about issues like: which values underpin the presently felt need to secure and assess quality?; what do these activities do to education?; and, how do these evaluative activities affect our perception of what education and human development is and should be about? led to the organisation of the mini-symposium The Quality Turn: Political and methodological challenges in contemporary educational evaluation and assessment.

The aim of the symposium was to offer opportunities for, and inspire dialogue about this Quality Turn, where taken-for-granted assumptions could be challenged and pave the way for future studies from critical perspectives. Another aim was to facilitate international contacts for younger researchers. Internationally recognised scholars were invited and most of the presentations are now published in this journal.

Systems of assessing quality in education

The interest in measuring quality in education displayed by policy-makers in several countries takes different forms in national and local contexts. Yet there are similarities and of particular significance are the efforts to develop ways to assess, assure and promote quality in education through comprehensive systems of evaluative activities. As already shown with the first Cambridge conference, ideas of how to do this are spread between countries. Ideas are transferred, translated, mediated and learned through organisations like the OECD (via e.g. country reviews, PISA), the EU (through e.g. the Lisbon Strategy), and the SICI (The Standing International Conference of Inspectorates) (Ozga et al. 2011).

In a time when scholars, policy-makers, evaluators and civil servants meet, trade, broker and learn from each other, from the high speed information of publications,
and from comparisons of different assessments, and when the use of certain semantic magnets like quality and learning are widespread, it is easy to conclude that the same practices will occur, spread and transgress national boundaries. But there is no simple mimicking of systems of evaluative activities across borders. Xavier Pons reports in *The turn and the paths. School external evaluation in England, France and Switzerland: A sociological approach* on a study of school inspection in those three countries, and their different take on inter- and transnational notions of external evaluation. Although similar terms are in motion in the education policy rhetoric (like ‘quality’, and ‘performance’), Pons shows that the particular national, regional (cantons in Switzerland) and local (academies in France and Local Education Authorities in England) contexts play a significant role in the translation of what could be called an international evaluation and quality assessment discourse. Further, Pons also demonstrates how the institutional, professional and cognitive sources of the interdependencies of policy actors who are involved in these evaluative activities are different in the three countries, and predetermine how international evaluation policies are translated and used nationally and locally.

Laura Hamilton, Brian Stecher and Kun Yuan give another example of how policy context matters in the configuration of educational evaluation and assessment policies and practices. Their article *Standards-based accountability in the United States: Lessons learned and future directions* shows findings on how outcomes (tests) based governing relates to accountability and student learning. They do this by disentangling the different suppositions that the No Child Left Behind Act (NCLB) and the reauthorisation of this law rest on, many of them quite similar to what have been adopted in several other countries. In so doing it becomes clear how the NCLB Act requires all states to develop standards and tests along with sanctions tied to the outcomes, resulting in a view of accountability as more or less solely tied to test scores. Further, they show how the stress to alignment of tests to the pre-defined standards and the considerable influence of the private sector in the development of the tests drive out curricular content that is not tested. The authors also discuss the problem of the quality of the standards themselves (also called learning outcomes, objectives), which is something common to all education systems that subscribe to this governing rationale, and were also observed at the first Cambridge conference. Hamilton, Stecher and Yuan also propose some measures to overcome the most threatening effects of the US system of evaluative activities.

As the article by Hamilton et al. clearly shows, mandated tests are parts of more comprehensive education policies, hence also related to accountability issues. When outcomes measures are closely linked to other areas of importance like student retention, teacher salaries, principals’ contracts, school finances, or to other policy areas like immigration policy, consequences are often noticeable. In his article *Constitutive effects as a social accomplishment: A qualitative study of the political in testing* Peter Dahler-Larsen presents a telling example of how the testing system in Danish for im-
migrants affect how the teachers’ talk about the students, how teaching is organised, what is taught, and the meaning of being a professional teacher. He also demonstrates the importance of taking the particular national policy context into account in order to understand these particular effects of a testing system, pointing to its connection to the permit to stay in the country, and to the financial incitements for the school when students pass the tests.

