



OECD Reviews of Evaluation and Assessment in Education

NETHERLANDS

Deborah Nusche, Henry Braun,
Gábor Halász and Paulo Santiago



OECD Reviews of Evaluation and Assessment in Education: Netherlands 2014

Deborah Nusche, Henry Braun,
Gábor Halász and Paulo Santiago

This work is published under the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of OECD member countries.

This document and any map included herein are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Please cite this publication as:

Nusche, D., et al. (2014), *OECD Reviews of Evaluation and Assessment in Education: Netherlands 2014*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing.
<http://dx.doi.org/10.1787/9789264211940-en>

ISBN 978-92-64-21193-3 (print)

ISBN 978-92-64-21194-0 (PDF)

Series: OECD Reviews of Evaluation and Assessment in Education

ISSN 2223-0947 (print)

ISSN 2223-0955 (online)

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Photo credits: Cover © iQoncept - Fotolia.com, © AKS - Fotolia.com, © Sergej Khackimullin - Fotolia.com.

Corrigenda to OECD publications may be found on line at: www.oecd.org/about/publishing/corrigenda.htm.

© OECD 2014

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of the source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to rights@oecd.org. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at info@copyright.com or the Centre français d'exploitation du droit de copie (CFC) at contact@cfcopies.com.

Foreword

This report for the Netherlands forms part of the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes (see Annex A). The purpose of the Review is to explore how systems of evaluation and assessment can be used to improve the quality, equity and efficiency of school education. The Review looks at pupil assessment, teacher appraisal, school leader appraisal, school evaluation and system evaluation, and how these help to improve pupils' learning.

The Netherlands' involvement in the OECD Review was co-ordinated by Hans Stegeman at the Ministry of Education, Culture and Science and Marian Hulshof at the Inspectorate of Education, assisted by Jessica Villamil. The Netherlands opted to participate in the country review strand of the international project and host a visit by an external review team. Members of the OECD review team were Deborah Nusche (OECD), co-ordinator of the review; Henry Braun (Professor of Education and Public Policy in the Lynch School of Education at Boston College), Gábor Halász (Professor of Education in the Faculty of Pedagogy and Psychology at the University Eötvös Loránd in Budapest) and Paulo Santiago (OECD) (see Annex C). This publication is the report from the OECD review team. It provides, from an international perspective, an independent analysis of major issues facing the evaluation and assessment framework in the Netherlands, current policy initiatives, and possible future approaches. The report serves three purposes: (1) provide insights and advice to the Dutch authorities; (2) help other OECD countries understand the Dutch approach; (3) improve the international knowledge on evaluation and assessment policies.

An important part of the Netherlands' involvement was the preparation of a Country Background Report (CBR) on evaluation and assessment policy developed by Jaap Scheerens, Melanie Ehren, Peter Slegers and Renske de Leeuw (2012) at the University of Twente. In addition, an Update to the Country Background Report was developed by Jaap Scheerens (2013) in preparation of the OECD review visit. The OECD review team is grateful to the authors and to all those who assisted them for compiling these high quality and informative reports as background to the review and analysis. The CBR and the Update are important outputs from the OECD project in its own right as well as an important source for the OECD review team. Unless indicated otherwise, the data for this report are taken from the Dutch CBR and the Update. The CBR follows guidelines prepared by the OECD Secretariat and provides extensive information, analysis and discussion in regard to the national context, the organisation of the school system, the main features of the evaluation and assessment framework and the views of key stakeholders. In this sense, the CBR and this OECD review report complement each other and, for a more comprehensive view of evaluation and assessment in the Netherlands, should be read in conjunction.

The review visit to the Netherlands took place on 3-10 June 2013. The itinerary is provided in Annex B. The visit was designed by the OECD in collaboration with the Ministry of Education, Culture and Science and the Inspectorate of Education. During the review visit, the team held discussions with a wide range of officials within the Ministry

of Education, Culture and Science, the Inspectorate of Education, the Central Institute for Test Development, the College for Examinations, the National Bureau of Statistics, the Court of Audit, the Education Council, the Primary and Secondary Education Councils, the Institute for Curriculum Development and the National Knowledge Centre on Special Education. The review team also met with stakeholder groups including representatives of teacher, school leader, student and parent organisations, teacher education institutions, educational service centres, private test providers, and researchers with an interest in evaluation and assessment issues. The team visited a range of schools, interacting with representatives of the school boards, school leadership teams, teachers, students and parents in The Hague and its surrounds and also in Hulshorst, Groningen and Rotterdam. The intention was to provide a broad cross-section of information and opinions on evaluation and assessment policies and how their effectiveness can be improved.

The OECD review team wishes to express its sincere gratitude to all those who took time to meet with us and to share their insights. Our overwhelming memory is of a warm welcome wherever we went and frank, open, constructive and stimulating discussions. We wish to thank the Ministry of Education, Culture and Science and the Inspectorate for Education for meeting our challenging demands to organise a review visit that allowed us to learn different perspectives from many different stakeholders (see Annex B). In particular, we warmly thank Hans Stegeman, Marian Hulshof and Jessica Villamil for accompanying us at different stages of the review, making sure that everything ran smoothly and allowing us to concentrate on learning from the people we met. The courtesy and hospitality extended to us throughout our stay in the Netherlands made our task as a review team as pleasant and enjoyable as it was stimulating and challenging.

In particular, during the OECD review, we asked all stakeholders to provide us with information and evidence on assessment and evaluation practices in the Netherlands. We wish to extend particular thanks to all those who chose to do so and to send research and information on practices to us during and after the review visit. This vast information base formed a rich resource and helped the OECD review team to develop a deeper understanding of evaluation and assessment in the Netherlands. While our report may not refer explicitly to some of this material, it was invaluable in underpinning our initial analysis. Any misunderstandings or misinterpretations are entirely our responsibility.

The OECD review team is also grateful to Liz Zachary for editorial and publication support on this report and to Heike-Daniela Herzog for administrative support.

This report is organised in six chapters. Chapter 1 provides the national context, with information on the Dutch school system, main trends and concerns, and recent developments. Chapter 2 looks at the overall evaluation and assessment framework and analyses how the different components of the framework play together and can be made more coherent to effectively improve student learning. Chapters 3 to 6 present each of the components of the evaluation and assessment framework – pupil assessment, teacher appraisal, school evaluation and system evaluation – in more depth, presenting strengths, challenges and policy recommendations.

The policy recommendations attempt to build on and strengthen reforms that are already underway in the Netherlands, and to build on the professionalism and strong commitment to further improvement that was evident among those we met. The suggestions should take into account the difficulties that face any visiting group, no matter how well briefed, in grasping the complexity of the Dutch school system and fully understanding all the issues.

Of course, this report is the responsibility of the OECD review team.

