Systematic quality work in Swedish schools. Intentions and dilemmas

Ulf Lundström*

Ulf Lundström
Department of Applied
Educational Science
Umeå University
ulf.lundstrom@edusci.umu.se

Abstract
This article analyses the intentions and internal logic of a systematic quality work programme in Swedish schools as an example of institutionalization of an international super-standard in large public organizations. A programme theory analysis is used to illuminate the intentions, inner logic, involved actors, possible limitations and underlying assumptions of the evaluation system, which together constitute the systematic quality work. What is realistically achievable is also highlighted and analysed in relation to current research and the welfare liberal, and social democratic education policy perspectives expressed in the National curriculum. The programme analysis indicates an ambitious, all-embracing, and ever-present evaluation system and exposes inherent potential contradictions. The programme could foster improvement and mirrors the distribution of responsibilities that characterizes Swedish school governance, which implies that *how* the programme is enacted in schools will be decisive for its results. The potential of the programme in practice is discussed in relation to the varying local context's inherent risks and dilemmas, concerning governance, reductionism, constitutive effects, and working conditions.

Introduction

In 2012, Sweden's National Agency for Education (NAE) (2012a) issued the document "General Advice for systematic quality work in schools" (henceforth, "General Advice"). The present article analyses this document as exemplifying the potential of an evaluation system within the international quality trend, which serves as an international super-standard or organizational recipe (Rövik 2000) adopted in most sectors of society today. The General Advice can be perceived as a response to the persistent signs of declining student achievement and declining student performance equity between schools and municipalities, problems prevalent in several countries whose self-perception before "the PISA shock"[1] was that they had high-quality educational systems.[2] The systematic quality work can also be regarded as a response to the OECD proposal that Sweden should develop a coherent framework for evaluation and assessment (Nusche et al. 2011). Furthermore, the enhanced quality work, prescribed in the General Advice, is in line with the international trend to couple mechanisms for generating performance data and school markets (Apple 2004: 18). Market and bureaucratic logics (Freidson 2001) may collide in the field of education (Lundström & Holm 2011) but an interest in accessible information on school/student performance is shared. School choice presupposes that the "customers" (i.e., students and parents) are well-informed, while democratic and managerial accountability both require communicable achievement indicators.

*Ulf Lundström is an associate professor at the Department of Applied Educational Science, Umeå University. He has a background as a teacher and principal in upper-secondary school and holds a Ph.D. in Educational Work. His research examines the teaching profession, evaluation and education policy. He is currently engaged in three research projects examining: how evaluation is enacted in comprehensive school, how inclusion aims can be paired with competition on a school quasi-market, and local safety policy.

Around 1990, drastic reforms, including far-reaching decentralization, de-regulation, school choice and governance by goals and results were introduced in the Swedish school system. Demands for more and better evaluations followed these changes. In the following period, the concept of "quality" (quality assurance, quality assessment, quality development, etc.) was established to denote the evaluation, assessment, and the increasingly extensive statistics of the education system. The present government has continued the efforts to realize these reforms, but has also taken several steps to recentralize education governance over the last few years (SOU 2014), as exemplified by the establishment of the Swedish Schools Inspectorate in 2008, and the issuance of the General Advice. The General Advice constitutes a framework for a local evaluation system whose ultimate aim is to improve school and student achievement. As the General Advice is strongly linked to the Education Act (2010), it is a crucial steering document for Swedish schools and their owners (i.e., municipalities and private providers). My intention is to analyse this evaluation system in the broader context of quality management in public administration. My analysis is not restricted to this case, but is intended to treat an example of the institutionalization of an elaborate quality concept in a large, decentralized public organization. This implies an interdisciplinary approach. The evaluation system in focus is related to three broad policies – international "organizational recipes" (Rövik 2000) or "travelling policies"[3] (Ozga & Jones 2006) – emphasizing quality, evaluation, and organizational learning/empowerment. A substantial body of research examines each of these concepts (see reviews in, e.g., Bergman & Klefsjö 2003, Dahler-Larsen 2012a, Fitzpatrick et al. 2004, Leeuw & Furubo 2008, Nusche et al. 2011, Rossi et al. 2004, Vedung 2010). However, little research has examined evaluation systems that unite all three concepts, even though they are dominant discourses in many organizations today. This article contributes to our understanding of the implications of that complexity. Furthermore, I have not found a programme theory analysis of a national evaluation system in society's largest institution, the school system.

This article seeks mainly to analyse the aims and internal logic[4] of the systematic quality work stipulated in the document "General Advice for systematic quality work in schools", which outlines an evaluation system for Sweden's school system. The external validity is analysed as well, i.e., the extent of the programme's consistency with its context and with other relevant research (Funnell & Rogers 2011).

The potential for unintended outcomes and what is realistically achievable will be highlighted and discussed in light of research knowledge and the "welfare liberal" (Olssen et al. 2004: 180) and "social democratic" (Arnesen & Lundahl 2006) education policy perspectives expressed in the national curriculum, in contrast to the "classical liberal" and "neoliberal" perspectives. In short, from the welfare liberal education policy perspective, education is egalitarian and a public good. The purpose of education is to promote both society's and individuals' interests and development: to develop ethical, social, cultural, and political awareness among all students, and to promote the integration of society in terms

of gender, ethnicity, and class. This definition largely overlaps Arnesen and Lundahl's (2006) definition of the social democratic perspective.

Guided by programme theory analysis, a set of questions is used to analyse the intentions, inner logic, potential, involved actors, possible limitations, and underlying assumptions of the systematic quality work programme. The analysis is extended "beyond a mere test of a given programme theory" (Dahler-Larsen 2001: 346), a constructivist inclination that makes the analysis sensitive to the programme's context. In their daily work, school staff enact – that is, interpret, translate, and adapt – policy from the programmatic level in forms assumed to work in their specific local contexts (Ball et al. 2012). Programme theory analysis is a tool for research into both evaluation systems and evaluation in practice, useful for better understanding programme strengths and weaknesses (Brousselle & Champagne 2010). Furthermore, it helps us to examine whether the programme has suffered from an implementation failure or a theoretical failure and to identify side effects (Dahler-Larsen 2001).

The following section considers the background of the international quality trend and its impact on the national educational context. Programme theory analysis is presented in the third section. In the next two sections, the programme theory analysis is conducted and discussed, and dilemmas and inherent risks related to the programme's context are explored in light of other relevant research. The final section presents the paper's conclusions.