The quality in student assessment

In several evaluation and quality assurance systems in education different kinds of aggregated measures of student performance or achievement/attainment are an integral part. This means that processes of student assessment, grading, testing and marking are part of the evaluation of quality in entire national education systems. Different kinds of student assessment are practiced and favoured in different education systems, policy contexts, and by practitioners. All students and teachers know that the assessments and grades matter for individuals’ future careers, but individual measurements and the quality of them may also have other consequences.

Drawing on a study conducted in two case schools, Barbara Crossouard in her article *Classroom assessment and education: Challenging the assumptions of socialisation and instrumentality* questions the view of formative assessment as the now preferred assessment practice by policy-makers in the United Kingdom. In a policy environment characterised by governing by pre-defined objectives and the measurement of outcomes (attainment), quality in formative assessment easily leads to a quest to document and keep track of individual students’ accomplishments. Crossouard points here to the risk of emphasising easily observable behaviours in the formative assessment. This is because formative assessment in such a context has to rely on language descriptors that never fully represent the tacit nature of assessment of more complex tasks and knowledge. Through the materials from the two case schools, Crossouard shows that formative assessment can become a technology sustaining an instrumental view of learning to pre-defined goals, thereby also neglecting reflexive thoughts about the inherently political in education and learning, i.e. the formation of human subjects. Instead, she argues for the use of formative assessment that enables the assessing of both students’ attainment in relation to pre-defined objectives and non-specified outcomes, the latter supporting open-ended and creative learning processes.

In a similar argument, Royce Sadler in *Assessment, evaluation and quality assurance: Implications for integrity in reporting academic achievement in higher education* takes up the issue of quality in education, as a matter of the quality in grading, and critically examines practices of grading in higher education. He does this by discussing how well student achievement is represented by course grades. Sadler argues that not only do course grades give information about the achievement of individual students, but they are quite frequently also used as indicators of the quality of a course, a programme, in aggregated forms to report on performances of students.
where these figures are understood as a measure of ‘production’, sometimes tied to state grants (as shown in Dahler-Larsen, this issue). Sadler describes three common grading practices in higher education and their underpinnings, and shows why they all lack the quality necessary for appraising student achievement adequately. Pointing to the need for course grade integrity based on fidelity, commensurability and comparability, he also outlines how to improve grade integrity and to thereby also improve quality in higher education.

**The quest for quality in education – balancing reasonable accountability and human responsibility**

The contemporary political intentions to improve education globally are hard to question. What needs to be questioned is the Quality Turn, its underlying assumptions about the meaning of education as a goal rational process, mainly aimed at advancing economic growth. Economic growth may perhaps be needed in order to build welfare societies, but welfare societies also need citizens who are engaged and able to find new solutions to new societal problems. Education as a pre-defined set of learning outcomes, measured, assessed, evaluated and assured through performance measures, and carried out as a rational, route free from failures and surprises, hardly prepare humans for an ordinary life for what it takes to build a sound, socially just and democratic society. Still we also need to have some idea about how public resources are used so that we can exercise our political and human rights in forming the future.

In the final article *Quality, standards, and accountability: An uneasy alliance*, Thomas Schwandt discusses this balance – the importance of standards, accountability and quality and the parallel risks of standardisation, over-regulation and control. As a conclusion, commentary on the previous articles, and response to the theme of the symposium, Schwandt inspires us to extend our concerns to social justice globally, and provides us with a route to critically examine our own standpoints and the subsequent challenges to the Quality Turn we ought to provide.

---

*Christina Segerholm* is Professor in Education at the Department of Education, MidSweden University. She performs studies on the influence of different evaluative activities on education policy and practice, and is currently leading a research project on school inspection and governing.
References

Bergh, A. (2010) Vad gör kvalitet med utbildning? Om kvalitetsbegreppets skilda innebörder och dess konsekvenser för utbildning. [What does quality do to education? Different meanings of the concept of quality and their consequences for education. In Swedish. Diss.] Örebro: Örebro Universitet.