Table of contents

Acronyms and abbreviations	9
Chapter 1 School education in the Netherlands.....	17
National context	18
Main features of the school system	18
Quality and equity in Dutch education.....	23
Main policy developments	27
References	33
Chapter 2 The evaluation and assessment framework.....	35
Context and features.....	36
Strengths.....	39
Challenges.....	45
Policy recommendations	50
References	57
Chapter 3 Student assessment	59
Context and features.....	60
Strengths.....	65
Challenges	69
Policy recommendations	74
References	89
Annex 3.1 Research on 21st century skills.....	93
Chapter 4 Teacher appraisal	95
Context and features.....	96
Strengths.....	100
Challenges	106
Policy recommendations	111
References	123
Chapter 5 School evaluation	127
Context and features.....	128
Strengths.....	136
Challenges	142
Policy recommendations	152
References	163
Chapter 6 Education system evaluation	169
Context and features.....	170
Strengths.....	177
Challenges	183

Policy recommendations	185
References	192
Conclusions and recommendations	193
School system context	193
Strengths and challenges	194
Policy recommendations	203
Annex A: The OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes	211
Annex B: Visit itinerary	213
Annex C: Composition of the Review team	215

Figures

Figure 1.1 Structure of the Dutch school system and student movements	20
Figure 1.2 Percentage of decisions taken at each level of government in public lower secondary education (2011)	21
Figure 3.1 The “education triangle”	75
Figure 5.1 Risk-based inspections in the Netherlands	131

Tables

Table 1.1 Performance of Dutch students at the international benchmarks of achievement in primary education	25
Table 1.2 Selected indicators of quality and equity in Dutch education, based on PISA 2012	27
Table 2.1 Main instruments used for evaluation and assessment in the Dutch school system (2013).....	38
Table 5.1 Instruments and mechanisms for school evaluation in the Netherlands	129
Table 6.1 Main attainment targets for specific Action Plans	172

Boxes

Box 3.1 New Dutch laws on student assessment developed in 2013.....	60
Box 3.2 Dutch student monitoring systems developed by Cito	62
Box 3.3 New South Wales, Australia: The School Measurement, Assessment and Reporting Toolkit	78
Box 3.4 United States: The National Education Technology Plan	79
Box 3.5 Moderation of examinations in upper secondary education in Queensland, Australia and New Zealand	82
Box 3.6 Norway: Centrally-supported networks on assessment.....	84
Box 3.7 Defining and communicating balanced assessment strategies in Canada and New Zealand	85
Box 4.1 The teaching profession in the Netherlands: Main features	99
Box 4.2 The seven domains of the Dutch teacher competency requirements	101
Box 4.3 The Dutch Inspectorate’s indicators for lesson observation.....	102
Box 4.4 Key elements of the Dutch Action Plan <i>Teaching 2020: A Strong Profession</i>	103
Box 4.5 Memphis, Tennessee, United States: Linking teacher appraisal to professional development..	113
Box 4.6 Building school leader capacity for teacher appraisal and evaluation in Norway and Northern Ireland, United Kingdom.....	114
Box 4.7 Support for beginning teachers in Northern Ireland, United Kingdom.....	116

Box 4.8 Teacher career structures in Australia and Singapore	117
Box 4.9 Teacher registration in Australia	120
Box 4.10 The National Board for Professional Teaching Standards in the United States	121
Box 5.1 Quality inspections in the Netherlands: Domains and quality aspects considered in primary education	132
Box 5.2 Windows for Accountability: Indicators for Dutch secondary schools.....	135
Box 5.3 Growth models and value-added models	139
Box 5.4 Main principles of the Dutch inspection reform	141
Box 5.5 Risks stemming from changes to the Dutch legislation on end-of-primary testing	148
Box 5.6 Norway: Local and regional networks for school quality improvement	153
Box 5.7 Targeted training on school self-evaluation for school principals in New Zealand.....	154
Box 5.8 Emerging school peer reviews in the Flemish Community of Belgium	155
Box 5.9 Factors promoting use of assessment data at the school level.....	157
Box 6.1 The Measurement Framework for Australian Schooling.....	187
Box 6.2 New Zealand's National Education Monitoring Project (NEMP).....	189

Acronyms and abbreviations

AST	Advanced Skills Teaching (Australia)
AVS	Dutch School Leaders Association
BRON	Basic Register for Education (<i>Basis Register Onderwijs</i>)
CPB	Netherlands Bureau for Economic Policy Analysis
CBS	National Bureau of Statistics
CASS	Curriculum Advisory and Support Services (Northern Ireland, UK)
Cito	Central Institute for Test Development
CIVED	Civic Education Study
CVE	College for Examinations
COOL	Cohort Survey School Careers (<i>Cohort Onderzoek Onderwijsloopbanen</i>)
DI	Differentiated Instruction
DUO	Implementation of Education Service (<i>Dienst Uitvoering Onderwijs</i>)
ECD	Evidence Centred Design
ECTS	European Credits Transfer System
ENIRDELM	European Network for Improving Research and Development in Educational Leadership and Management
EPMS	Enhanced Performance Management System (Singapore)
EPD	Early Professional Development
ESIS	Electronic School Information System (<i>Elektronisch School Informatie Systeem</i>)
ETI	Education and Training Inspectorate (Northern Ireland)
ERO	Education Review Office (New Zealand)
EU	European Union
GTCNI	General Teaching Council for Northern Ireland
HAT	Highly Accomplished Teacher (New South Wales, Australia)
HAVO	Tertiary education (first three years of general secondary education)
HBO	Higher Vocational Education (<i>Hoger beroeps onderwijs</i>)
HRM	Human Resource Management
ICCS	International Civic and Citizenship Education Study
ICILS	International Computer and Information Literacy Study
ICT	Information and Communications Technology
IEA	International Association for Educational Achievement
IPB	Integrated Personnel Policy (<i>Integraal Personeelsbeleid</i>)
JPON	Annual Survey of Educational Levels (<i>Jaarlijks Peilingsonderzoek van het Onderwijsniveau</i>)

LAKS	National Student Organisation
LECSO	National Knowledge Centre on Special Education
LVS	Student Monitoring System (<i>Leerling Volg Systeem</i>)
LWOO	Learning support programmes
MBO	Further vocational training
NBPTS	National Board for Professional Teaching Standards (United States)
NEET	Not in Employment, Education or Training
NEMP	National Education Monitoring Project (New Zealand)
NETP	National Education Technology Plan (United States)
NSA	Dutch School Leaders Academy
NRO	Netherlands Initiative for Education Research
NWO	Netherlands Organisation for Scientific Research
OECD	Organisation for Economic Cooperation and Development
PARCC	Partnership for the Assessment of Readiness for College and Career
PIRLS	Progress in Reading Literacy Skills
PISA	Programme for International Student Assessment
PO	Primary education
PPON	Periodical Survey of Education (<i>Periodieke Peiling van het Onderwijsniveau</i>)
PRIMA	Primary Education and Special Education Cohort Studies (<i>Primair onderwijs en special onderwijs cohortonderzoeken</i>)
PRO	Programme of practical training
ROA	Research Centre for Education and the Labour Market
SER	Social-Economic Council
SLO	General Institute for Curriculum Development
SBAC	Smarter Balanced Assessment Consortium (United States)
SEN	Special Educational Needs
SICI	Standing International Conference of Inspectorates
SMART	School Measurement, Assessment and Reporting Toolkit (New South Wales, Australia)
TALIS	Teaching and Learning International Survey
TIMSS	Trends in Mathematics and Science Skills
UNESCO	United National Educational, Scientific and Cultural Organisation
VAM	Value Added Models
VET	Vocational Education and Training
VMBO	Pre-vocational secondary education (<i>Vorbereidend Middelbaar Beroepsonderwijs</i>)
VO	Secondary education
VOCL	Secondary Education Cohort Study (<i>Voortgezet Onderwijs Cohort Leerlingen</i>)
VSO	Special secondary education
WVO	Pre-university education (<i>Vorbereidend Wetenschappelijk Onderwijs</i>)
WO	Academic higher education