Background and context

As the concept of quality is often taken for granted, there is good reason to describe its application and development in the Swedish school context from a broader perspective. The described Swedish development is part of an international trend: "In recent decades, we have seen a tremendous interest in quality as a strategic issue in the Western world" (Bergman & Klefsjö 2003). Education is not an exception: "The quest for quality has become even more focussed in the last decade as economic globalisation has increased the significance of quality education and international assessments of student performance provide measures for comparative appreciation of education results" (Ehren et al. 2013: 4).

Today's quality discourse is largely rooted in commercial manufacturing contexts. The quality discourse followed the dissemination of new public management ideas from the private to the public sector, where an ongoing struggle over the meaning of the concept started. Dahler-Larsen (2008) reasons about various meanings of the quality discourse: first, it is omnipresent; second, public concerns and problems are increasingly treated as quality issues; and, third, quality is organized, i. e., in every situation or context, there is always someone who defines the meaning of quality. In Europe, quality has been described as the key to success in the global knowledge economy and to the construction of the European project (Grek et al. 2009).

The international quality trend first had an impact on the Swedish public sector in the 1980s. This influence, in combination with others, such as the severe economic recession in Sweden at the time, public education debate, and new public management ideas, paved the way for extensive school reforms and new ways to govern schools around 1990. At that time, the previously strongly centralized and regulated school system became governed by goals and results, and highly decentralized – changes that presupposed stricter evaluation systems. Evaluation was described as "the corner stone of the prospective management by objectives" and "an important pre-requisite for school development" (Government Bill 1989/90:41, 9, *author's translation*). Evetts (2009, 249) describes the dilemma faced by several countries at the time, namely to create "smaller, cheaper and more effective governments and ... better public services and more professional practitioners". In such a context, the quality concept, including quality assessment and quality steering, was adopted and developed in Swedish education policy.

At almost the same time, far-reaching school choice and marketization reforms were introduced, reforms in which the quality discourse was central as well. Bunar (2010: 11) claims that "improving the overall educational quality in the country, students' achievement, and lowering the costs have always been the strongest arguments made by the proponents of school choice policy". This is also evident in the state policy documents connected to these reform decisions, for example in the Government Bill proposing favourable conditions for tax-funded free schools (versus in most other countries): "stimulating competition ... can contribute to higher quality and productivity in the school system" (Government Bill, 1992/93:230: 27, *author's translation*). School choice presupposes well-informed customers, which in the school sector implies that pupils and their parents need to base their school choices on comprehensive information. This view is embraced by the present government, for example as exemplified by new directives to the NAE to develop indicators and channels for accessibly communicating school quality information, to help students and parents choose school. "In a free school-choice system it is necessary that students and parents be given the conditions to make as well-grounded decisions as possible. To do that, solid decision support is needed" (Swedish Government, 2012, *author's translation*).

Bergh (2011: 712) demonstrates that quality has been pursued by successive Swedish governments between 1990 and 2010 and calls the government's 1996 development plan (Skr 1996/97:112), the "official starting point for quality thinking in Swedish education". Bergh (2011: 714) describes the growing rhetorical emphasis on the connection between education quality and economic growth in contrast to the previous aim of providing a good general education for all:

> …speech acts concerning the relationship between quality and *education* increasingly come to focus on goal achievement and emphasise Sweden's future role as a leading knowledge nation.

Bergh's analysis indicates that educational quality (i.e., traditional educational values such as democracy, knowledge and Bildung) was a dominant criterion in the rhetoric until the beginning of the 1990s. However, in that decade, the traditional understanding of quality was challenged by results-based quality, market quality, and system quality – concepts that imply a shift in the meaning of quality.

In 1997, the government made a decision about quality assessment in the school system (SFS 1997: 702) that implied that all schools and municipalities were obliged to carry out annual quality reports as a part of the ongoing assessment and evaluation work. In 2006, the National Agency for Education was assigned the task of developing general guidelines to advance goal achievement and development; the outcome was the first quality work document, the 2006 "General advice for systematic quality reports in schools" (NAE 2006). A few years later, this was replaced by a new document (NAE 2012a) in response to the introduction of a new Education Act (2010). This new document, "General Advice for systematic quality work in schools", is the focus for the present analysis.

Programme theory analysis

Programme theory analysis is useful for analysing the intentions and internal logic of the systematic quality work as decreed in the document "General Advice" (NAE 2012a). As programme theory underlines the possibility to use interviews and other methods to derive programme theories (Funnell & Rogers 2011: 120), the analysis is also based on interviews with two policy actors involved in formulating the General Advice, representing the National Agency for Education and the Swedish Schools Inspectorate. Research and public reports are used as part of the frame of reference as well, not least for illuminating the potential for unintended outcomes and what is realistically achievable. "A program theory can be improved by including key factors that are likely or are known to affect its success" (Funnell & Rogers 2011: 110).

The General Advice is a central document that frames a multitude of evaluation, assessment, and follow-up systems and activities in the Swedish school system. This article treats the document as a social intervention programme, i.e., "an organized, planned and usually ongoing effort designed to ameliorate a social problem or improve social conditions" (Rossi, et al. 2004: 434).[5] The programme to be analysed is set forth in the General Advice, which constitutes an *evaluation system* in the sense that "evaluations are no longer commissioned and conducted on an ad hoc basis, but through more permanent arrangements, which aim to guarantee, in advance, the supply of evaluative information" (Leeuw & Furubo 2008: 159). Furthermore, systems "are characterized in terms of organizational capacity, sustainability, money, power and interactions with clients, stakeholders and user systems" (Leeuw & Furubo 2008: 14).

