

OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes

UPDATE TO THE COUNTRY BACKGROUND REPORT FOR THE NETHERLANDS

May 2013

Jaap Scheerens
University of Twente, the Netherlands

This report was prepared by the Ministry of Education, the Netherlands, as an input to the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes. The document was prepared in response to guidelines the OECD provided to all countries. The opinions expressed are not necessarily those of the OECD or its member countries. Further information about the OECD Review is available at www.oecd.org/edu/evaluationpolicy.

CONTENTS

Introduction

PART I System Evaluation

- 1) Extension of schematic overview evaluation and assessment provisions presented in Chapter 3**
- 2) Further reflection on the utilization of system level evaluations**
- 3) The new program “*School aan Zet*” (Schools taking the Initiative), and implications for evaluation and assessment**
- 4) Tensions between autonomy and quality (in the sense of centralized evaluation provisions) in the Netherlands**

PART II School Evaluation

- 5) Update of information on Windows for Accountability**
- 6) Update on revision supervision frameworks; the issue of excellence in schooling**
- 7) Developments with respect to pilots “Value added”**

PART III Teacher and school leader appraisal

- 8) Update on policies concerning teacher evaluation**
- 9) School leader appraisal**

PART IV Student evaluation

- 10) Pro and cons of using achievement tests for different purposes (the case of the end test primary education, the CITO test)**

INTRODUCTION

This document was prepared as an input to the *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes*. After publication of the Country Background Report prepared for this project in 2012, the Netherlands invited the OECD to conduct an in-depth country review of evaluation and assessment frameworks in the Netherlands. During the preparatory visit to the Netherlands in February 2013, it was decided to provide an update of the report to further support the main review visit of the OECD review team in June 2013. For this purpose a small study was contracted out to the first author of the original country background report, Prof. dr. Jaap Scheerens. Agreement was reached on ten issues on which additional information would be required. In this report these ten issues have been categorized, according to the structure of the original report, into four parts: system evaluation, school evaluation, teacher appraisal and student assessment.

PART 1: SYSTEM EVALUATION

1. Extension of schematic overview evaluation and assessment provisions presented in Chapter 3

Introduction

The schematic overview of system evaluation, school evaluation, teacher appraisal and student assessment that is part of Chapter 3 in the original report has been complemented by an additional set of tables in which, for each type of evaluation, information is added on the type of instruments used, and the periodicity of the administration. The original table from the report is reproduced once more as: Overview A; and the new extension is indicated as Overview B.

Overview A: type of evaluation, short description, formal responsibility and implementation and use.

Chapter 3 System evaluation

<i>Type of evaluation</i>	<i>Short description</i>	<i>Formal responsibility</i>	<i>Implementation and use</i>
Policy & program evaluations	Evaluation of educational policies and programs	Minister of Education, in one occasion Parliament	Strong resistance from the field against early program evaluations. Little evidence on actual use.
PPON	Periodic national assessment primary schools	Central Test Agency, CITO	Relatively low profile.
Cohort studies	Achievement and attainment indicators of cohorts of primary and secondary school students	Joint responsibility of the Ministry of Education and the Foundation for Scientific research, NWO	Question marks with respect to use by education policy planners and schools
Annual report Inspectorate	Comprehensive report on the state of education	The Inspectorate of Education	Relatively high profile for policy use. Modest press coverage
Monitors	Partial effect and evaluation studies contracted out by the Ministry	The Ministry of Education	Extensive information, no clear evidence about synthesis and policy use

	of education, some of them longitudinal (monitors)		
Key data, “trends in beeld”	Comprehensive annual reports containing key data and indicators	The Ministry of Education	Appear to have high potential for policy use, given active dissemination and user friendly formats
International indicators & international assessments	Dutch participation in IEA, EU and OECD studies	The Ministry of Education	Have obtained high profile in public debate on education concerning the quality of education

Chapter 4 School assessment

<i>Type of evaluation</i>	<i>Short description</i>	<i>Formal responsibility</i>	<i>Implementation and use</i>
School self evaluation	Internal quality care by schools	Schools	Hampering implementation, substantial underutilization
School Inspection	Systematic school supervision, using structured formats and check-lists	Inspectorate of Education	No implementation problems, schools have a positive attitude to inspections
Quality cards	User friendly set of key indicators on school functioning to inform general public and parents. Recently reduced to an indication of the inspection regime a school has to follow, which is indicative on good or failing	Inspectorate of Education	Disappointing use by parents for purposes of school choice.

	performance		
Windows for Accountability	Information dossiers on each school, consisting on centrally delivered quantitative indicators and qualitative indicators provided by schools	A new foundation resorting under the Councils for Primary and Secondary Education, as of 2012	No use and impact information available as yet.

Chapter 5 Teacher appraisal

<i>Type of Evaluation</i>	<i>Short description</i>	<i>Formal responsibility</i>	<i>Implementation and use</i>
New initiative Inspectorate to appraise teachers	Inspection of personnel policy of schools and the quality of teaching in a school; classroom observations in a national sample of schools	The Inspectorate of Education	Results are published in the Annual Inspection Report
Within school teacher supervision	Individual teacher appraisal by school leadership and governance	The competent authorities of the school	No systematic information available

Chapter 6 Student assessment

<i>Type evaluatie</i>	<i>Short description</i>	<i>Formal responsibility</i>	<i>Implementation and use</i>
examinations	Formal assessments at the end of	The Ministry of Education, with delegated	Implementation is obligatory. Use and application is straightforward.

	secondary schools for purposes of individual certification	responsibility to the CVE and CITO. Schools, monitored by the Inspectorate are responsible for the internal school examination	
Cito test	The CITO test is a school leaver test at primary school level, used by 85% of schools.	Schools are responsible for taking part. CITO takes care of technical aspects.	The test is used in supporting students' choice of a specific secondary school track. In aggregated form, use for school and system level evaluation.
CITO LVS	A pupil monitoring system for primary schools, all grades and broad coverage of subjects.	Schools are responsible for taking part; i.e. they buy into the system. CITO takes care of technical aspects.	Tests are used for didactic diagnosis and formative student assessment. In addition aggregated data are sometimes used for school self evaluation. Actual use by schools is still far from optimal.

Overview B: type of evaluation, short description, main instruments and periodicity

Chapter 3 System evaluation

<i>Type of evaluation</i>	<i>Short description</i>	<i>Main instruments</i>	<i>Periodicity</i>
Policy & program evaluations	Evaluation of educational policies and programs	(Failed) attempts at quasi-experimental designs in 1970-ies Ex post facto research in evaluations 80-ies, 90-ies and first decade of 2000	1970 Evaluation of Innovation programs 1980-1990 Retroactive evaluations secondary and primary education 2008 Parliamentary Committee Educational Innovation (Dijsselbloem)
PPON	Periodic national assessment primary schools	Achievement tests	Since 1987 Mathematics/Arithmetic: 1999/2000, 2005, 2012, 2013 Dutch language: 2002 and 2005, 2011, 2013 Reading skills: 2007, 2008, 2011 World orientation, diverse subjects; 9 reports between 2003 and 2011. Non cognitive skills, social outcomes, citizenship, 2011, 2012
Cohort studies	Achievement and attainment indicators of cohorts of primary and secondary school students	Achievement tests in mathematics/arithmetic and language; educational attainment data (progress, graduation); questionnaires to obtain school background information	1977 SMVO-cohort 1982 SLVO-cohort VOCL –cohort: 1989, 1993, and 1999/2000 PRIMA –cohort: 1988, 1994/95, 2000/01, 2002/03, 2004/05 2007 COOL- cohort, integration primary and secondary cohorts
Annual report Inspectorate	Comprehensive report on the state of education	Secondary analyses Questionnaires Direct observations by Inspectors	Annually, since 1801
Monitors	Partial effect and evaluation studies contracted out by the Ministry	Varied methodology; mostly based on questionnaires; panel studies; last 5 years attempts at randomized	Permanently, yearly updating

	of education, some of them longitudinal	control studies	
Key data, “trends in beeld”	Comprehensive annual reports containing key data and indicators	Education statistics	Since 1996, on an annual basis
International indicators & international assessments	Dutch participation in IEA, EU and OECD studies	Internationally comparative assessment tests; school and teacher questionnaires	TIMSS since 1995 PISA, since 2000