List of Dutch organisations referred to in English throughout the report

Central Institute for Test Development	Centraal Instituut voor Toetsontwikkeling, Cito
College for Examinations	College voor Examen, CVE
Dutch School Leaders Association	Algemene Vereniging Schoolleiders, AVS
Dutch School Leaders Academy	Nederlandse Schoolleiders Academie, NSA
Education Cooperative	Onderwijscoöperatie
Education Council	Onderwijsraad
Education Foundation	Stichting van het Onderwijs
Foundation Schoolinfo	Stiching Schoolinfo
General Institute for Curriculum Development	Instituut voor Leerplan-Ontwikkeling, SLO
Inspection Council	Inspectieloket
Inspectorate of Education	Inspectie van het Onderwijs
Ministry of Education, Culture and Science	Ministerie van Onderwijs, Cultuur en Wetenschap, OCW
Ministry of Finance	Ministerie van Financiën
National Bureau of Statistics	Centraal Bureau voor de Statistiek, CBS
National Knowledge Centre on Special Education	Landelijk Expertise Centrum Speciaal Onderwijs, LECSO
National Student Organisation	Landelijk Aktie Komitee Scholieren, LAKS
Netherlands Bureau for Economic Policy Analysis	Centraal Planbureau, CPB
Netherlands Initiative for Education Research	Nationaal Regieorgaan Onderwijsonderzoek, NRO
Netherlands Organisation for Scientific Research	Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO
Primary Education Council	PO-raad
Research Centre for Education and the Labour Market	Onderzoeksinstituut voor onderwijs en de arbeidsmarkt, ROA
Secondary Education Council	VO-raad
Social-Economic Council	Sociaal-Economische Raad, SER
Netherlands Court of Audit (High Council of State)	Algemene Rekenkamer

Executive summary

Compared internationally, the Dutch education system achieves very good results. It has made progress on many fronts and has a high standing on international assessments. Yet, there is a general appreciation that the system must continue to improve and strive for the next level. The nature of that next level, however, has not yet been specified. According to various groups interviewed by the OECD review team, it could mean further improving the country's standing on international assessments, with particular focus on increasing the proportions of students in the highest category; or enhancing general academic achievement and responding better to the learning needs of different student groups; or turning attention to “21st century skills” such as creativity, collaboration and ICT literacy. Evaluation and assessment can play an important role in supporting the attainment of whatever educational goals are decided upon.

In many ways, the evaluation and assessment policies developed in the Dutch school system are in line with the principles identified by the OECD to develop a comprehensive and balanced evaluation and assessment framework. The Dutch evaluation and assessment approach stands out internationally as striking a good balance between school-based and central elements, quantitative and qualitative approaches, improvement and accountability functions and vertical and horizontal responsibilities of schools. Building on recent reforms and developments already underway, this report suggests a range of policy options to further develop and integrate the overall evaluation and assessment framework while maintaining its balanced approach.

Embed the evaluation and assessment framework with broader learning goals

A crucial aspect for the successful implementation of evaluation and assessment is their alignment with student learning objectives. Thus, it seems essential to begin as soon as possible a broad-based consultative process to build consensus on a set of long-term learning goals for Dutch students that will prepare them well for the mid-21st century. A national conversation on how traditional learning goals should be augmented to meet the challenges of the 21st century will help identify the changes that must be made to the evaluation and assessment framework to support innovative, future-oriented and reflective teaching and learning. To make the evaluation and assessment system coherent, it is important that the learning goals to be achieved are placed at the centre of the framework and that all other evaluation and assessment activities align to work towards these goals. For example, approaches to student assessment, competency descriptions for school professionals and quality indicators for school and system evaluation should reflect the broad set of learning goals that the school system is aiming to achieve.

Integrate teacher appraisal into the evaluation and assessment framework

As the most important school-level factor in student achievement, teachers are key to improving education outcomes. The OECD review team commends the current focus of the Dutch government to build on and further develop teacher professionalism in the Netherlands. Defining and rethinking the framework for evaluation and assessment also provides an opportunity to place teacher professionalism firmly at the heart of the evaluation and assessment agenda. The development of a comprehensive appraisal and feedback system for teachers could help increase the focus on teaching quality and continuous professional learning while also bringing recognition to effective teachers and spreading good practice more widely. To this end, the teaching profession in the Netherlands could benefit from a revised and refined set of teaching standards; strengthened school-based appraisal processes linked to professional learning opportunities, especially for beginning teachers; and an enhanced registration system that could be linked to teacher career development. To build synergies between different aspects of evaluation and assessment, teacher appraisal and professional development should also be linked to school development priorities, and the external review of schools should comprise the monitoring of school-based practices to review and enhance teaching quality.

Continue to adapt school evaluation to emerging needs

Given the multi-actor nature of school evaluation in the Netherlands, the Inspectorate should continuously map the environment in which it operates at both the national and local level, and take into account the potential impact of the other partners when designing its school evaluation approaches. The current move towards the introduction of differentiated inspections provides a good opportunity to reconsider the relationship between internal and external school evaluation, and self-evaluation could gain a greater role in school inspections. To evaluate and further stimulate schools that have developed innovative approaches, external evaluators need in-depth knowledge of innovative learning environments and an understanding of how to evaluate practices that are not yet widely proven as resulting in good quality. The policy focus on teacher professionalism means that external evaluators will need to understand the complexities of internal human resource management and development practices in schools. Further innovations may be required, for example, modifying the Inspectorate's classroom observation framework or involving teachers as part of inspection visits. In light of the general concerns regarding the use of value-added models (VAM) for accountability, the Inspectorate should move cautiously in formally incorporating VAM estimates into the school indicator system. The current focus should be on helping participating schools to use such information as a component of more comprehensive self-evaluation efforts and improvement strategies, while enlisting a larger sample of schools to participate in future pilot studies.

*Strengthen evaluation and assessment
competencies across the education system*

In the context of freedom of education, the successful implementation of evaluation and assessment in the Netherlands crucially depends on the competencies of school professionals. Teachers are responsible for student assessment and need to build strong competencies in this domain. To this end, assessment-related competencies should be reflected in teacher standards, appraisal and registration, and be addressed in a coherent way across initial teacher education and publicly funded professional development. School leaders and their teams at the school level are responsible for both teacher appraisal and school self-evaluation and would benefit from a comprehensive strategy to support them in these tasks. This could include: considering the implementation of mandatory training for school leadership; ensuring that school leaders receive adequate appraisal and feedback from their employers; promoting peer evaluation, professional networks and partnerships among schools; and allowing greater access for school leaders to participate in external reviews and development work with other schools. School governing boards also share responsibility for human resource management and school quality assurance and could benefit from professional learning opportunities to enhance their capacities to appraise school leaders and monitor school quality.