Funnell and Rogers (2011: 31) define *programme theory* as "an explicit theory or model of how an intervention contributes to a set of specific outcomes through a series of intermediate results". This is in contrast to a "black box eval-

uation", in which input and output are described but not the processes occurring between them, and programme theory "aims to help policymakers and practitioners 'open up the box' of successful programmes to understand how it works rather than having to buy the whole package and plug in". Programme theory is also useful for adapting interventions to new situations as it illuminates the causal mechanisms by which they are supposed to work, making it possible to "determine whether this is different for different people and in different implementation context" (Funnell and Rogers 2011: 9). This is a relevant idea in the Swedish school context, which is characterized by the extensive devolution of school policy realization to school-principals, teachers' teams, and individual professionals in schools. This devolution presupposes substantial reflection and development at the local level, where various professions are expected to exercise their discretion in order to fulfil state policy goals based on professional judgement. In this article, I regard programme theory analysis as a tool for illuminating the intentions of an evaluation system and how it is intended to work at a programmatic level. Furthermore, in line with Scott's (2003) description of organizations as open, natural systems and Dahler-Larsen's (2012b) understanding of evaluation systems as social and political phenomenons, I indicate the importance of the wider environment in which organizations are embedded by putting the programme into a wider context in the introduction, background and the analysis sections.

The study is part of a research project, "Consequences of evaluation for school practice", funded by the Swedish Research Council, in which a programme theory analysis approach has been developed (see Hanberger 2014). The approach is based on Funnell and Roger's (2011) definition of programme theory and on well-established assumptions and models of programme theory analysis, including external validity, as described by Leeuw (2003), Dahler-Larsen (2012b), and Funnell and Rogers (2011). The approach will be used for the analysis, including to answer the following questions:

Step 1: Reconstructing the programme -
General Advice's programme theory (PT):
1. What are the assumptions of General Advice?
2. What problem(s) was General Advice intended to resolve?
3. What are the pre-requisites for General Advice, and what activities are assumed to produce what effects?
4. Are the problems that General Advice is to manage described and substantiated?
a. Whose knowledge needs does General Advice meet?
b. Are arguments provided as to why General Advice is needed?
5. Are the intended effects clearly described/specified?
a. Are the intended short- and long-term effects defined?
b. For whom and where are the effects expected to occur?

Step 2: Assessing the PT's internal validity

6. Is General Advice's PT consistent?

a. Is there a logical/coherent description of how the programme activities are to achieve the intended effects?

b. Are there activities for all intended effects?

c. Are there activities that do not logically match their intended effects?

Step 3: Assessing the PT's external validity

7. Does General Advice have scientific support?

a. Does the programme reflect/do justice to the objects and activities it is intended to measure?

b. Is General Advice's PT consistent with existing knowledge of conditions/factors creating improved teaching/ learning/ education (system)?

c. Is the PT consistent with governance and accountability theories?

d. Does General Advice provide information about data sources, data quality, non-response, who reports information to the system, etc.?

8. Is the knowledge that General Advice produces useful in helping resolve the problems that General Advice was set up to manage?

A programme theory of systematic quality work in schools

Reconstructing General Advice's programme theory

The fundamental assumption of the programme (Q1[6]) is that systematic quality work in Swedish schools will help "achieve the national goals for education" (NAE 2012a: 45)[7]. The Education Act stipulates that the quality work should be systematic and continuous. The assumption is that continuous assessments and evaluations will identify areas needing improvement, making goal achievement possible. The employees are assumed to participate and be competent, motivated, and engaged in the work; students and parents are also expected to take part. Enhanced participation is an important goal of systematic quality work (NAE 2013a).

Management by objectives and results (and the assumption that such management effectively promotes goal achievement if applied in line with the programmatic ideas) is a prerequisite underlying the whole system, and within that framework the quality concept is crucial. The General Advice connects quality to management by objectives and results in the following definition of *quality:*

a generic term for how well the schools:
- fulfils the national goals,
- meets national demands and guidelines,
- fulfil other goals, requirements and guidelines that are consistent with the national goals, and
- are characterized by striving for renewal and continuous improvement based on prevailing conditions. (NAE 2012a: 45)

29

The problems to be resolved by systematic quality work are not explicitly expressed (Q2), though, they are implied in the recurrent formulations about goal achievement, which leads the reader to think that current goal achievement is unsatisfactory. In this respect, the quality work constitutes "a risk reduction practice", similar to how Power (1999: 5) speaks of auditing in general. In our case, the quality work is a means for the state to reduce the risk of a school system producing unsatisfactory results. An alternative interpretation is that the goal achievement is already satisfactory, but that the intention is to improve even more. The same goes for notions such as development, quality, and equity: the programme does not explicitly state that these are unsatisfactory but it appears to be assumed – but could instead aim to improve something that is already good.

One official (Interview 1) says that the state has been dissatisfied with the varying standard of the local quality reporting. For example, overarching educational goals like values and norms are often ignored. The Inspectorate's concern is not primarily about students' declining subject area knowledge though it is linked to the overarching goals. The Schools Inspectorate (2013) describes the problems: "The schools' reports and analyses often concern just a few national subject tests. Beyond that, the analytical work at the schools is insignificant" (*author's translation*). Another official (Interview 2) points out that the analyses are often confined to the student level, while the overarching level is missing, and that too many school principals are poorly informed about the daily work of the school.

The following overall reconstruction of a programme theory for the General Advice (NAE 2012a) illustrates what prerequisites and activities are assumed to cause what effects, that is, the internal logic of the programme (Q3, 5a and 6).

If the following conditions are met:
The school owners (i.e., municipalities and independent providers) and principals, in cooperation with school staff
- create routines for the quality work
- ensure that management, organization and assessment support the quality work
- ensure that staff are competent to conduct assessment and analysis
- ensure student and parent participation in the quality work

And if the following activities are carried out:
The school owners and principals
- ensure that the documentation constitutes a sufficient basis for the analyses and decisions needed.
- create documentation routines and forms that are efficient and appropriate for the quality work
- strive to create an overall picture of education quality at all levels
- compile results indicating how the prerequisites and realization of the education influence goal achievement

- ensure that evaluations concerning specifically identified areas are carried out, in addition to the ongoing assessments
- in cooperation with the staff, analyse what influences and produces results and goal achievement, on the basis of the assessments.
- analyse whether the causes appear clear or whether further assessments and evaluations are required
- use the analysis as a basis for dialogue on the need for development
- together with the staff, identify areas for development and then decide what efforts and measures are needed in order to fulfil the national goals

The following short- and long-terms effects should occur:
- the educational goals in the Education Act, national curriculum, and other steering documents are fulfilled
- student learning and results are improved
- other school goals are achieved, such as study and vocational guidance and student welfare/health
- an equal education of high quality is achieved
- the education is continuously developing

In this programme theory analysis, we searched for phrases and passages in the 52-page document that distinctly stipulate conditions, activities, and outcomes, and these phrases and passages were interpreted in the light of the model above. Such expressions were often found in sections of text introduced by quotations from the Education Act. Long passages of reasoning as well as comments were used to a lesser extent and sections with little relevance to our purpose (e.g., a passage about complaint routines) were ignored. The overview indicates that the document was at least partly formulated based on programme theory, which was confirmed by an official (Interview 1).