Chapter 4 School assessment

<i>Type of evaluation</i>	<i>Short description</i>	<i>Main instruments</i>	<i>Periodicity</i>
School self evaluation	Internal quality care by schools	Check-lists, questionnaires; narrative reports, self-assessments	Strong stimulation in the period between 1999 and 2006. Continued press caused by the Inspectorate observing “quality care” as a core quality facet
School Inspection	Systematic school supervision, using structured formats and check-lists	Explicit evaluation frameworks and structured observations, since about 1986	Permanently, according to selection schemes of schools that have changed over time. Currently each school is inspected at least once every four years.
Quality cards	User friendly set of key indicators on school functioning to inform general public and parents. Recently reduced to an indication of the inspection regime a school has to follow,		Quality cards from the Inspectorate came into existence in 1998 for secondary schools and 2003 for primary schools. In 2007 replaced by much reduced cards on the “supervision arrangements” for schools; related to current proportional inspection. In addition so called “attainment cards”, for

	which is indicative on good or failing performance		secondary schools came into being.
Windows for Accountability	Information dossiers on each school, consisting on centrally delivered quantitative indicators and qualitative indicators provided by schools	Administrative data available at the national level for the “central” Indicators. Local indicators are based on information provided by schools, e.g. on student and parent satisfaction.	Windows for secondary schools is operational since 2010. Windows for primary schools started in 2012.

Chapter 5 Teacher appraisal

<i>Type of Evaluation</i>	<i>Short description</i>	<i>Main instruments</i>	<i>Periodicity</i>
New initiative Inspectorate to appraise teachers	Inspection of personnel policy of schools and the quality of teaching in a school; classroom observations in a national sample of schools	Adapted school evaluation framework of the Inspectorate, 2013 revised supervision framework. Structured observation by inspectorates.	Part of the regular schedule for school inspections, including proportional inspection and basic inspection of all schools, every four years
Within school teacher supervision	Individual teacher appraisal by school leadership and governance	Criteria to evaluate teacher competencies by school boards and school leaders. National register of qualified teachers	Since 2006. Teacher register since October 2011.

Chapter 6 Student assessment

<i>Type evaluatie</i>	<i>Short description</i>	<i>Main instruments</i>	<i>Periodicity</i>
examinations	Formal assessments at the end of secondary schools for purposes of individual certification	Central and school examinations consist of multiple choice tests and open questions.	Each year
Cito test	The CITO test is a school leaver test at primary school level, used by 85% of schools.	Standardized achievement test, multiple choice (language, mathematics/arithmetic, study skills)	Each year
CITO LVS	A pupil monitoring system for primary schools, all grades and broad coverage of subjects.	Standardized tests for longitudinal assessment in all main subject matter areas in all grades of the primary school (see table 6.1, country background report.	Test taking as part of the LVS typically occurs twice, during a school year.

2. Further reflection on the utilization of system level evaluations

Introduction

In the original report it was established that there are no empirical studies available that had looked into the use of system level evaluations. Literary the report stated:

“The most that can be said about the probable impact of system evaluation in the Netherlands is that conditions for the information to be used seem to be favorable: there is an abundance of actual and relevant evaluative information; major “carriers” of system level evaluative information are produced close to the main users (i.e. the Ministry of Education) in the form of the Inspectorate’s Annual Report and the annual publications Key Figures and Trends. Monitors and smaller scale evaluation studies are also used to periodically inform the government, i.e. the Minister and Parliament. Cases in point are the various monitors on teacher policies. Finally, current educational policy is increasingly being formulated in terms of measurable targets and standards, which can be seen as another favorable condition to the use and impact of system level evaluation and assessment”.

In this addendum a closer look will be taken at three cases of knowledge transfer from system level evaluation and research information to the national level of policy planning and educational policy, namely: the functioning of the Knowledge Chamber, the way the Committee for Policy Oriented Research Primary Education (BOPO) fulfills a brokerage role between (evaluation) research and policy planning, and the impact of recent results from international assessment studies, (TIMSS and PIRLS 2011), on national educational discourse.

Case descriptions

The Knowledge Chamber

Since 2007, the Dutch public service has so called Knowledge Chambers, operating in the various Ministries. <http://www.nieuwsbank.nl/inp/2008/07/30/H131.htm>

These Knowledge Chambers are meant to address the demand and supply of policy relevant knowledge. Usually the Knowledge Chambers, and this is also the case for the Knowledge Chamber in the Ministry of Education, consist of top officials of the Ministry, chairpersons of major Advisory Councils, the Planning Bureaus, representatives of societal organizations, knowledge institutes, and scientists, who are usually invited on an ad hoc basis.

Writing about the Chamber in the Ministry of Education, Rouw (2011)¹ summarizes the motives for it. “The Knowledge Chamber was established for three reasons:

- 1) As part of our effort to promote evidence based policy, to strengthen the knowledge base of our policies, the Ministry having concluded that it needed to include researchers and experts in policy making to share their views and insights with policy advisors in order to bring in scientific evidence.
- 2) The second reason is the need to formulate an explicit knowledge policy. Without such a policy the government runs several risks:
 - a. First there is the risk of an excess of knowledge and information. As the amount of data and information is constantly rising it is becoming more difficult to pick

¹ Rouw, R., (2011) *Gevoel voor Bewijs. Naar vloeiende verbindingen tussen kennis en beleid*. Den Haag, Atelierreeks, NSoB.

up relevant research, to interpret it correctly and to link it to knowledge already available.

- b. The second risk is compartmentalization in knowledge domains. The compartmentalization between and within departments is reflected in the way the knowledge infrastructure is organized, namely in separated domains. An integral approach is hindered by the compartmentalization of knowledge.
- c. The third risk is that government officials, especially at the top, concentrate on the process of policy making rather than on the content of a certain policy.

It is to minimize these risks that the Ministry has made the Knowledge Chamber into a crucial component of its knowledge policy and formulates knowledge policy at the top, starting from a strategic vision of the role of knowledge in policy.

- 3) The third reason was the restructuring of the system of advisory councils and knowledge institutions (such as planning offices and research institutes) which began in 2006. The results of this initiative were set down in a letter from the cabinet to parliament. In this letter the government stated that the direct interaction between policymakers on the one hand and knowledge institutions and researchers on the other hand should be improved. The letter also states that the form in which this interaction is organized is up to each ministry to decide. A knowledge chamber is strongly recommended but not prescribed. Every ministry must devise an arrangement that suits the conditions within their domain the best.”