Chapter 1

School education in the Netherlands

The Netherlands has one of the OECD's most devolved education systems, with schools enjoying a high degree of autonomy. School autonomy is grounded in the principle of "freedom of education", which gives the right to any natural or legal person to set up a school and to organise teaching. This constitutional arrangement puts public and private schools on an equal footing, with all schools receiving public funding, provided that they meet the requirements for schools in their sector. While "freedom of education" implies that schools are free to determine the content and methods of teaching, the central government sets learning objectives and quality standards that apply to both public and private schools. The Inspectorate of Education monitors school quality and compliance with central rules and regulations. Current policy priorities include promoting excellence among students and schools, raising teacher professionalism and enhancing results-oriented work in schools. Evaluation and assessment are key elements in the drive to achieve these goals, as reflected in recent laws on student assessment making it mandatory for primary schools to implement regular student monitoring systems as well as a standardised end-of-primary test.

National context

Population

The Netherlands has 16.8 million inhabitants.¹ It is densely populated and highly urbanised, with 84% of the total population living in urban areas in 2012.² As in other European countries, the birth rate is declining, but with 1.8 children per woman, it remained higher than the European Union (EU) average of 1.6 in 2011.³ Migration is the main driver of population growth and potential educational expansion. Immigration to the Netherlands has continuously increased since 2005. In early 2012, 21% of Dutch residents were non-native, with the largest immigrant groups coming from Turkey, Indonesia and Germany (OECD, 2013a).

Governance

Dutch policy-making has a long tradition of power-sharing and consensus decision-making. The government is typically formed by a coalition of political parties. Both trade unions and employer organisations are consulted in economic and social decisions, and meet regularly with the government as part of the Social-Economic Council (*Sociaal-Economische Raad*, SER) (Bal and De Jonge, 2007). In 2012, 57% of the Dutch population expressed that they had confidence in their national government, compared to 66% in 2007. The proportion of citizens with trust in their government remains above the OECD average of 40% in 2012 (OECD, 2013b).

Economy

The Netherlands is a wealthy country, with its GDP per capita ranked eighth in the OECD in 2012⁴. Globalisation has had a major impact on the Dutch economy, given its important role as a European trade and transportation hub. After a long period of uninterrupted economic growth, the Dutch economy has weakened since the beginning of the global economic and financial crisis in 2008. There has been a significant increase in unemployment rates in recent years, from 3% in 2008 to 7% in December 2013, but the proportion of unemployed persons remains slightly below the OECD average of 8%, and well below the EU average of 11%.⁵ The increase in unemployment has affected people differently depending on their age and educational attainment, with the youngest and least-educated cohort being hit the hardest (OECD, 2013c).

Main features of the school system

Structure of the school system⁶

Education is compulsory from age five to sixteen, but students can voluntarily enter primary education at age four. Figure 1.1 illustrates the structure of the Dutch school system and student movements across different school types. The Dutch school system is organised into two phases:

1. **Primary education** (PO, typical ages 4-12) lasts for eight years. There is also an offer of special primary education for children with special educational needs aged 3 to 12. Schools are free to determine the content and methods of teaching, but their work must be based on national attainment targets and reference levels for literacy and numeracy. At the end of primary education, students receive a school report describing their achievement levels and potential. They transfer into

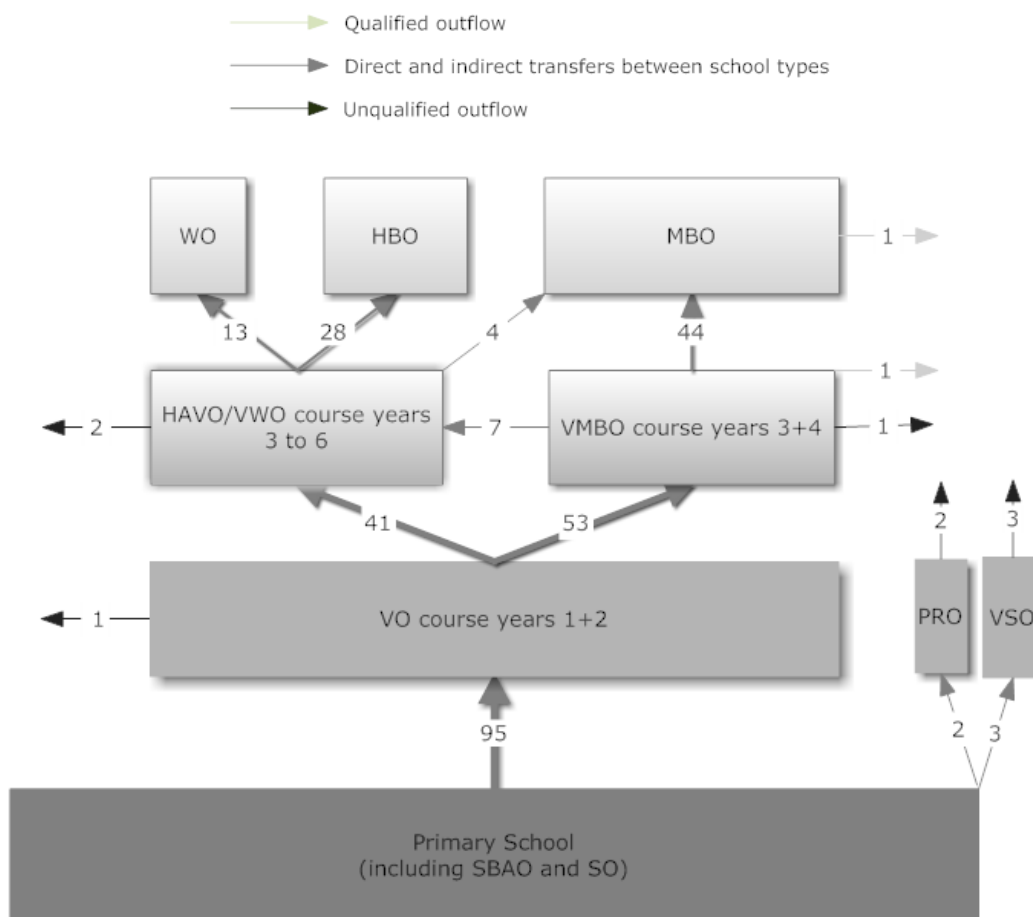
different types of secondary education, based on their achievement and the advice of their primary school teacher.

2. **Secondary education** (VO, typically from age 12) lasts between four and six years. The secondary school system is highly stratified, even though most secondary schools offer a number of different programmes in the same building. The first years of secondary education (VO 1/2) are intended to provide students with a shared curriculum of basic general education⁷. However, most secondary schools stream their students at this stage, foreshadowing the later tracking into different programmes. There are two main learning pathways: pre-vocational education (VMBO) and education preparing for tertiary education (HAVO and VWO).
 - **Pre-vocational education** (VMBO, typical ages 12-16) caters to vocationally-oriented students and lasts for four years. It is intended as a foundation course providing a basis for further vocational training (MBO), which is focused on preparing students for the labour market. Students in VMBO can choose among four learning programmes⁸. Students who complete the theoretical VMBO programme may choose to transfer to HAVO. In 2010, 53% of students having completed primary education entered VMBO programmes.
 - **Pre-tertiary education** exists in two main forms. Senior general secondary education (HAVO, typical ages 12-17) lasts for five years and is intended to provide students with a basic general education, preparing them for professional higher education (HBO), typically provided at universities of applied sciences. Pre-university education (VWO, typical ages 12-18) lasts for six years and is intended to prepare students for progression to academic higher education (WO), typically provided at universities⁹. In 2010, 41% of students having completed primary education entered HAVO or VWO programmes.