The programme theory reconstruction shows that the school owners and principals are the key actors in the quality work. They are responsible for crucial decisions, for example regarding "what material to collect, how to do it, when and in what purpose" (NAE 2012a: 28) (Q 4a, 7d). However, the document strongly emphasizes everybody's participation in the systematic quality work, in order to make quality apparent and to contribute to improved goal attainment. The other actors, i.e., the school staff, students and their parents, are given "opportunities to participate, make choices, and influence the improvement of the education" (NAE 2012a: 23). The students' right to participate is emphasized, while it is said that the parents should take part to a lesser degree. The school staff are expected to work actively in all phases of the quality work: follow-up, analysis, planning, and realization. Documentation is central in all phases.

The programme theory analysis demonstrates that fully implemented systematic quality work demands substantial resources, competence, and coordination (Q 3). Ball et al. (2012: 19) claim that "the fact that policies are intimately

31

shaped and influenced by school-specific factors which act as constraints, pressures and enablers of policy enactments tends to be neglected". This issue is recognized in the General Advice programme, but is not described as problematic, probably due to the division of responsibility for school management. The state seems to assume that the municipalities and the private providers will take responsibility. This is in line with the governance system but not with evidence of how various school owners run schools. The decentralization and deregulation reforms of the early 1990s have led to great variations in local resource allocation, priorities, and strategies (SOU 2014). Performance differences have increased at all levels - between municipalities, schools, classes, and individual pupils - since the end of the 1990s (NAE 2012b). Moreover, from a legal governance perspective, the NAE (2011b: 58) is not satisfied with how municipalities have taken responsibility: "The devolution from the state to the municipalities of creating conditions and estimating economic resources needed in order to achieve the national goals of the school system, cannot be said to have been handled in an appropriate and equal way" (*author's translation*). This knowledge is a relevant motive for enhancing the systematic quality work, but we cannot take it for granted that the school owners will become more responsible.

The quality work is primarily intended to satisfy the knowledge needs of school owners and staff, as a basis for their work to improve goal achievement (Q4a and 5). However, the General Advice does not explicitly express the knowledge needs of the state. It does not specify whether the new and upgraded quality work will be assessed by the state (e.g., by the Schools Inspectorate) and, if so, what material will be collected. It is conceivable that the quality work is expected to improve pupil performance, which will be detected in the ordinary statistics (e.g., regarding grades, test results, and student throughput) and in international knowledge comparisons, such as PISA and TIMMS. This is confirmed by interviewee 1, who also stresses that it is the Schools Inspectorate's task to scrutinize the local quality work.

Assessment of programme theory validity

The programme theory is largely consistent (Q6), provided the basic assumptions of management by objectives and results (MBOR) work in practice. Laegrid et al. (2006: 269) conclude that MBOR may be useful if it is "adjusted to the complex political and administrative context of central agencies". However, they claim that MBOR is not based on a consistent theory (Q7c). Instead, it is derived from both economic organization theories and management theories and "prescribes both centralization and decentralization" (Laegrid et al. 2006: 252). Considerable autonomy is allowed but "the price public bodies have to pay for their freedom is to accept a more rigid performance-management system".

The fact that several actors at various levels are responsible and have the right to influence the realization of the quality work is a demanding and possibly contradictory challenge (Q 7a and 7c). The strong authority of the national steering document is clear. At the same time, decisions made at the owners' level as

well as staff, student, and parent influence and participation are emphasized. Moreover, as schools are expected to formulate various specific improvement needs at school level, there is also a potential tension between these and the owners' priorities. This is noted in the General Advice, which warns of the risk that school principals become mere administrators instead of "drivers" of the development work (NAE 2012a: 34).

It is a major task to balance the many goals of the national steering documents against the owners' goals, the 'students' right to influence and responsibility' (NAE 2012a: 18) (Q7c), and the "participation of all" in the quality work which should be "ever-present at all levels: individually, in work teams, in the unit as a whole, and at the owners' level" (NAE 2012a: 11). The multifaceted nature of the task is underlined by the fact that other steering documents must also be considered, for example, United Nations Agreements on Human Rights and the NAE's General Advice on several other areas than quality work, for example, planning and grading. In addition, the Schools Inspectorate (2013) develops Advice and guidance on their own (Interview 2).

The previously mentioned assumption that all employees should participate in the quality work, being competent, motivated, and engaged, cannot be taken for granted. There is substantial research into the increasing work-load in schools (Hargreaves & Fink 2006; Lärarförbundet, 2012; Swedish Work Environment Authority 2012), as a result of the last few decades' "policy epidemic" (Levin 1998: 137).

The main objective of the programme clearly reflects and does justice to the objects and activities it is intended to measure (Q7a). However, the concept of goal achievement is problematic in a system with a substantial number of goals at various organizational levels. A rough estimate of the number of goals in the national curriculum is approximately 950. They are formulated at three levels: fundamental values and tasks of the school (approximately 20 goals); overall goals and guidelines for education (approximately 30 goals); and syllabuses, which are supplemented by knowledge requirements (approximately 900 goals). In addition, there are goals at the municipal/private provider and school levels as well. The extension and importance of these vary and they are not examined here. In all, this rough description tells us that goal achievement is a contested concept in a context such as the Swedish school system[8].

The wide-ranging and intangible nature of the quality concept complicates the analysis of the programme's consistency (Q7a). Dahler-Larsen (2008) describes five perspectives on quality, making it possible to talk about quality as a principle while not taking definitions for granted. These perspectives are based on various problem areas and quality criteria: reducing variations around a defined standard (for an acceptable quality level); obtaining certain effects (i.e., improvements in people's lives and in society at large); achieving declared political goals; satisfying the preferences of users, be they customers, clients or users; and ensuring quality by means of the organizational system, i.e., quality assurance should be built into the organizational structure. The General Advice programme embraces all these quality perspectives but focuses mainly on achieving

declared political goals and ensuring quality by means of the organizational system. However, the other perspectives are not absent.