The Knowledge Directorate, at the Ministry of Education, which has been mentioned in the original report, played a supportive role with respect to the Knowledge Chamber. Although the Knowledge Chamber had no explicit role in considering policy implications from system evaluations, it might have functioned as a good platform to do so, particularly with respect to bringing information from the many sources of system evaluation together.

Presently the Knowledge Chamber at the Ministry of Education has entered a period of suspension, while the future of of consultation, as well as effective forms thereof, are being considered.

BOPO research reviews

The Committee for Policy Oriented Research Primary Education (BOPO), recently contracted out four review studies on policy oriented research about major priority areas in primary school policy; “adaptive education” (Dutch: *Passend Onderwijs*; literary “Education that Fits”), “policy on disadvantaged students” (*Onderwijsachterstandsbeleid*), “educational governance” and “quality”, more specifically the policy aimed at achievement oriented education (Ledoux and Smeets, 2013,² Meijnen and Mulder, 2013³, Scheerens and Doolaard, 2013⁴ and Frissen, Hofman and Peeters, 2013)⁵ Each of the review studies was based on three to four separate research studies that had been conducted between 2009 and 2012. The results

² Ledoux, G., & Smeets, E., (2013). Programmalijn Passend Onderwijs. Presentatie BOPO studiedag, 5 maart

³ Meijnen, G., W., & Mulder, D., L., (2013) Programmalijn Onderwijsachterstanden. Presentatie BOPO studiedag, 5 maart

⁴ Scheerens, J., & Doolaard, S. Review studie Onderwijskwaliteit PO. Groningen: GION

⁵ Frissen, P.H.A., Hofman, R.H., & Peeters, P. (2013) Het ongemak van Autonomie. Onderwijsbeleid tussen vrijheid en verantwoording

of the review studies were discussed at a conference at the Ministry of Education, on 5 March, 2013.

Procedurally the review studies as well as the final conference at the Ministry of Education can be seen as a strong effort to synthesize research outcomes and discuss policy implications with policy planners and representatives of the relevant councils, in this case, the PO Council and the Education Council.

Substantively each of the review studies came up with evaluative conclusions that “had an edge”, in the sense that they were either supportive of the policy objectives or provided considerable dissonance.

“Education that fits” (adaptive education) is the policy line that followed earlier policy on inclusive education, which was aimed at having a larger share of students with certain difficulties or handicaps in regular primary schools and diminishing the share that goes to special education. The research results showed that, though the participation rates in special education have gone down, schools and teachers are likely to experience problems, when the proportion of students with difficulties increases, and the handicaps are more severe.

The Netherlands has policy on enhancing the position of disadvantaged learners for decades. Basically schools obtain extra funding, if the proportion of socially disadvantaged students is above certain thresholds. More recently pre-school provisions have been created for children from 2 – 4 years. The studies that were analyzed in the review, had particularly looked at pre-school and programs in the lower grades of primary school (4-6 year old students). The results were that no effects could be established, not for the pre-school, and neither for the primary school programs. These outcomes stand in a long tradition of evaluation studies showing no effects of the special and priority programs to enhance equity in Dutch education. The reviewers noted that the programs used in primary schools were not evidence based. Children from minority groups appear to be gradually overcoming their staying behind to Dutch pupils, but minority group children from socio economically disadvantaged homes still do worse than Dutch students with the same socio economic background.

The study on quality showed some modestly positive effects of “achievement orientation”, and noted that there is considerable support for this policy among schools. Schools appeared to still experience a lot of difficulty with the didactic analysis and differentiation facets inherent in the achievement oriented approach. The reviewers noted a tension between school autonomy and effective improvement on two issues: the difficulty that researchers had in realizing their planned randomized control group designs; in fact, due to lack of cooperation from schools, such designs could not be realized. Secondly, the predominant “bottom up” style of school improvement precludes structured external support and intervention, also in areas where schools are obviously in need of didactic and curriculum advice.

The review on “Good governance” in the primary school sector, noted the same tension between school and teacher autonomy on the one hand, and central monitoring and accountability on the other. This review too, observed that the demand of implementing evidence based programs and input is not lived up to in the improvement oriented work in the sector. Finally this review noted an unbalance between the position of the centre, educational organizations representing school boards and school boards on the one hand and the much weaker position of teachers and parents on the other.

Earlier this year the results from the International Reading Literacy Study (PIRLS- 2011 and Trends in International Mathematics and Science Study (TIMSS- 2011) were made public: <http://www.utwente.nl/gw/timssenpirls/nieuws/Resultaten%20TIMSS%20en%20PIRLS/> Generally speaking Dutch students did well on both studies, and the slight decline noted in earlier waves of these studies appeared to have come to a stop. A further qualification of these outcomes is that Dutch students do exceptionally well in the lowest part of the distribution, but less well in the highest achievement segment. This trend had already been noted in earlier dated studies, PIRLS 2001, and PIRLS, 2006 (Vermeer et. al., 2011)⁶. It is quite striking how in the media much attention was given to the Dutch students doing somewhat less well in the high segment, whereas very little was said about the positive results for the students at the low end of the distribution. From a policy point of view the evidence of the small proportion of excellent Dutch students is well aligned to the current emphasis on excellent students and excellent schools. This link is also made explicitly in policy documents. The positive results on the position of lower achieving students in Dutch primary and secondary education, however, do not appear to be related to decades of priority policies to enhance the position of low achieving and socially disadvantaged students. This is the more striking, since practically all internal evaluation studies on the effects of these policies came up with “no effects”; see the previous section on the BOPO review studies.

Conclusions

The original report documents that a broad range of system level evaluation instruments is being applied in Dutch education, and that synthetic reporting is well taken care of, for example in the Annual Inspection Report and the annual reports Key Data and Trends. The information that was provided in the cases that are described in the above adds the following:

- a) There appear to be sufficient fora where policy makers and researchers can meet in discussing research and evaluation results; e.g the Knowledge Chamber and the BOPO reviews and study conferences. The Knowledge Chamber would seem to have the potential to address general themes that cut across educational sub sectors and links themes.
- b) Illustrative substantive outcomes were presented that underline the policy relevance of the research information

Still, as the conclusion was in the original report, it is not possible to provide research based evidence on the use of the many sources of system evaluation in the Netherlands. All one could say is that there is a fairly favorable infrastructure for use at the policy-making level. Use and application of information from system evaluation by schools and teachers would be even harder to document, as schools and teachers would have only indirect access to them.

⁶ van der Steeg, M., Vermeer, N., & Lanser, D. (2011). Niveau onderwijs daalt. Vooral beste leerlingen blijven achter *Nederlandse onderwijsprestaties in perspectief*. The Hague: CPB

3. “Schools have the Initiative

Introduction

During the meeting at the Ministry of Education that was organized in preparation of the OECD review (site visit in June) the program “Schools have the Initiative”, which was also referred to in the original report, was presented. The program is very interesting as a case of educational evaluation:

- first of all, since a thorough process and effect evaluation has been planned, and is being executed; with the intention to use a quasi experimental design, and make use of various existing data streams;
- secondly, the concept of experts who come to the school as “critical friends”, has an evaluation facet to it. Critical friends not just assist and support schools, but also review practice in a friendly and critical way, and provide feedback on the basis of this review;
- thirdly, the way “Schools have the initiative” relates to the intentions of evidence based reform is interesting.