At the secondary level, there is also an offer of special secondary education (VSO) intended for students with special educational needs. In 2010, 3% of primary school leavers entered VSO. In addition, there is a programme of practical training (PRO), intended for students who are considered unlikely to obtain a qualification through one of the other learning pathways, even with special support. Practical training prepares students for direct entry to the labour market. In 2010, 2% of primary school leavers entered PRO. Many VMBO schools also provide learning support programmes (LWOO) to students with learning difficulties.

Figure 1.1 Structure of the Dutch school system and student movements

In percentages of a cohort of students leaving primary education, 2010



Source: Dutch Ministry of Education, Culture and Science (2012), *Key Figures 2007-2011: Education, Culture and Science*, Dutch Ministry of Education, Culture and Science, The Hague.

Notes: HAVO = General secondary education; HBO = Professional higher education; MBO = Vocational education; PRO = Elementary vocational training; SBAO = Special primary education; SO = Special education; VMBO = Pre-vocational secondary education; VO = Secondary education; VSO = Secondary special education; VWO = Pre-university education; WO = Academic higher education.

Governance of the school system

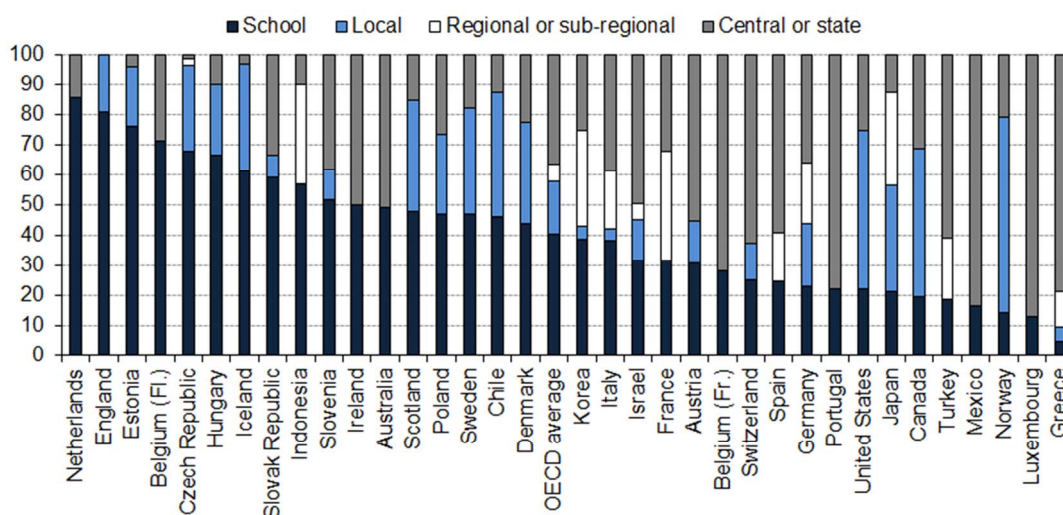
School autonomy and freedom of education

The Netherlands has one of the OECD's most devolved education systems, with schools enjoying a high degree of autonomy (Figure 1.2). Since the 1980s, Dutch schools have acquired increasing levels of responsibility while local governments play only a minor role. In lower secondary education, schools make 86% of the decisions (compared to an OECD average of 41%), and the central government makes 14% of the decisions. A more detailed look at the different domains of decision making reveals that schools in the Netherlands make 100% of the decisions regarding the organisation of instruction,

personnel management and resource management, but only 43% of the decisions regarding planning and structures¹⁰ (OECD, 2012).

School autonomy is grounded in the principle of “freedom of education”, which has been guaranteed by the Dutch Constitution since 1917. Freedom of education gives the right to any natural or legal person to set up a school, to organise teaching, and to determine the (educational, religious or ideological) principles on which teaching is based. This constitutional arrangement puts public and private schools on an equal footing, with all schools receiving public funding, provided that they meet the requirements for schools in their sector. Parents have free school choice, and funding “follows the student”, which lays the foundation for potentially strong competition among schools.

Figure 1.2 Percentage of decisions taken at each level of government in public lower secondary education (2011)



Source: OECD (2012), *Education at a Glance 2012: OECD Indicators*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/eag-2012-en>.

Note: Countries are displayed in descending order of the percentage of decisions taken at the school level.

The role of school boards

Every school is governed by a legally recognised competent authority, or school board, which oversees the implementation of legislation and regulations in the school and employs teachers and other staff. Traditionally, public schools were established by the state and governed mostly by local government. But since the 1980s, there has been a trend of local government setting up separate legal entities or foundations as school boards and transferring parts or all of the governance of schools to them. In addition, the Dutch education system has a large sector of government-dependent private schools, i.e. schools that are privately run but publicly funded. These schools are typically based on religious or pedagogical-didactical principles. While public schools are open to all students, government-dependent private schools may refuse students whose parents do not subscribe to the school’s profile or principles. In 2011, 24% of students in the primary sector were enrolled in Protestant schools, 34% in Roman Catholic schools and 11% in

other private schools. The remaining 31% of students were enrolled in public schools (Dutch Ministry of Education, Culture and Science, 2012).

The composition of school boards varies widely across the Netherlands. Governors may be volunteers (laypersons receiving an honorarium) or professionals (receiving a salary). While volunteer governors join a school board through co-optation, professional governors are typically appointed by their internal supervisors.¹¹ Since they are appointed, governors are trustees of the school rather than representatives. Governors are typically citizens from the local community, members of a religious or life philosophy community or professionals with specific expertise such as law, finance, human resource management or education. In 2011, 48% of all school governors in primary education were parents, but they were not evenly distributed across school boards. School boards may comprise only parents, a mix of parents and professional governors, or no parents at all (Hooge and Honingh, 2014). A governing board can be responsible for more than one school. In 2012, there were 1 169 school boards in the primary sector, of which 75% were responsible for fewer than 10 schools and 25% were responsible for more than 10 schools¹².

Central steering, inspection and support

The Ministry of Education, Culture and Science has responsibility for the overall education system, but it does not interfere with the organisation of individual schools, unless there is evidence of insufficient quality or financial mismanagement. The Ministry's responsibilities relate mainly to setting legislation and determining the structure and funding mechanisms of the education system. In the context of teacher policy, the Dutch Government (2011, p.3) described the distribution of responsibilities for educational reform as follows: "the government will establish the objectives of the policy measures ('*what*') while the field itself will decide how best to pursue those objectives ('*how*')."

The Ministry can also control the system by setting quantitative or qualitative standards, attainment targets and examinations. In addition, the Minister of Education is responsible for school inspection, which is carried out by the Inspectorate of Education. The Inspectorate has professional independence and develops an annual planning of activities, which is sent to the Parliament after approval by the Minister of Education¹³. The Inspectorate monitors both the quality of education provided by schools and their compliance with statutory and financial rules and regulations.¹⁴ It also checks that schools assure their own quality effectively (Chapter 5).

There is a large intermediary structure of school support organisations, some of which are organised according to religious denominations. The Council for Primary Education (*PO-raad*) and the Council for Secondary Education (*VO-raad*) represent the employers (school boards) of their respective sectors and offer support services to schools, such as a team of 'flying brigades' that work with schools identified by the Inspectorate as weak or unsatisfactory (Chapter 5).