The General Advice gives a very ambitious description of the phases of quality work, but its formulations are mostly fairly general and school staff are expected to realize the underlying intentions in practice, which is in line with the principle of decentralized governance. However, practice is often characterized by work overload, as mentioned, implying the risk that the staff will try to find shortcuts in order to meet the demands. This may result in achievement being captured only by easily measurable aspects of education (Interview 2), which will devalue educational quality in a deeper sense. The purposes of schooling from a broader perspective may disappear, as these broader goals are difficult to quantify without reducing their meanings (Q7a). This is an example of a "reductionist view", which Dahler-Larsen (2012b: 37) claims is a risk in complex contexts. Some of the broader goals are captured in the following:

- Fundamental values (e.g.,democracy, equality, human rights, and the environment…)
- A lifelong desire to learn
- Live with and appreciate the values inherent in cultural diversity
- Preparation for active participation in society
- Ability to take personal responsibility
- All-round personal development of students into active, creative, competent and responsible individuals and citizens.
- A cultural heritage – values, traditions, language, and knowledge
- Basic knowledge that constitutes the common frame of reference of everyone in society
- Ability to keep one's bearings in a complex society
- Ability to critically examine facts and appreciate the consequences of different alternatives
- Ability to communicate
- Creativity, curiosity and self-confidence
- Desire to explore one's own ideas and solve problems
- Perspectives: ethical, historical, international, and environmental
- Self-development and personal growth
- Learn, research, and work independently and together with others
- Make use of critical thinking

(Key-words compiled from the national curriculum for compulsory school; NAE 2011a)

The risk of reductionism is accentuated by the striving for "precise, concrete, specific and hierarchically structured indicators" Laegrid et al. (2006, 251) inherent in MBOR. The broad goals can, to some extent, be captured in tests and grades, but they tend to disappear or recede into the background due to the ever-present pressure to measure, mark, exhibit, and rank performance in what are

assumed to be reliable and objective ways (Hargreaves & Fink, 2006; Power & Frandji 2010). For example, two of the most commonly used and respected evaluation systems for Swedish schools, SIRIS (from the NAE) and Open Comparisons (from the Swedish Association of Local Authorities and Regions), publish quantitative, recurrent data on, for example, grades, test results, and resources. The abovementioned broad goals are not measured by these indicators, except for what is indirectly captured in tests and grades, which implies a devaluation of the education (Q7a and c). Levin's (1998: 137) statement on testing gives a hint of this dilemma:

> We know quite clearly how difficult it is to design testing that addresses the broad range of school objectives, is culturally fair, does not distort teaching practices and provides information that is actually helpful both to educators and policy makers.

To what extent programme theory is consistent with existing knowledge of conditions and factors known to create the intended effects (Q7b) depends on the theoretical perspectives used. On the basis of certain management- and implementation perspectives, it is consistent. According to perspectives that emphasize social constructive, bottom-up, governance or street-level contexts (Ball et al. 2012; Colebatch 2009; Lipsky 1980), the programme theory instead appears to be a vision. The programme, is "a normative projection of the hopes invested in the practice" (Power 1999: 4) and the specific conditions (including time and competence) created in practice are important determinants of the extent to which hopes will be fulfilled. In cases in which the work-load is actually a problem or the competence is insufficient, there is a risk that the actual outcome of the quality work will be disappointing. This risk is anticipated by the General Advice as it states that staff need adequate conditions, time, and competence in order to fulfil the demands of the quality work. For example, it is said to be important that "the staff have both the competence to follow up the teaching processes and the time and support they need for shared reflection on what works more or less well in education" (NAE 2012a: 17).

However, as several reforms have been implemented over the last few years without supplying staff with additional resources, only time will tell what conditions have actually been created in order to realize the intentions of systematic quality work. The same goes for the perception that successful organizational change requires appropriate infrastructure and opportunities for staff to make the deeper meanings of the changes their own (Fullan 2001). As the 290 municipalities and the private providers are in charge of organizing and operating school services, such resource allocation will depend on local priorities.

There needs to be awareness of constitutive effects, that is, "the ways in which evaluation systems shape behaviour and redefine the meaning of public activities because evaluation indicators become goals in themselves" (Dahler-Larsen 2012b: 37). One example is the abovementioned risk of reductionism. Another one is if the results of the quality system are linked to the MBOR idea to reward good and punish bad performance (Laegrid et al. 2006: 251). An

OECD-report (Nusche et al. 2011) recommends a system for teacher appraisal as a part of the evaluation framework. This brings us close to linking the substantial statistics on grades, test results, and other performance indicators to individual performance-related pay. However, as this pay system seems to fail to meet its stated aims (Lundström 2012), a stronger link between assessments and pay would probably be counterproductive.

The grade inflation that has occurred since the 1990s (NAE, 2012c; Vlachos 2011) can partly be explained as an effect of the natural need for employees to be rewarded and to avoid punishments. Another possible explanation is the increasing importance of displaying apparent good performance in school rankings in a competitive school market. Losing this competition may result in teacher redundancy or school closure, which creates incentives to produce the appearance of what is assumed to be good quality. Such constitutive effects of the evaluation system can result in what Ball (2003: 222) calls "game-playing" or "fabrication" and what Power (1999: 94) calls "creative accounting" or "fiddling", which he describes as part of the mutual construction of information systems and forms of behaviour.

Vedung (2010: 263) says that evaluation as a governance formula has "completely exploded", since around 1990, and that the message is simple: "Good intentions, increased funding and exciting visions are not enough; it is real results that count. The public sector must deliver. It must produce value for money". The motives underpinning the programme are consistent with such a simple definition of democratic accountability (Q7c).

The programme theory is also consistent with other steering documents, for example the two most important: the Education Act and the national curriculum. The division of responsibility that characterizes the governance of the Swedish school system is crucial for understanding the design of quality assessments. The Swedish NAE (2013b) describes the governance:

> The Swedish Parliament and the Government set out the goals and guidelines for the preschool and school through i.a. the Education Act and the Curricula. The mission of the Agency is to actively work for the attainment of the goals. The municipalities and the independent schools are the principal organisers in the school system, allocate resources and organise activities so that pupils attain the national goals.