Below a brief description of the program will be presented, and then, the issue of connecting it to evaluation and assessment will be taken up again in the discussion.

Description of “Schools have the Initiative”

“Schools have the Initiative” is based on an agreement between the Ministry of Education and the PO and VO Councils.

In this agreement the partners state that they give space and confidence to school boards to do the right things. Each school can make its own choice in improving its quality. Schools will not be burdened with additional accountability requirements. The responsibility lies with the school boards, which are expected to use their regular channels such as the annual school report.

“Schools have the Initiative” is a three year program (2012 – 2015) for school boards and schools in primary and secondary education, aiming for higher educational outcomes within the framework of national ambitions. “Schools have the Initiative” have translated these national ambitions into program goals in six domains. According to the philosophy of “Schools have the Initiative” schools ought to be the “owners of their change processes”. Therefore the program is called: “Schools have the Initiative”. The six domains of “Schools have the Initiative” are:

- 1) Achievement oriented work
- 2) HRM/learning organization
- 3) Basic skills (language and arithmetic in primary schools, and core subjects English language, Dutch Language and Mathematics in secondary schools)
- 4) Dealing with differences between pupils
- 5) Excellence/gifted students
- 6) Promotion of Beta skills, science and technology

More specifically the program “Schools have the Initiative” intends:

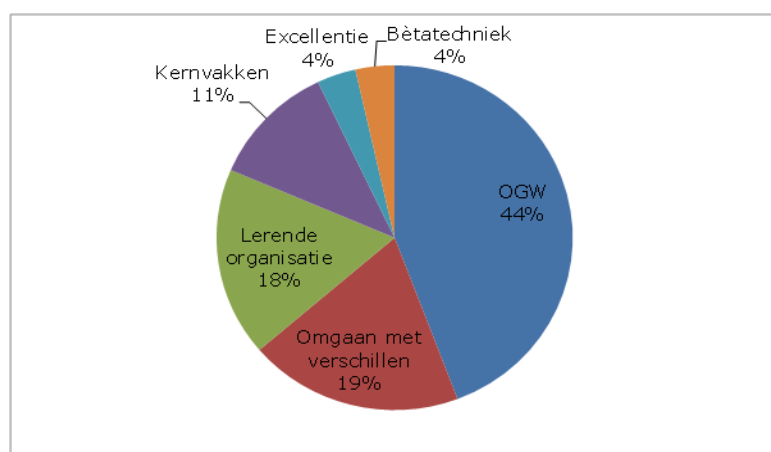
- to stimulate school boards and schools to formulate ambitions and goals

- to provide school boards and schools with relevant knowledge and expertise, and give them insight into good practices, relevant parties and resources and how to get “from A to B”
- keep a sharp eye on the progress that boards and schools make (given their own goals and ambitions), and provide transparency at national level

The actual program is based on schools soliciting a budget from the Ministry to carry out work within one or several of the domains, and attracting external experts, from a national pool of experts and from the existing educational support infrastructure.

“Schools have the Initiative” aims for participation of 3000 primary schools, 450 regular secondary schools and 150 secondary schools offering special education. In November 2012, a total of 2800 schools (both primary and secondary schools, were already involved.⁷

At this time the division of school projects over the six domains, was as indicated in the figure below (same source).



Achievement oriented work (OGW); dealing with differences between students and the Learning Organization, appeared to be the mostly chosen domains.

According to Frissen, Hofland and Smeets (2013) the annual budget for “Schools have the Initiative” is about 150 million per year for elementary schools, for a period of three years. For secondary schools about the same funding is available (bringing the total close to a billion EURO).

A core element in “Schools have the Initiative” is the way schools are supported by external experts. During the three year period each school that takes part is entitled to four meetings with an expert. The experts are indicated as either “ambition experts” or “thematic experts”; some experts are specialized in primary schools, others in secondary schools and or special education. <http://www.schoolaanzet.nl/over-school-aan-zet/de-experts/>

When looking at the profiles of the “ambition experts” (primary schools) it appears that of the 29 experts the large majority has a background of organization and management consultant; 6 experts also have a background in didactics and teaching and learning. For the “thematic experts” (again, primary education) the picture is reversed. Of the 43 experts that are listed, 33

⁷ Source: Informatory note for the Council of Directors at the Ministry of Education, November, 2012.

have a background in subject matter didactics (language, arithmetic, mostly), and teaching and learning; 10 others have more of a background of organization consultant, or hrd manager.

Discussion

“Schools have the Initiative” is a very interesting case, speaking to the way central policy initiatives get to be worked out in a system that is characterized by considerable autonomy of schools. In a subsequent section, the tensions between “quality and autonomy” will be addressed more fully, and in that section reference will again be made to “Schools have the Initiative”.

In the introduction three possible connections of “Schools have the Initiative” and evaluation and assessment were mentioned: a) program evaluation, b) the role of experts as critical friends, and c) interpretations of evidence based educational policy.

Program evaluation

Re a) The Ministry of Education has contracted out a program evaluation of “Schools have the Initiative” that comprises process and effect evaluation.

The central research questions for process evaluation are as follows:

- Which kind of operations have been implemented in “Schools have the Initiative” and which instruments and means have been used for this?
- Is “Schools have the Initiative” being accepted by school boards and schools and how satisfied are the participating schools?
- How satisfied and committed are stakeholders belonging to the existing educational infrastructure?

The core questions for effect evaluation are the following:

- What is the starting position of schools that participate in “Schools have the Initiative” on core indicators (as stated in the agreement between the Ministry and the PO and VO Council), as compared to non participating schools?
- Can differential growth on the core indicators be detected, when comparing program and control schools?
- How did the implementation process in the program schools develop in the period between the pre- and post measures, and how does the program implementation process influence the results of the previous core question, about differential growth between program and control schools.

As implied in the evaluation questions for the effect evaluation a comparison group design was chosen for the evaluation. In the actual evaluation plan a quasi experimental design, where control schools are drawn from schools which enter the program one year later, was given preference over a randomized control group design (as this was considered as practically unfeasible).

The effect evaluation is to be based on quantitative indicators, related to participation objectives (number of participating schools) and realization of performance indicators that vary for the different sub-programs.

Examples of indicators are:

- By 2015 all participating schools in primary and secondary education have identified their (top 20%) excellent students, developed a targeted approach for

these students, and visible performance effects of this approach (sub program Excellence)

- In 2015, fixed percentages (ranging from 30% to 55% for various types of secondary schools) choose a Beta or technical profile (sub program science and technology)
- In 2015 all participating schools score significantly higher on (self selected sub-domains) of basic subjects as measured by standardized achievement tests (sub program achievement oriented work)
- In 2015 all participating schools in secondary education have implemented a learning school organization, and are developing a targeted approach with respect to HRM and achievement oriented work (sub program Learning Organization and HRM)
- In 2015 “practically all” teachers of the participating schools are sufficiently able to provide differentiation in their teaching and deal with individual differences between students (Professionalization in differentiation).