Curriculum and learning objectives

While the principle of freedom of education implies that schools are free to determine the content and methods of teaching, the central government sets quality standards that apply to both public and private schools. There is no national curriculum, but the Ministry of Education, Culture and Science establishes core learning objectives that students are expected to meet at the end of primary and lower secondary education. These objectives

have been set for Dutch, English, mathematics, social studies, science, arts and physical education, as well as for a number of subjects within the different secondary education programmes. For example, in lower secondary education, 58 core objectives cover all subjects. The core learning objectives provide a legal prescription for the knowledge and skills students are expected to achieve at the end of primary and lower secondary education. Each school is responsible for developing its own curriculum in agreement with the core objectives. For the upper cycle of secondary education, the Ministry of Education, Culture and Science has formulated learning targets, which have been translated into centrally set examination syllabi. These strongly influence the curricula taught in upper secondary education.

As the core learning objectives are described in very broad terms, the Ministry of Education, Culture and Science has developed additional reference levels for literacy and numeracy. Schools have been required to implement these since 2010 and will be required to report on their students' performance in relation to the reference levels from the 2015/16 school year. The reference levels provide a general description of knowledge and skills to be achieved, a description of the types of tasks to be mastered, and the criteria that these tasks should meet. The reference levels are defined for the end of primary education and the end of each of the educational tracks in secondary education (VMBO, HAVO and VWO). For each of these stages, they indicate a "fundamental" level to be achieved by all students and an "advanced" level for gifted students. In the 2014/15 school year, all schools will participate in a pilot study to evaluate the coverage of the reference levels in available end-of-primary tests.

Quality and equity in Dutch education

Attainment and participation in education

Attainment rates of the Dutch population are similar to the OECD average and show positive trends. In 2011, 72% of the population aged 25-64 had completed at least upper secondary education, compared to an OECD average of 75%. Among the generation of 25-to 34-year-olds, 82% had attained at least upper secondary education (the same as on average across the OECD) and 40% had completed tertiary education (compared to an OECD average of 39%). Current estimates indicate that 92% of today's young people in the Netherlands will complete upper secondary education over their lifetime, compared to 83% across the OECD (OECD, 2013c).

In 2011, the Netherlands had the lowest rate of 15-29 year-olds not in employment, education or training (NEET) across all OECD countries: 7% compared to an OECD average of 16%. However, this proportion has been increasing since the beginning of the financial crisis in 2008, particularly among young people with a tertiary degree where the proportion of NEETs more than doubled from 2% in 2008 to 5% in 2011. However, it still remained significantly below the OECD average of 13% (OECD, 2013c).

As in other OECD countries, there has been a trend in the Netherlands of students opting increasingly for academic rather than vocational paths. Between 1990 and 2011, the proportion of students in pre-vocational education (VMBO) decreased from 58% to 39%, while the share of students in general and pre-university education (HAVO and VWO) rose from 32% to 44% (Dutch Ministry of Education, Culture and Science 2012). However, there is evidence of regional disparities in the distribution of students across different programmes. For example, in 2011, the proportion of students who passed the diploma in pre-university education (VWO) compared to all final examination candidates

ranged from below 10% in some municipalities to over 60% in others (Dutch Inspectorate of Education, 2013a).

National data also indicate that the distribution of students across different educational programmes in secondary education is closely related to parental income and origin. Among the cohort that entered secondary education in 2005, students of the top quartile of parental income were four times more likely to be in pre-university education (VWO) four years later (2008/09) than children from the bottom quartile. By contrast, pupils from the bottom quartile of parental income were more than five times more numerous in the basic vocational programme (VMBO-BL) than their top-quartile counterparts. Similarly, students with an immigrant background were underrepresented in general and pre-university programmes (HAVO and VWO) and overrepresented in vocational programmes. For example, in 2010/11, only 30% of non-Western ethnic-minority pupils were enrolled in general education or pre-university programmes (HAVO or VWO), compared to almost 50% of the native Dutch population (Dutch Ministry of Education, Culture and Science, 2012).

There are concerns about increasing numbers of students being diagnosed as having special educational needs. According to the Dutch Ministry of Education, Culture and Science (2012), the proportion of special needs students increased significantly over the past two decades: from 9.3% in 1990 to 17.3% in 2011. This increase is concentrated in the secondary sector, which has seen enrolments in secondary special education rise between 2007 and 2011, despite a decrease in the number of children in primary special education (SBAO) and special education (SO) (Dutch Ministry of Education, Culture and Science 2012).

Educational outcomes

Dutch students have shown strong performance in international student assessments. At the primary level, the Netherlands participates in the IEA's (International Association for the Evaluation of Educational Achievement) Trends in Mathematics and Science Study (TIMSS) and Progress in International Reading Literacy Study (PIRLS). Results from the 2011 TIMSS and PIRLS assessments indicate very good performance for Dutch students aged nine to ten. Among all participating countries¹⁵, the Netherlands was significantly outperformed by only seven countries in mathematics and science, and by nine countries in reading. While the Dutch test scores decreased after the first assessments in mathematics (1995) and reading (2001), they have remained stable since 2006/2007 in reading and mathematics and increased in science (Mullis et al., 2012a; 2012b; Martin et al., 2012).

For the distributions of TIMSS and PIRLS scores, the Netherlands has relatively small variance in comparison to the distributions of other participating countries. The Dutch education system appears to cater particularly well to students at the lower end of the performance distribution. The Netherlands is the only country participating in PIRLS where all students achieved at least the low international benchmark of performance in reading. In addition, 99% of the Dutch students achieve at least the low international benchmark in mathematics and science in TIMSS. The Netherlands also has above average proportions of students achieving the intermediate and high international benchmarks in all three areas assessed, but only between 3% and 7% performing at the advanced international benchmark (Table 1.1). While the proportions of top performers are close to the international averages, they remain far below the participating East Asian countries and some other OECD countries. For example, in mathematics the advanced

international benchmark was reached by 43% of students in Singapore, 30% in Japan, 18% in England and 10% in Belgium (fl.), compared to 5% in the Netherlands (Mullis et al., 2012b).

Table 1.1 Performance of Dutch students at the international benchmarks of achievement in primary education

Percentages of students reaching international benchmarks in TIMSS and PIRLS

International benchmark	Area tested	Netherlands	International median
Low	Reading (PIRLS)	100	95
	Mathematics (TIMSS)	99	90
	Science (TIMSS)	99	92
Intermediate	Reading (PIRLS)	90	80
	Mathematics (TIMSS)	88	69
	Science (TIMSS)	86	72
High	Reading (PIRLS)	48	44
	Mathematics (TIMSS)	44	28
	Science (TIMSS)	37	32
Advanced	Reading (PIRLS)	7	8
	Mathematics (TIMSS)	5	4
	Science (TIMSS)	3	5

Sources: Martin, M. O., et al. (2012), *TIMSS 2011 International Results in Science*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam and TIMSS and PIRLS International Study Center, Boston.

Mullis, I.V.S., et al. (2012a), *PIRLS 2011 International Results in Reading*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam and TIMSS and PIRLS International Study Center, Boston.

Mullis, I.V.S., et al. (2012b), *TIMSS 2011 International Results in Mathematics*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam and TIMSS and PIRLS International Study Center, Boston.