Furthermore, the wide autonomy granted school principals and teachers is connected to quality and is reflected in the national curriculum:

> Both the daily pedagogical leadership of the school, as well as the professional responsibility of the teachers are necessary conditions for the qualitative development of the school. This necessitates continuous review, following up and evaluating results, as well as assessing and developing new methods. Such work has to be carried out in active co-operation between school staff and pupils, and in

close contact with the home and the local community. (NAE 2011a: 13)

As mentioned, the programme theory is based on, and consistent with, MBOR, but we cannot take for granted that it is ideal for all aspects of education (Q7c). The components precise objectives and rewards and punishments have been mentioned. Moreover, the emphasis on management (Drucker, 1954; Gill & Whittle 1992) does not conveniently agree with the governance and leadership culture of Swedish schools, the emphasis on collaboration and participation in the professional educational culture, and to the fundamental values of the National curriculum, which acknowledge human growth and inclusion. Teachers are allowed substantial discretion and many issues are devolved to the teachers' teams. From a social constructive, governance or bottom up perspective, it would be necessary to examine the practitioners' interpretations and translations of the goals in practice, professional values, and, moreover, working conditions.

Systematic quality work is described as an ongoing, cyclic process comprising various phases: analysis, planning, realization, and follow-up – that are repeated iteratively. The starting point of all phases is goal achievement and each phase requires analysis and documentation. The overall intention is that the continuous assessments should produce information about work–organizational conditions, realization, and results.

The General Advice provides an overall description of the division of labour, what is to be analysed and how this should be done – at the owner and school levels. The main focus of the follow-up should be on the students' knowledge results. The owner should have ongoing dialogue with the principal concerning needs for improvement and these, together with decisions on priorities, should be compiled into an annual situation assessment. The owner is also expected to summarize views from pupils and parents. The statistics should be compiled "in an accessible way" (NAE 2012a: 28).

The General Advice points out that the quality work should be based on science and proven experience (Q7d) and also on the staff's competence, supported by various statistics and other material: "The NAE's statistics, follow-ups, evaluations, and support material, the School Inspectorate's reports and quality assessments and various research studies. Judgements from The Board of Appeal for Education can also be used as support" (NAE 2012a: 33). Both management and staff should use self-evaluations, a method described as the "critical scrutiny of one's own work processes, which leads to discoveries of what each one needs to do in order to improve goal achievement" (NAE 2012a: 18).

Realizing the systematic quality work in practice is delegated to the school owners, school staff, students and parents, which makes it difficult to answer the question regarding whether the quality work produces relevant knowledge (Q8). The General Advice is a programme that fosters improvement potential and mirrors the distribution of responsibilities characterizing Swedish school governance. As municipalities and private providers are in charge of organizing and

operating school services within the framework of national regulations and goals, it is decisive for the results how they, in interaction with their varying contexts, interpret and translate systematic quality work.

Conclusions

The analysis of the General Advice programme indicates an ambitious, all-embracing, and ever-present evaluation system in an extensively decentralized goal- and result-oriented school system. The analysis demonstrates that there are inherent potential contradictions, mainly concerning MBOR as well as quality assessment in a highly decentralized system, in what seems to be a consistent programme. This identifies governance problems that may be one of several clues to problems in the school system, such as declining PISA results.

Furthermore, the analysis mirrors and problematizes "governing by the use of new techniques to 'steer and guide' rather than command" (Colebatch 2009: 61) – a key aspect of the governance concept. The programme appears to be a relevant response to the OECD's call to develop a coherent framework for all those working on evaluation and assessment in education (Nusche et al. 2011). However, the analysis identifies a number of prerequisites that must be put in place and several dilemmas that must be resolved wisely if the quality work is to contribute to improved goal achievement in line with the programme aims. The risk of theoretical failure is mostly connected to taking for granted the basic assumptions of MBOR and the implicit or explicit definitions of quality, while the risk of implementation failure mainly concerns the difficulty of realizing a comprehensive and all-embracing programme in a complex context character-ized by decentralization, high reform intensity, and work overload.

It is a challenge to balance the large number of goals contained in the na-tional steering documents (approximately 950) against the owners' goals, "stu-dents' right to influence and responsibility" (NAE 2012a: 18), and "participation of all", including parents, in the quality work, which should be "ever-present at all levels: individually, in work teams, in the unit as a whole, and at the owners' level" (NAE 2012a: 11). There is a risk of reductionism in such a complex situa-tion. Broad educational goals, such as personal fulfilment, citizenship, and social inclusion/justice – which constitute the soul of education – tend to disappear under such pressures. The goals of education may be reduced to narrow measur-able outcomes, which is in line with Biesta's (2009: 35) question of "whether we are indeed measuring what we value, or whether we are just measuring what we can easily measure and thus end up valuing what we (can) measure". This risk is accentuated by the striving for precise, concrete, specific, and hierarchically structured indicators inherent in MBOR and, furthermore, if the staff expected to realize the quality work are stressed by work overload, which is often the case in schools. In such a case, education may be devalued and the teachers downgraded to managed service-oriented workers in a performance context (Ball 2003).

Other risks are connected with the constitutive effects of MBOR, such as the idea of rewarding good and punishing bad performance. This may result in the

"fabrication" of desirable results and, furthermore, clashes with teachers' professional culture in which teamwork and cooperation are crucial. The abovementioned risks imply a shift from the welfare liberal education policy perspective characterizing the national curriculum to a neoliberal perspective (Olssen et al. 2004).

The far-reaching decentralization and devolution of responsibility to school owners and schools constitute another dilemma. The systematic quality programme assumes that existing resources, competence, and coordination are sufficient to implement the programme. However, as there is plenty of evidence of other policies lacking sufficient prerequisites, it cannot be assumed that there are sufficient resources for the demanding task of fulfilling the aims of this systematic quality work programme.

Michael Power opens his influential book *The Audit Society* (1999) with two diametrically opposed scenarios of the presence of checking in society: one in which everything and everyone is checked, and one without any checking at all. Power concludes that neither of these scenarios is imaginable: "What we need to decide, as individuals, organizations, and societies, is how to combine checking and trusting" (Power 1999: 2). The General Advice programme is an example of an ambitious attempt to deal with this tension. The present analysis exposes tensions between the apparently logical programme theory and inherent potential contradictions and critical aspects in the work context. How the programme will be interpreted and enacted in practice in interaction with various school contexts is a question for further research.