The actual program evaluation is carried out by a research institute. A striking characteristic is that the evaluation approach uses existing data streams to a large extent. Data sources like statistical data from DUO, process indicators from the Inspectorate of Education, achievement results from the COOL cohort studies, and outcomes of various monitors, and even information from the register of teachers, are all intended to be used for measuring the quantitative indicators. Process evaluation partly uses more qualitative data and inventories of “good practices”.

Critical friends

Re b) Little documentation was found on the possible monitoring role of the experts involved in “Schools have the Initiative”. The way the role of the experts is described is as informants, counselors, supporters and advisors of schools, and not as evaluators, not even “friendly critics”. In the way the experts present themselves on the web site, very few of them elaborate on specific monitoring and evaluation experiences and skills.

Evidence based reform

Re c) Initiated by the advice of the Parliamentary Committee on the Innovation of Secondary Education, the Committee “Dijsselbloem” in 2008, educational policy was to become evidence based. New educational policies would need to be piloted, and these pilots had to be rigorously evaluated, preferably by means of randomized field trials. “Schools have the Initiative” was launched at a time when scientific evaluation studies on Achievement Oriented Work, were still in progress, and results not yet available. Moreover, the program is much broader than Achievement Oriented Work, for which at least explicit rationales exist (Visscher and Ehren,⁸ 2010, Scheerens and Doolaard, 2013)⁹. Some authors express doubts about the programmatic

⁸ Visscher, A. J., & Ehren, M. (2011). *De eenvoud en complexiteit van opbrengstgericht werken*. [Enschede]: [Universiteit Twente, Vakgroep Onderwijsorganisatie en -management].

⁹ Scheerens, J., & Doolaard, S. Review studie Onderwijskwaliteit PO. Groningen: GION

activities of schools as part of “Schools have the Initiative” being evidence based (cf Frissen, Hofland & Smeets, 2013)¹⁰.

4. Tensions between autonomy and quality (in the sense of centralized evaluation provisions) in the Netherlands

In the original report some of the tensions between central evaluation and monitoring requirements and school autonomy in the Netherlands have already been highlighted. These tensions manifest themselves most strongly with respect to mandatory testing and the implications of the ambitions with respect to “evidence based” reform. In the Netherlands, a strong intermediary field of educational organizations has always been an important third party, next to the government and the autonomous schools. As indicated in the original report, since about the end of the 1990’s autonomy was further increased, and the interplay between government, educational organizations (particularly those representing school boards) was summarized in a steering philosophy for which the term “governance” was applied. According to this philosophy “the government tries to organize the relationships between students, parents, teachers, school boards, municipalities and the Inspectorate in such a way that they correct one another in keeping a balance and stimulate one another to innovation and entrepreneurship” (Frissen et al., 2013, p 4). The idea of “governance” is that schools not only make themselves accountable to the government, but especially to parents and the public at large (Ibid. p. 4). In actual practice school autonomy is conditioned by the way the government seeks to stimulate the quality of education. This policy does not only specify outcomes but also stimulates certain approaches that are seen as instrumental to realizing improved outcomes, such as mandatory testing and the approach of “achievement oriented work”. Frissen et al (2013) conclude that the ideal of governance, in the sense of actors operating in a balanced structure, through a system of checks and balances, is not realized, because the position of teachers and parents is much weaker than that of school boards, intermediary organizations and the government.

When taking the set of government initiatives to enhance the quality of education as the focus, one could characterize the playing field by observing that central quality norms and preferred work approaches (such as “achievement oriented work”) might be seen as limiting school autonomy, while, at the same time, school autonomy could be seen as constraining the possibilities of the intended “evidence based” way of reform and school improvement. Here the second interpretation will be discussed; the first line of analysis is followed in the report by Frissen et al., 2013, cited earlier.

Evaluation, particularly the intention to evaluate policy effects, both ex ante and ex post, is a contested area in this arena. The actual context of system level evaluation can best be seen as characterized by compromises that involve providing the main actors with different shares of “ownership”, both with respect to policy implementation and evaluation. This will be illustrated by comparing two partly differing and partly overlapping approaches to realizing the Quality Agendas and Action Plans for better performance. The first one is an agenda for effect evaluations of policy interventions, under the heading “Outlook on Effectiveness” (Dutch: *Zicht op Effectiviteit*), the second is the “Governance Agreement”, regarding support

¹⁰ Frissen, P.H.A., Hofman, R.H., & Peeters, P. (2013) Het ongemak van Autonomie. Onderwijsbeleid tussen vrijheid en verantwoording

for the implementation of major facets of the overall quality enhancement policy (such as achievement oriented work). The latter approach is covered in the program “Schools have the Initiative”, discussed in section 3 of this paper.

“Outlook on Effectiveness” lists a total of 50 current policy interventions that have been analyzed for their being amenable to effect evaluation by the Central Planning Bureau¹¹. Of these 50 interventions 34 were considered evaluable. A more in depth study of a more limited number of policy interventions (Van Elk et al., undated)¹² indicates that in actual practice, there are many problems in realizing strong evaluation designs, in quite a few cases, for example the interventions had already started, so that a base line measure was not feasible. In a review of studies that assessed quality enhancement in primary schools Scheerens and Doolaard (2013)¹³ concluded that intended randomized field trials could not be carried out, because of the lack of cooperation from schools. “Outlook of Effectiveness” enforces the idea of evidence based policy, where policy plans are scrutinized for research based support and evaluated with strong research designs. In actual practice this approach seems to experience difficulties, firstly because policy interventions are implemented before science based analyses and baseline measures have had the chance to take place and secondly because autonomous schools are hard to be persuaded to take part in experimental studies.

“Schools have the Initiative” is seen, on the one hand, as the implementation of government policy regarding quality enhancement; but on the other hand as a set of interventions “owned” by schools. The role of the Experts that schools can hire by the special funding they obtain is motivated twofold: as assisting in implementing government policy but also as supporters of school initiatives. Next, they also have a kind of evaluative role, indicated by the concept of “Critical friends” (see section 3).

Generally speaking the combination of “Outlook on Effectiveness” and “Schools have the Initiatives”, can be seen as a typically Dutch construction of giving different stakeholders their share, also known as the “Polder Model”. tradition of bottom up reform, with schools and small networks of schools, having different roles, and complementing one another.

¹¹ Ministerie van Onderwijs en Wetenschappen (2012) Stand van Zaken Zicht op Effectiviteit. Internal report

¹² Van Elk, R., Van der Meer, F., Van der Stee, M. & Webbink, D. (undated). Zicht op Effectiviteit. Studie naar Evaluatie Ontwerpen voor Onderwijs en Wetenschapsmaatregelen

¹³ Scheerens, J., & Doolaard, S. Review studie Onderwijskwaliteit PO. Groningen: GION

PART II: SCHOOL EVALUATION

5. Update on Windows for Accountability

Introduction

Windows for Accountability (*Vensters voor verantwoording*) is a relatively new form of information provision and evaluation on schools, in which central administrative information available from the DUO organization is combined with school level information, to provide a set of central and decentral indicators.