At the secondary level, Dutch students show overall strong performance. In the 2012 round of the OECD's Programme for International Student Assessment (PISA), 15-year-olds in the Netherlands achieved results significantly above the OECD average in the 3 areas tested (mathematics, reading and science). Only two other OECD countries (and six partner economies) achieved significantly higher performance levels in mathematics. Similarly, the Netherlands was significantly outperformed by only five OECD countries (and four partner economies) in reading and by four OECD countries (and three partner economies) in science (OECD, 2014).

The Dutch PISA results also show positive indicators in equity, with fewer low performers and more high performers than the OECD average. Compared to the OECD average, significantly fewer Dutch 15-year-olds showed scores below the PISA performance level 2, believed to mark the basic competency which enables active participation in a society. In mathematics, 15% of Dutch 15-year olds were positioned below level 2, compared to an OECD average of 23%. At the top end of the performance distribution, 19% of the Dutch students reached the PISA performance level 5 and above, compared to 13% at the OECD average (OECD, 2014). The impact of student

socioeconomic background on performance in mathematics was less pronounced than at the OECD average: 12% of the variance in the Dutch performance in mathematics can be explained by socioeconomic background, compared to 15% at the OECD average (OECD, 2013d). The Netherlands also has an above average proportion of “resilient students”, i.e. students who manage to overcome difficult socio-economic circumstances and exceed expectations, when compared with students in other countries¹⁶ (Table 1.2).

However, there are also indications that more can be done to improve both excellence and equity in the Dutch education system. When looking at trends in PISA results over time, the Dutch scores have decreased in mathematics and remained unchanged in reading and science. Between 2003 and 2012, the proportion of top performers in mathematics decreased from 26% to 19% while the proportion of low performers increased from 11% to 15%. There are also marked performance differences across schools in the Netherlands. The performance variance between Dutch schools for mathematics was among the highest of all participating countries at 65%, compared to an OECD average of 37% (OECD, 2013e). This is probably related to the highly tracked education system, with most of the performance variation being explained by the study programme in which a student is enrolled. The Netherlands also has particularly high proportions of grade repetition, with 28% of 15-year-olds reporting that they have repeated a grade at least once, compared to 12% at the OECD average (OECD, 2013e). A closer look at this data reveals that the prevalence of grade repetition is concentrated in primary education, with 20% of 15-year-olds reporting that they repeated a grade in primary education, compared to 7% in lower secondary education (OECD, PISA 2012 Database¹⁷).

In the Netherlands, 11% of the students assessed by PISA 2012 had an immigrant background. These immigrant students tend to be socio-economically disadvantaged in comparison to non-immigrant students. They scored an average of 57 points lower in the PISA mathematics assessment than non-immigrant students, and an average of 35 points lower after accounting for socio-economic differences. This performance difference is significantly larger than on average across OECD countries, where on average immigrant students perform 34 points lower than non-immigrants, and 21 points after accounting for socio-economic differences. In the Netherlands, 29% of students with an immigrant background were low performers (i.e. scoring below the PISA level 2) in mathematics, compared to 12% of the non-immigrant population. Compared to the OECD average, however, the Netherlands has a smaller proportion of immigrant students who are low performers (Table 1.2).

Table 1.2 Selected indicators of quality and equity in Dutch education, based on PISA 2012

	Netherlands	OECD average
% of top performers	Mathematics	19
	Reading	10
	Science	12
% of low performers	Mathematics	15
	Reading	14
	Science	13
% of immigrant students who are low performers	29	36
% of students who repeated a grade	28	12
% of “resilient” students	9	6
% of variance in mathematics performance explained by socio-economic status	12	15

Sources: OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science*, PISA, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264208780-en>.

OECD (2013d), *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed*, PISA, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264201132-en>.

OECD (2013e), *PISA 2012 Results: What Makes a School Successful (Volume IV): Resources, Policies and Practices*, PISA, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264201156-en>.

Notes: Top performers = students performing at PISA level 5 and above; low performers = students performing below PISA level 2; resilient students = students in the bottom quarter of the socio-economic scale within a country or economy who perform among the top 25% of students across all participating countries, after taking their socio-economic status into account.

Main policy developments

Support for school improvement initiatives in priority domains

In 2012, the Ministry of Education, Culture and Science launched a three year programme called *Schools have the initiative (School aan Zet)*. The programme is intended to stimulate school improvement by leveraging internal motivation to increase the effectiveness of the education provided. Under this programme, the Ministry designated six domains of interest and encouraged schools to apply for funding to improve their performance in these domains. The six domains designated by the *Schools have the initiative* programme reflect current national priorities for education:

1. Results-oriented work
2. Human resource management / learning organisation
3. Basic skills
4. Dealing with differences between students
5. Excellence / gifted students
6. Science and technology skills

Participation in the programme is voluntary and begins with schools defining their own goals and ambitions. Schools can use the funding connected to the programme to

hire external consultants (critical friends) from a pool of experts vetted by the Ministry. Together with these external experts, school professionals conduct three “ambition conversations” to set objectives and three “evaluation conversations” to monitor achievements in line with their own expectations. There are no particular reporting requirements connected to the funding allocated through the programme.

The *Schools have the initiative* programme is based on an agreement between the Ministry of Education, Culture and Science and the Primary and Secondary Education Councils (*PO-raad* and *VO-raad*). In this agreement, both partners state their trust in school boards to determine their own strategies for improvement and choose external consultants whose expertise fits their needs. The programme intends to set a framework of national ambitions and priority domains while respecting the principle of freedom of education and ensuring that schools have ownership of their improvement strategies.

The *Schools have the initiative* programme has set participation objectives and performance indicators to monitor the success of the programme. It aims for the participation of 3 000 primary schools, 450 regular secondary schools and 150 secondary schools offering special education. Performance indicators to monitor the effectiveness of the programme are formulated for each of the six domains. For example, in the domain of *results-oriented work*, the programme aims that in 2015 all participating schools score significantly higher on self-selected sub-domains of basic subjects as measured by standardised achievement tests. A programme evaluation has been contracted out to a Dutch research institute (Scheerens, 2013).

A focus on achieving excellence

The Dutch government has launched a range of initiatives to promote excellence in education. This includes a focus on providing better support to gifted and talented students and increasing the proportion of students achieving the highest levels on national and international assessments. To offer guidance regarding the expectations for gifted and talented students, the reference levels in mathematics and Dutch language indicate not only the basic level to be achieved by all students but also an advanced level to be achieved by the country’s top performing students.

As mentioned above, one of the domains for the *Schools have the Initiative* programme is focused on excellence and gifted students. The expectation is that all schools that have selected this domain will identify the top 20% of their students and implement strategies to accelerate their development. Student assessment results are likely to play an important role in the identification process. In several of the schools visited by the OECD review team, interviewees indicated the importance of providing extra attention and support to gifted students to enable them to fully realise their academic potential.

In a paper titled “From good via differentiation to better”, the Inspectorate of Education states its ambition to promote excellence in schools (Dutch Inspectorate of Education, 2013b, cited in Scheerens, 2013). While the Inspectorate plans to continue to work closely with weak and unsatisfactory schools, it aims to pay more attention to schools that are already achieving average or good results. This includes the differentiation of quality indicators in order to raise ambitions in schools and help schools with good results to go further towards excellence. The Inspectorate is also conducting pilot studies to develop “value-added measures”, with a view to obtaining more detailed information regarding the contribution of schools to the learning and growth of their students (Chapter 5).