References

Arnesen, Anne-Lise & Lundahl, Lisbeth (2006) Still Social and Democratic? Inclusive Education Policies in the Nordic Welfare States, *Scandinavian Journal of Educational Research*, 50 (3): 285-300

Apple, Michael W. (2004) Creating Difference: Neo-Liberalism, Neo-Conservatism and the Politics of Educational Reform, *Educational Policy,* 18(1):12-4

Ball, Stephen J. (2003) The teacher's soul and the terrors of performativity, *Journal of Education Policy,* 18 (2): 215-228

Ball, Stephen J., Maguire, Meg & Brown, Anette (2012) *How schools do policy. Policy enactment in schools.* London & New York: Routledge

Bergh, Andreas (2011) Why quality in education – and what quality? A linguistic analysis of the concept of quality in Swedish government texts, *Education Inquiry,* 2 (4): 709 - 723

Bergman, Bo & Klefsjö, Bengt (2003) *Quality from customer needs to customer satisfaction.* Lund: Studentlitteratur.

Biesta, Giert (2009) Good education in an age of measurement: on the need to reconnect with the question of purpose in education, *Educational Assessment, Evaluation and Accountability,* 21, 33 – 46

Brousselle, Astrid & Champagne, Francois (2010) *Program theory evaluation: Logic analysis. Evaluation and Program Planning* 34 (2011): 69-78

Bunar, Nihad (2010) Choosing for quality or inequality: current perspectives on the implementation of school choice policy in Sweden, *Journal of Education Policy,* 25 (1): 1–18

Colebatch, Hal K. (2009) Governance as a conceptual development in the analysis of policy, *Critical Policy Studies* 3 (1): 58-67

Dahler-Larsen, Peter (2001) From Programme Theory to Constructionism. On Tragic, Magic and Competing Programmes. *Evaluation* 7 (3): 331-349

Dahler-Larsen, Peter (2008) *Kvalitetens beskaffenhed (The nature of quality).* Odense: Syddansk universitetsforlag.

Dahler-Larsen, Peter (2012a) The Evaluation Society. Stanford: Stanford Business Books

Dahler-Larsen, Peter (2012b) Evaluation as a situational or universal good? Why evaluability assessment for evaluation systems is a good idea, what it might look in practice, and why it is fashionable, *Scandinavian Journal of Public Administration*, 16(3): 29 - 46

Drucker, Peter F. (1954) *The Practice of Management.* New York: Harper & Row Publishers, Inc.

Education Act (2010) SFS 2010:800. Stockholm: Government Offices of Sweden.

Ehren, M.C.M., Altrichter, H., McNamara, G. & O'Hara, J.O. (2013) Impact of school inspections on improvement of schools – describing assumptions on causal mechanisms in six European countries. *Educational Assessment, Evaluation and Accountability* 25 (1): 3-43

Evetts, Julia (2009) New Professionalism and New Public Management: changes, continuities and consequences, *Comparative Sociology,* (8): 247 - 266.

Fitzpatrick, Jody L., Sanders, James, R., and Worthen, Blaine R. (2004) *Program Evaluation. Alternative Approaches and Practical Guidelines (3rd Ed.)* Boston: Pearson Education Ltd.

Freidson, Eliot (2001) *Professionalism, the Third Logic.* Cambridge: Polity Press.

Fullan, Michael (2001) *The NEW meaning of Educational Change (3rd Ed.).* New York and London: Teachers College Press

Funnell, Sue C. & Rogers, Patricia, J. (2011) *Purposeful Program Theory. Effective Use of Theories of Change and Logic Models.* San Fransisco: Jossey-Bass

Gill, John & Whittle, Sue (1992) Management by panacea: Accounting for transience. *Journal of Management Studies* 30 (2): 281-295

Government Bill (1989/90:41) *Om kommunalt huvumannaskap för lärare, skolledare, biträdande skolledare och syofunktionärer. (On municipal governing of teachers, school-leaders and career guidance practitioners)* Stockholm: Ministry of Education

Government Bill (1992/93:230) *Valfrihet i skolan. (Freedom of choice in school).* Stockholm: Swedish Government

Grek, Sotiria, Lawn, Martin, Lingard, Bob, Varjo, Janne (2009) North by northwest: quality assurance and evaluation processes in European education, *Journal of Education Policy,* 24(2): 121 - 133.

Hanberger, Anders (2014) What PISA Intend to and Can Possibly Achieve: a critical programme theory analysis. *European Educational research Journal 13*(29): 183 - 196

Hargreaves, Andy & Fink, Dean (2006). *Sustainable Leadership.* San Fransisco: Jossey-Bass.

Leeuw, Frans L. (2003) Reconstructing Program Theories: methods available and problems to be solved, *American Journal of Evaluation, 24*(1): 5-20

Leeuw, Frans L. & Furubo, Jan-Eric (2008) Evaluation Systems: What Are They and Why Study Them? *Evaluation,* 14(2):157 - 169

Levin, Benjamin (1998) An Epidemic of Education Policy: (what) can we learn from each other?, *Comparative Education Volume 34*(2): 131 - 141

Lipsky, Michael (1980) *Street-level bureaucracy. Dilemmas of the individual in public services*. Russell Sage Foundation: New York.

Laegrid, Per, Roness, Paul G. & Rubecksen, Kristin (2006) Performance Management in Practice: the Norwegian Way, *Financial Accountability & Management,* 22(3): 251 - 270

Lubienski, Cristopher (2009) 'Do Quasi-markets Foster Innovation in Education?: A Comparative Perspective'. *Education Working Papers*, No. 25, OECD Publishing.