In the original report “Windows” was described as follows:

“*Windows for accountability*” is a project in which all quantitative information, as well as additional qualitative information, about secondary schools is united in one system. The system uses (amongst others) information from DUO, (a central administrative unit of the Ministry of Education), the Inspectorate and the schools themselves. The information is summarized and visualized. “Windows” lead up to information dossiers about individual schools, which schools can use to inform their direct environment and stakeholders. The project is aimed at the development of quality standards that allow, among others, to compare schools with each other (benchmarking). The idea is to connect vertical accountability (towards the Inspectorate of Education) and horizontal accountability, (towards parents, other school sectors – e.g. Windows on secondary school are made available to primary schools- and municipalities). The horizontal element, informing the stakeholders of the school, is predominant (VO-Raad, 2011)¹⁴. The connection with the “vertical” element implies that there is an agreement with the Inspectorate that for schools with well developed “Windows” the administrative burden of supplying information to the Inspectorate will be decreased”.

Windows for secondary schools has been in the field since 2010. The indicators for Windows secondary schools are cited in the original report.

Recent developments

A recent development in the “Windows” – project is the gradual implementation of “Windows for primary schools” (to be indicated in the sequel as Windows PO). Currently pilots on the application of Windows in primary schools are being run. A fully fledged instrument is planned to be ready by 2015. The web-site on Windows PO,

<http://www.vensterspo.nl/over-dit-project/doelstellingen>

describes the aims of the instrument as follows:

“Windows PO is a project in which numerical information about primary schools is being collected in one system, for example, data on educational outcomes, the student population, financial data, and data on the school staff.

Windows PO aims at developing an instrument together with school boards and schools, which has the following characteristics:

- enables a more professional governance of the school organization and ‘achievement oriented work’;

¹⁴ VO Raad: Werkboek Horizontale Verantwoording VO. VO Raad, april 2011.

- enables to make the school accountable to interested parties;
- enables to make information available to parents, among others to assist in the choice of a primary school;
- enables making information available to parents, among others to assist in the choice of a primary school.”

Three different Windows are distinguished:

- a) the Management Window: internal application by the school, related to “achievement oriented work” and offers the opportunity to benchmark (compare with other schools)
- b) the School Window: the web-site of the school, which specifically contains information for parents and other stakeholders (e.g. the municipality)
- c) the School Choice Window: a web-site which contains information to facilitate parents to compare primary schools.

Information on the indicators for Windows PO is provided on the web site:

<http://www.vensterspo.nl/images/Overzicht%20indicatoren%20VSO.pdf>

The indicator set is translated and rendered in the table below.

Central indicators of the horizontal and management window	Decentral indicators for the horizontal and management window
<ol style="list-style-type: none"> 1. General school data 2. Number of students 3. Market share catchment area 4. student flows, intake and school leaving 5. Characteristics of staff 6. Financial key data 7. Financial benefits from sponsors and parent contributions 8. Attainment results 9. School advice on secondary education 10. Transfer to specific types of secondary schools 11. The Inspectorate’s supervision arrangement 12. Score on the Inspectorate’s evaluation framework (quality indicators) 13. The position of students three years after school leaving 	<ol style="list-style-type: none"> 20. Provisions the school is offering 21. (Educational) profile of the school 22. Educational support profile of the school 23. Teaching time and school time 24. Organization and grouping as part of the education process 25. Results indicated by Student Monitoring System and mid time results 26. Pupil satisfaction 27. Parent satisfaction 28. School climate and safety 29. Learning gain** <p>Additional indicators, specifically for Special Education Schools:</p> <ol style="list-style-type: none"> 30. Development perspective 31. Teacher competency* 32. Stages in Special education VSO 33. Use of ICT 34. Cooperation with parents 35. Important chain partners (e.g welfare and care organizations) 36. Quality care* 37. Social outcomes*** 38. Achievement oriented work***
Additional central indicators only for the management window	Additional decentral indicators only for the management window
<ol style="list-style-type: none"> 14. Prognoses of number of students 15. Absence due to illness 	<ol style="list-style-type: none"> 39. Staff satisfaction 40. Cycle of conversations**

16. Financial burdens 17. Financial benefits and costs per students 18. Housing arrangements** 19. Student absence***	41. Expenditure on professionalization and training of staff*** 42. Number of complaints***
--	--

*); new indicators

**): ready after 2015

***): indicators that were dropped

The source of this version of the Windows PO indicator set is Van Gennip, Kuyk and Smeets, 2013¹⁵

A striking point is that the indicator on achievement oriented work has been dropped, whereas monitoring the way the schools takes care of this approach is mentioned as one of the major aims of the Management Window.

Discussion

The development of Windows can be seen as a consequence of the Governance debate in education, in which the organizations that represent school boards have been united in powerful central boards, the PO and VO Councils. These organizations took an active role in the issue of information gathering to support good governance of schools. The information gathering that constitutes Windows is motivated by the call for horizontal supervision, in which the school accepts to be accountable to local stakeholders such as parents and municipalities. Horizontal supervision is contrasted to vertical supervision, which is taken care of through school inspection, by the Inspectorate of Education.

An interesting technical feature of Windows is the combination of central and decentral information, and bringing this information together for each school in a number of information formats, available through the web. Then inclusion of central administrative data from DUO lessens the burden of schools to provide basic information and data themselves.

6. Reflection on inspection methods

Introduction

In the original report proportional, “risk based” school inspection is described in detail. In general terms the current form of risk based inspection puts a strong emphasis on the inspection of low performing schools, and to guiding the improvement of these schools by more intensive inspection. Given the emphasis on excellent performance of students in current educational policy, also supported by findings from international assessment studies, showing that the Netherlands is not doing exceptionally well in the higher segment of the performance distribution, the Inspectorate initiated activities to pay more attention to excellence in schooling. In this section some of the new ambitions of the Inspectorate in the domain of “excellence” will be reviewed.

¹⁵ Van Gennip, H., Kuyk, J. and Smeets, E. (2013) *Indicatoren voor Vensters voor Verantwoording*. Nijmegen: ITS

Intension to further differentiate school inspection

In the “Ambition paper” titled “From Good, via Differentiation to Better” the Inspectorate states that it would like to answer the call for excellence in education by means of differentiation in its approach to school inspection. This means that the focus on the identification and stimulation of weak and very weak schools remains, but that more attention will be given to moderate, average and good schools as well.¹⁶

In addition to weak and very weak schools, moderate, average and good schools are explicitly described, as follows

Moderate (Dutch: Matige) schools, meet the base norms, but improvement is desirable and necessary. Examples of moderate schools are schools where student performance each year is close to the base norm, but are declining or below the standard for specific groups of students. Moderate schools can also be schools that realize insufficient “value added”. Such schools, obtain under the current supervision practice usually a “base arrangement”, but accompanied with agreements and warnings. Part of the schools that are currently judged as “weak”, belong to this new category of moderately performing schools.

Schools of average quality. This is the large middle-range of the distribution of schools that are functioning fair, but are not judged as “good” for various reasons. Their capacity to further improvement, better leadership, self regulation and professionalization is what matters. These schools would need to be challenged to actively improve in these areas.

Good schools. These are schools with above average results, a professional culture of permanent school improvement, where teachers review and coach one another, where there is continuous monitoring of student and school achievement and where ambitions are high. The ambition report indicates that the quality standards (the way the scores on the quality indicators of the Supervision Frameworks are interpreted according to certain norms) need to be differentiated. It also refers to value-added performance measures as being important for the intended further differentiation in school evaluation.