Dahler-Larsen, P. (2008) Kvalitetens beskaffenhed. [The nature of quality. In Danish.] Odense: Syddansk Universitetsforlag.

Elliott, J. and Kushner, S. (2007) The need for a manifesto for educational programme evaluation. Cambridge Journal of Education 37(3), 321–336.

MacDonald, B. (1974) Evaluation and the control of education. In MacDonald, B. and R. Walker (eds.) Innovation, evaluation, research and the problem of control. Norwich: CARE, University of East Anglia.

MacDonald, B. and Parlett, M. (1973) Re-thinking evaluation. Notes from the Cambridge conference. Cambridge Journal of Education 3(2), 74–82.

Owen, J. and Rogers, P. (1999) Program evaluation: Forms and approaches. Thousand Oaks: Sage Publications.

Ozga, J., Dahler-Larsen, P. Segerholm, C. and Simloa, H. (eds.) (2011) Fabricating Quality in Education. Data and Governance in Europe. London: Routledge.

Parlett, and Hamilton, D. (1972) Evaluation as illumination: A new approach to the study of innovative programmes. In D. Hamilton, D. Jenkins, C. King, B. MacDonald and M. Parlett (eds.) Beyond the Numbers Game, A Reader in Educational Evaluation. London: Macmillan.

Proposition 2009/10:139 Fokus på kunskap – kvalitet i den högre utbildningen. [Focus on knowledge – quality in higher education.] Accessed 19 March 2012 at http://www.regeringen.se/sb/d/12489/a/142211

Shadish, W.R., Cook, T.D. and Leviton, L.C. (1995) Foundations of Program Evaluation. Theories of Practice. Newbury Park: SAGE.

Simons, H. (1971) Innovation and the case study of schools. Cambridge Journal of Education 1(3), 118–123.

Stake, R.E. (1967) The countenance of educational evaluation. Teacher College Record 68(7), 523–540.

Tylor, R.W. (1949) Basic Principles of Curriculum and Instruction. Chicago: University of Chicago Press.

Vedung, E. (2010) Four waves of evaluation diffusion. Evaluation 16(3), 263–277.
Endnotes

1 The articles in the thematic section are all based on presentations given at the mini-symposium The Quality Turn: Political and Methodological Challenges in Contemporary Educational Evaluation and Assessment, 15–16 June 2011 at Umeå University. The symposium was made possible by economic support from the Swedish Research Council, Educational Sciences (grant no. 722-2010-0640); from Umeå University’s conference fund; from the Department of Education at Umeå University; from the Department of Applied Education at Umeå University; and from the Unit of Educational Measurement at the Department of Applied Education at Umeå University.
CONTENT

Editorial

THEMATIC SECTION
Christina Segerholm The Quality Turn. Political and Methodological Challenges in Contemporary Educational Evaluation and Assessment
Xavier Pons The Turn and the Paths. School External Evaluation in England, France and Switzerland: A Sociological Approach
Laura S. Hamilton, Brian M. Stecher & Kun Yuan Standards-Based Accountability in the United States: Lessons Learned and Future Directions
Peter Dahler-Larsen Constitutive effects as a social accomplishment: A qualitative study of the political in testing
Barbara Crossouard Classroom assessment and education: Challenging the assumptions of socialisation and instrumentality
D. Royce Sadler Assessment, evaluation and quality assurance: Implications for integrity in reporting academic achievement in higher education
Thomas A. Schwandt Quality, Standards and Accountability: An Uneasy Alliance

OPEN SECTION
Maria Hedlin Admission policy of Swedish teacher education favouring men: Discussion in Parliament in 1962
Stig-Börje Asplund Being a skilled reader: Reception patterns in vehicle engineering students’ literature discussion
Anna-Carin Jonsson & Dennis Beach Predicting the use of praise among pre-service teachers: The influence of implicit theories of intelligence, social comparison and stereotype acceptance
Kerstin Bergqvist “Own work” in primary school – A teaching and learning practice in the context of administration and control