Lundström, Ulf (2012) Teachers' perceptions of individual performance-related pay in practice: A picture of a counterproductive pay system, *Educational Management Administration & Leadership,* 40(3): 376 - 391

Lundström, Ulf & Holm, Ann-Sofie (2011) Market competition in upper secondary education: perceived effects on teachers work, *Policy Futures in Education,* 9(2):193 - 205

Lärarförbundet (2012) *Låt lärare vara lärare. Perspektiv på läraryrket. Arbetsbelastning (Let teachers be teachers. Perspectives on the teaching profession. Workload)* Stockholm: Lärarförbundet

National Agency for Education (2006) Kvalitetsredovisning. *Allmänna råd och kommentarer (General Advice for systematic quality reports in schools)*

National Agency for Education, NAE (2009) *Rustad att möta framtiden? PISA 2009 om 15-åringars läsförståelse och kunskaper i matematik och naturvetenskap. (Equipped to meet the future? PISA)* Rapport 352. Stockholm: Fritzes

National Agency for Education (2011a) *Curriculum for the compulsory school, preschool class and the leisure-time centre 2011.* http://www.skolverket.se/om-skolverket/in_english/publications (accessed 2013-03-14)

National Agency for Education (2011b) *Kommunalt huvudmannaskap i prakti-ken. En kvalitativ studie*. Rapport 362. Stockholm: Fritzes

National Agency for Education (2012a) *Systematiskt kvalitetsarbete – för skol-väsendet. Skolverkets allmänna råd med kommentarer(General Advice for systematic quality work in schools)* Stockholm: Fritzes

National Agency for Education (2012b) *Likvärdig utbildning i svensk grund-skola? En kvantitativ analys av likvärdighet över tid (Equal education in Swedish compulsory school?)* Rapport 374. Stockholm: Fritzes

National Agency for Education (2012c) *Betygsinflation – betygen och den fak-tiska kunskapsutvecklingen. PM. (Grade inflation. Grades and the actual knowledge development)* Dnr 2012:387

National Agency for Education (2013a) *Allmänna råd för kvalitetsarbete. Sam-manfattning på Skolverkets hemsida. (General Advice for systematic quality work. Summary)* http://www.skolverket.se/om-skolverket/publicerat/nyhetsbrev/2.5111/2.6040/2.6582/allmanna-rad-for-kvalitetsarbete-1.185850 (accessed 2013-03-19)

National Agency for Education (2013b) *The National Agency for Education.* http://www.skolverket.se/om-skolverket/in_english. (accessed 2013-05-02)

Nusche, Deborah, Halász, Gábor, Looney, Janet, Santiago, Paulo & Shewbridge, Claire (2011) *OECD Reviews of Evaluation and Assessment in Education. Sweden*

Olssen, Mark, Codd, John & O'Neill, Anne-Marie (2004) Education Policy: Globalization, Citizenship & Democracy. London, Thousand Oaks, New Dehli: SAGE

Ozga Jenny & Jones Robert (2006) Travelling and embedded policy: the case of knowledge transfer. *Journal of Education Policy* 21(1): 1–17.

Power Michael (1999) *The Audit Society: Rituals of Verification.* Oxford: Ox-ford University Press.

Power, Sally & Frandji, Daniel (2010) Education markets, the new politics of recognition and the increasing fatalism towards inequality. *Journal of Edu-cation Policy,* 25(3): 385 – 396

Rasmussen, Jens (2013) *Competence goal-driven education in school and teach-er education.* Keynote lecture. International Conference on Learning and Teaching 2013. Ministry of Education, Malaysia and Taylors University, Kuala Lumpur, Malaysia

Rossi, Peter H., Lipsey, Mark W. & Freeman, Howard E. (2004) *Evaluation. A systematic Approach (7th ed.).* Thousand Oaks, CA: Sage

Rövik, Kjell-Arne (2000) *Moderna organisationer. Trender inom organisations-tänkandet vid millennieskiftet*. Malmö: Liber

SFS 1997:702 *Förordning om kvalitetsredovisning inom skolväsendet m.m. (Regulation on quality reports in the school system)* Stockholm: Ministry of Education

Scott, Richard W. (2003) *Organizations: Rational, Natural, and Open Systems (5th Ed.).* Prentice-Hall: Upper Saddle River, NJ.

Schools Inspectorate (2013) *Kvalitetsarbete och rektors ledarroll (Quality work and the principal's leadership)* http://www.skolinspektionen.se/sv/Tillsyn--granskning/Vanliga-brister/Rektors-ledarskap-och-systematiskt-kvalitetsarbete/ (accessed 2013-04-05)

Skr 1996/97:112. *Utvecklingsplan för förskola, skola och vuxenutbildning – kvalitet och likvärdighet (Development plan for schools – quality and equity)* Stockholm: Ministry of Education

SOU (2014: 5) *Staten får inte abdikera - om kommunaliseringen av den svenska skolan (The state cannot abdicate – regarding municipalisation of the Swedish school).* Betänkande av Utredningen om skolans kommunalisering. Stockholm: Utbildningsdepartementet

Swedish Government (2012) *Uppdrag om skol- och resultatinformation (Commission about school and resultinformation)* U2012/4307/S. Stockholm: Ministry of Education

Swedish Work Environment Authority (2012) *Arbetsmiljön 2011 (Work environment 2011)* Arbetsmiljöstatistik. Rapport 2012:4

Vedung, Evert (2010) Four Waves of Evaluation Diffusion, *Evaluation,* 16(3), 263 - 277

Vlachos, Jonas (2011) Friskolor i förändring. In Laura Hartman (ed.) *Konkurrensens konsekvenser. Vad händer med svensk välfärd (The consequences of competition)* 66 – 111. Stockholm: SNS.

Interviews

Interview 1: Representative for the NAE (2013-03-27)

Interview 2: Representative for Swedish Schools Inspectorate (2013-04-04)

Notes

[1] The international knowledge assessment, the Programme for International Student Assessment (PISA), was launched by the OECD in 2000. "The results of the first PISA survey came as a shock. At that time it was the general opinion (in the public and among politicians) that the quality of the school systems in these countries was very high" (Rasmussen 2013).

[2] In the Swedish case, the problems are at least partly attributable to reforms making the school system more decentralized and market like than most other countries' systems (Lubienski 2009; NAE 2009, 2012b).

[3] Travelling policy denotes the dissemination and adoption of emergent global education policy trends in local contexts.

[4] Internal logic refers to whether the programme "is designed in a way that can logically produce the desired results" (Brousselle & Champagne 2010).

[5] The definition is close to one of several general definitions of program, made by Fitzpatrick, Sanders and Worthen (2004: 54): "an ongoing, planned intervention that seeks to achieve some particular outcome(s), in response to some perceived educational, social, or commercial problem".

[6] Q1 refers to question number 1 in the programme theory approach.

[7] Author's translation. All quotations from the General Advice are translated by the author.

[8] How to count and define goals in the national curriculum is an open question. For example, I define the knowledge requirements for each subject as goals and estimate that there are approximately 50 requirements per subject, which I multiply by the number of subjects (18). However, the point that there is a substantial amount of goals is hardly contested.