7. Developments with respect to pilot projects “value added”

Introduction

Impulses to explore the possibilities of applying “value added” analyses of educational outcomes came from the Quality Agenda’s and Action Programs on “Enhanced Performance”, and a heightened interest at the Inspectorate of Education in assessing “school excellence”. Tied to the intention of the government to establish mandatory achievement tests in primary and secondary education, baseline information is also considered necessary, so that computation of learning gain and valued added performance indicators becomes feasible. As for “excellence” the increased attention was stimulated by findings from international

¹⁶ Inspectie voor het Onderwijs, 2013a. Van Goed, via Differentiatie, naar Beter. Ambitie-paper van de Inspectie van het Onderwijs, 2013-2017

assessment studies (PISA, TIMSS, PIRLS) indicating that Dutch students do very well in the lower tail of the score distribution, but are about average in the highest score range.

Against this background pilot studies on various aspects of using value added performance indicators were started, first for primary schools, since 2011, and currently also for secondary schools. Some preliminary results of the pilot studies for primary schools will be indicated, and some of the arguments on pros and cons of value added indicators will be discussed.

Some preliminary outcomes of the pilots in primary schools

The pilot projects in primary schools are aimed at two main issues: one more technical and the other more applied. The technical part of the projects compares alternative statistical models to calculate value added indicators. The more practice oriented part of the project is concentrated on practical feasibility, such as the use that can be made of assessment results that schools are already collecting, and the acceptability of applying “value added” and indicators on learning gains for schools and teachers.

The following three models to compute learning gains were considered to be applied in the pilots:

- 1) the “Cukum” model (catching up, keeping up and moving up), which analyzes the progress of three groups of students (low- average and high achievers) and concentrates on the percentage of students that have changed categories;
- 2) learning gain correcting for development during the summer holidays, which is based on the idea that by comparing learning gain during the school year and learning gain during the summer holidays it is possible to control for out-of-school factors, so that an estimate of the “net gain”, solely based on school factors, becomes possible
- 3) Learning gain on subject matter sub-categories within spelling and arithmetic, so that a more precise indication of strong and weak points of students can be obtained.

Next two value-added models will be explored in the pilots:

- 1) contextualized growth-curve models; these models look at contextualized learning gain, which means that additional student background characteristics are incorporated in the computations;
- 2) variance-component and random effect models; these models are also contextualized, but are technically more flexible, because the requirements of vertical equation of tests do not apply.¹⁷

With respect to the practical issues the readiness of schools to follow developments in obtaining learning gain and value-added measures appear to be fair. Next, there are many practical problems with respect to feasibility because of incomplete data. Missing scores, changes in test services, retarded and advanced school careers, school change, and class repetition are frequently encountered.

Discussion

¹⁷ Rekers-Mombarg, L. & Luyten, H., (2012) Het meten van Leerwinst en Toegevoegde Waarde. Groningen, Enschede: GION, Universiteit Twente

The context for applying value added performance measures is “in motion”. The debate about the mandatory school leavers test in primary school indicates that testing and applying tests for accountability purposes is a delicate issue in the Netherlands. Several contexts of application for value added performance measure have changed somewhat. The new government has ended the pilot program on “merit pay” of teachers. With respect to school excellence, the developmental work of the Inspectorate has been overtaken by a Committee established by the Ministry of Education, which has organized a contest for being recognized as an excellent school in the Netherlands. This committee has more or less copied the approach of the German school price and sent jury’s to schools; in their overall judgments output indicators are used, but are given about the same weight as process indicators and self-appraisal by the schools. Moreover, schools enter the competition on a voluntary basis, so that the school price cannot be seen as a form of accountability. Finally, the link of value added performance measures and national program evaluations, such as the one with respect to “Schools have the Initiative” does not seem to have been made. Nevertheless, application of value added measures as part of the intended more differentiated school inspection, remains a relevant context.

PART III: TEACHER AND SCHOOL LEADER APPRAISAL

8. Update on policies concerning teacher evaluation

Introduction

In the original report, teacher appraisal appeared to be the “white spot” in the whole of educational evaluation and assessment provisions in the Netherlands. There is no school-external appraisal of teachers and limited insight in the way “competent authorities”, like school leaders and school boards, take care of their responsibilities in this area. At the same time there is growing recognition of the vital role that teachers have in assuring education quality. After the deliverance of the original country background report a few new developments in the domain of teacher policies and teacher evaluation took place, in the form of a new advice on teachers by the Education Council, and more space for evaluating facets of teacher and personnel policies at schools in the revised, 2013, version of the Supervision Framework for Secondary Education.

Recent developments

In March 2013 a new report of the Education Council, described as an “Exploration” on the subject “Being a teacher” came out¹⁸. The focus of the report is the “personal side of teaching as an occupation”. Teaching is described as a complex endeavor, both internally, when teaching is confronted with lots of dilemmas, and externally, by an ever changing environment. The personal side of teaching and “personal professionalization” is served by teachers being aware of the complexity and unpredictability of their environment, being aware of their own professional values and goals, making “wise personal choices”, using and creating professional space, and continually developing themselves. The report does not deal with external teacher evaluation, but generally puts down external accountability as one of the factors that is likely to limit the professional space of teachers. Nor does the report say very much about internal teacher evaluation. Action-research and design-oriented research are mentioned as possibly serving professional development.

In the context of the missing out of external teacher evaluation in the Netherlands, this report of the Education Council does not offer any new insights or initiatives, if anything it would seem to support the existing status quo.

Since 2011 the Inspectorate of Education is developing adaptations/ extensions of the Supervision Frameworks (*Toezichtskaders*) in the realm of teacher issues at school (Human resources management and professional development). In the autumn of 2012 these developments became more concrete in the “Experimental Evaluation Framework on Teachers”. In this paper teacher quality is described as the most important factor in educational quality. Problematic areas are the image of the teaching profession and the threat of teacher shortages. The extensions in the supervision frameworks are designated in three areas: the

¹⁸ Onderwijsraad (2013) Being a Teacher. The Hague: Education Council

school as the working context of teachers, the actual working process of teachers at schools and “the professional space of teachers”. These three areas are formulated more specifically in a number of questions (check points):

The school as the working context of teachers

Is the school’s personnel policy aimed at professional development of teachers that fits both the school’s educational policy of the school and the development needs of the teachers?

Does the school make good use of the formal instruments and possibilities available (e.g. the “function mix” and grant possibilities for teachers)?

Does school quality care address the educational policy of the school and orientation towards school improvement?

Which organizational measures are relevant for the organization of education, such as working in teams or departments and open learning centers?

Professional space

Do teachers experience “ownership” regarding the school’s educational policy?

Do teachers experience sufficient leeway and discretion to develop education as well as themselves?

Are teachers involved in organizing and improving their teaching, based on the results of the students, and do they make sure that their personal development is matched to these?

Is the school’s or the team’s culture aimed at a joint effort at school improvement and making themselves accountable; what is the role of feedback in this?

Teachers at work

What is the quality of feedback processes and how is this related to the engagement of students with their own learning process?

What are the actual capacities of observed teachers like, and what are their own perceptions of the phase in their professional development and career?

Are the subject matter didactic choices and the improvement priorities of the school or the team recognizable in the observed teaching situations and thus this contribute to the quality of the teaching?

Concerning these last two points attention is to be given to the subject matter oriented skills and subject matter didactics, and not just about general education skills.

On the basis of these additional points of attention the Inspectorate intends to monitor possible pitfalls and sub-optimal functioning of teachers more closely, as part of the risk-based inspection; and also consider the facilitating role of school leadership in this.

In the adapted Supervision Framework for secondary schools, 2013, these points of attention are included in more operational terms. This resulted in two extra “quality aspects”, namely “The school leadership sees to it that teachers can achieve and develop themselves in accordance with the school’s vision” and “The teachers use the professional space that is necessary for good education”. The quality indicators for these two quality aspects are shown in the overview below.

Quality aspect 14: The school leadership sees to it that teachers can achieve and develop themselves in accordance with the school’s vision
--

14.1 The school leadership takes care of the educational vision of the school being translated in concrete professional norms for teachers
--

- 14.2 The school leadership sees to it that there is support for the school vision and ambitions and improvement goals that are deduced from this vision
- 14.3 The school leadership steers towards realizing the educational goals of the school through its personnel policy
- 14.4 The school leadership steers teachers and teams of teachers in a way that they give shape to realizing the ambitions of the school in their own way
- 14.5 The school leadership makes itself internally accountable for the policy aimed at enhancing quality.

Quality aspect 15: The teachers use the professional space for delivering good education

- 15.1 Teachers evaluate the quality of their teaching
- 15.2 Teachers teach in correspondence to the educational vision of the school/department
- 15.3 Teachers work on improving their capacities and competencies to realize the vision of the school
- 15.4 Teams of teachers make themselves accountable for the results that they have obtained on the basis of their professionalization and improvement activities

It should be noted that these indicators are more associated to “secondary” organizational issues of the functioning of teachers at school: professionalization, improvement policies and responding to school leadership. The expectation is that such secondary organizational issues are supportive of educational quality. At the same time, the supervision framework has maintained observational categories more directly related to the primary process of teaching. This is expressed in quality aspect 7, “the didactic acting of teachers”. Indicators for this quality aspect are:

- educational activities are goal oriented
- clear explanation is provided
- students are actively engaged
- the teacher uses adequate didactic approaches
- the teacher stimulates the students to high achievement
- the teacher stimulates the students to apply think- and learning strategies

9. School leader appraisal

Introduction

Evaluation of school leaders was not included in the original report of the Country Background Review. If it had been, it would have become a short section, even shorter than the chapter on teacher appraisal. Formally school leaders are in exactly the same position as teachers, as far as responsibility with respect to evaluation and appraisal is concerned, which means that there is no external appraisal of individual school leaders. New developments too, are quite parallel to what was described with regards to teachers. This means that first steps to new initiatives are present in the realm of School Inspection. Following up on a general tendency in school inspection to explicitly address the level of School Boards, school leaders and school

leadership are now getting more attention in the revised and expanded supervision frameworks.

School leaders as a focus in revised supervision frameworks

In section 6, new quality indicators with respect to teachers' work in schools were cited. School leaders are explicitly mentioned in quality aspect 14; "The school leadership sees to it that teachers can achieve and develop themselves in accordance with the school's vision". In quality aspect 14 school leaders are depicted as stimulators and coordinators of teachers in their work at school; and, moreover, as responsible for quality enhancement of the school (indicator 14.5).

Currently the Inspectorate of Education is preparing a so called thematic research study on educational leadership (Inspectorate of Education, 2012)¹⁹

Points of attention in this (survey) study will be:

- the school leader's level of formal training and competencies
- the division of roles and tasks within the leadership team
- the division of roles and tasks of the school leadership vis a vis the governing board of the schools
- the way educational leadership is given shape

The results of this study are to be published in the 2014 version of the Annual Inspection Report.

Discussion

More attention for the role of the school leader in the Supervision Frameworks of the Inspectorate and in their thematic research studies seems a logical step, given a stronger focus on School Boards on the one hand²⁰ and the functioning of teachers on the other. Still appraisal of school leadership is a part of school evaluation, and not a case of the appraisal of individual school leaders.

¹⁹ Inspectorate of Education. (2012) Inspectie Breed Projectplan. Themaonderzoek JWP 2013. Onderwijskundig Leiderschap van de Schoolleiding in PO, VO, VO-SP en BVE. Utrecht: Inspectie van het Onderwijs

²⁰ Inspectorate of Education (2013) Toezicht op Bestuurlijk Handelen. Stand van Zaken en Plannen, Inspectie van het Onderwijs. Utrecht: Inspectie van het Onderwijs.

PART IV: STUDENT EVALUATION

10. Using the CITO school leavers test at the end of primary school for different purposes

After the original country background report was finished in June 2012, there has been a lot of debate concerning the function of the CITO school leavers test. The original report summarized the situation in June 2012 as follows:

“In a general sense the emphasis on result orientation, improvement of achievement in basic subjects, and public accountability have created a climate that is favorable to a further development and optimization of student assessment. At the same time the intention to move to a central obligatory test at the end of primary schools, appears to be a sensitive process. In a recent advice the Education Council, which is the most important advisory body to the Minister of Education, states that it prefers a situation where the central level formulates substantive demands to the test to a uniform obligatory test. The Council fears undesirable side effects of a centrally established, obligatory test for all primary schools (Education Council, 2012, p. 13). As a matter of fact the advice in question, titled “Controlled Space”, breathes concern for school autonomy in realizing a broad set of educational and pedagogical values, in a situation in which much emphasis is given to basics, and measurement of outcomes in basic subjects.”

The debate was running so high that one of the teacher unions actually called for a strike against the CITO test. The direct reason was the decision of the Secretary of State for Education to make the CITO test mandatory and make the school results public. As a kind of compromise solution the date of administration was pushed back, in such a way that the test results will now only be available when the decision for secondary school choice has already been made. Advice on secondary school choice to parents will now be based exclusively on the “non test based” advice of the primary school. The implication of this change in policy is that the CITO test has lost its original purpose, namely to guarantee an objective measure in primary schools’ advising parents on the kind of secondary school their child should attend. In stead school-accountability has now become the prime purpose of the CITO test.

It is interesting to review the kind of arguments that were used in the debate:

- the Educational Council appears to be concerned with an institutional monopoly by CITO, and has pleaded for other tests being admissible as well;
- the Educational Council also fuels the debate on the CITO test being a too narrow achievement measure, and repeatedly emphasizes the broader developmental aims of schools
- quite a few critical articles in the press just express aversion to standardized tests, and claim that these tests dehumanize education and are a threat to the professional autonomy of teachers (Van Leeuwen and Jongerius, Volkskrant, 9 March, 2013)
- the debate about the CITO test is also connected to protests against using tests to evaluate teachers; although the plans for merit pay of teachers of the previous government have been abandoned by the current Cabinet.

On the issue of it being admissible, given psychometric properties, to use one and the same test for different objectives, in this case advising parents and judging the quality of schools, experts seem to hold divergent opinions. In actual practice such eventual problems have not played a

decisive role in the Inspectorate's use of the CITO test to judge schools. A more serious issue, connected to administering the CITO test later in the school year and diminishing its role for student placement in secondary schools, is that this is likely to affect the motivation of students to do well on the test. The recently announced shift in purpose of the CITO tests takes away the high stakes nature of the test for students.