To stimulate higher levels of achievement, the government has implemented an “excellent school prize”, which is attributed to schools by a committee of experts each year.

Reforms to provide “education that fits” to students with special education needs

At the other end of the spectrum, there is also a recognition that more must be done to support struggling learners and students with special education needs (SEN). There has been a substantial increase in the numbers of students diagnosed as having special educational needs and a concomitant increase in the costs and challenges in providing them with an appropriate education (Dutch Ministry of Education, Culture and Science, 2012). Under the heading of “education that fits”, the Ministry of Education, Culture and Science has developed a policy aimed to better address the educational needs of SEN students.

The OECD review team was informed by the National Knowledge Centre on Special Education (LECSO) that the Netherlands will be divided administratively into 75 educational regions responsible for the organisation of educational services for SEN students. School boards within each region will be expected to work together to set up a legal institution for this purpose. There will not be a common national approach to cater to the needs of SEN students but each region will receive a lump sum to develop the necessary educational services. The regions will be autonomous in determining standards and criteria to assess and identify special educational needs. Regions that are unable to provide adequate services to their SEN students may transfer parts of this funding to another region to take charge of the students concerned.

An ambition to raise teacher professionalism

As a cornerstone to promoting excellence in education, the Dutch government has made the enhancement of teacher professionalism a policy priority for school education (Chapter 4). The 2006 Education Professions Act sets minimum standards for teacher competencies and introduces a requirement for school boards to monitor their teachers’ professional development through teacher competency files. The government’s Action Plan on teacher policy *Teaching 2020: A Strong Profession!* further highlights the importance of increasing teacher professionalism and summarises key measures intended to contribute to this goal, such as: the establishment of a professional registration process, the implementation of enhanced personnel policies and the extension of the Inspectorate’s remit to intensify its focus on teaching quality and teacher professionalism (Dutch Government, 2011). The Action Plan on teacher policy particularly highlights the importance of building teachers’ professional competencies in “results-oriented work” (see below).

Measures to promote “results-oriented work” in schools

Since 2010, the Ministry of Education, Culture and Science and the Inspectorate of Education have been promoting a policy on “results-oriented work” in schools (Chapter 3). Under this heading, they are encouraging schools to develop systematic, goal-oriented processes to maximise student performance. This involves setting learning targets, working with the national reference levels, using various kinds of assessment to measure student progress, providing feedback to students and designing strategies to adapt teaching to student learning needs (Dutch Inspectorate of Education, 2010).

New laws on student assessment

At the time of the OECD review visit in June 2013, important changes to the legislation on student assessment were being prepared. In December 2013, the First Chamber of Parliament accepted a law proposal making it mandatory for primary schools to administrate regular student monitoring systems as well as a final summative test at the end of Year 8. In implementing this policy, schools will be allowed to choose between different tests developed by the Central Institute for Test Development (*Centraal Instituut voor Toetsontwikkeling, Cito*) or other companies, provided that they meet central quality requirements. These laws will be implemented from the 2014/15 school year (Chapter 3).

Notes

- ¹ CBS, Population Counter, available at:
<http://www.cbs.nl/enGB/menu/themas/bevolking/cijfers/extra/bevolkingsteller.htm>.
- ² The World Bank, World Development Indicators, available at:
 at:http://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?order=wbapi_data_value_2012+wbapi_data_value+wbapi_data_value-last&sort=desc.
- ³ CBS, Birth Rate Statistics, available at:
<http://www.cbs.nl/enGB/menu/themas/bevolking/publicaties/artikelen/archief/2013/2013-3863-wm.htm>.
- ⁴ OECD Statistics, Gross domestic product (GDP) per head, US \$, current prices, current PPPs, available at: <http://stats.oecd.org/>.
- ⁵ OECD Statistics, Short-Term Labour Market Statistics, Harmonised Unemployment Rates (HURs), available at: <http://stats.oecd.org/>.
- ⁶ This Review covers primary and secondary education (ISCED 1-3). At the secondary level, it covers VO 1/2, VWO, HAVO and VMBO programmes, but not the first years of HBO and MBO programmes. The review does not cover special education nor the apprenticeship part of pre-vocational education.
- ⁷ Depending on the track, schools can have one, two or three years of shared curriculum (bridge period).
- ⁸ The four learning pathways are: (1) theoretical programme: best suited to pupils who want to continue on to secondary vocational education (MBO) or to senior general secondary education (HAVO); (2) combined programme: offers a mix of theoretical and practical subjects; (3) middle-management vocational programme: tailored to pupils aiming at further vocational training (e.g. as manager of a food franchise operation); and (4) basic vocational programme: a mix between general education and practical, on-the-job experience.
- ⁹ ‘Gymnasia’ offer such pre-university education with a programme including the subjects Latin, Greek and classical culture.
- ¹⁰ The four domains of decision-making comprise the following areas: **Organisation of instruction:** student admissions; student careers; instruction time; choice of textbooks; choice of software/learningware; grouping of students; additional support for students; teaching methods; day-to-day student assessment. **Personnel management:** hiring and dismissal of principals, teaching and non-teaching staff; duties and conditions of service of staff; salary scales of staff; influence over the careers of staff. **Planning and structures:** opening or closure of schools; creation or abolition of a grade level; design of programmes of study; selection of programmes of study taught in a particular school; choice of subjects taught in a particular school; definition of course content; setting of qualifying examinations for a certificate or diploma; accreditation (examination content, marking and administration). **Resource management:** allocation and use of resources for teaching staff, non-teaching staff, capital and operating expenditure, professional development of principals and teachers.

- ¹¹ Internal supervisors are laypersons from the school's community appointed through co-optation.
- ¹² Primary Education Council, Figures from the Sector, available at: <http://www.poraad.nl/content/cijfers-uit-de-sector>.
- ¹³ The relationship between the Inspectorate and the Ministry of Education is described in the 2006 Regulation on Educational Inspection, available at: http://wetten.overheid.nl/BWBR0019615/geldigheidsdatum_07-12-2009.
- ¹⁴ The tasks of the Inspectorate are stipulated by the 2012 Education Supervision Act (*Wet op het onderwijstoezicht, WOT*), available at: http://wetten.overheid.nl/BWBR0013800/geldigheidsdatum_04-03-2014.
- ¹⁵ In 2011, 52 education systems and 7 benchmarking participants took part in the 4th grade assessment of TIMSS and 49 education systems and 9 benchmarking participants took part in PIRLS.
- ¹⁶ Resilient students are defined by the OECD (2013e) as disadvantaged students (those in the bottom quarter of the socio-economic scale within a country or economy) who perform among the top 25% of students across all participating countries, after taking their socio-economic status into account.
- ¹⁷ OECD, PISA 2012 Database, Interactive Data Selection, available at : <http://pisa2012.acer.edu.au/interactive.php> .

References

- Bal, J. and De Jonge, J. (2007), *Improving School Leadership: Country Background Report for the Netherlands*, available at: www.oecd.org/education/school/38639469.pdf.
- Dutch Government (2011), *Teaching 2020, A Strong Profession!*, Dutch Centre for Expertise in Vocational Education and Training, www.ecbo.nl/ECBO/ReferNet/docs/11-0315_Teacher_2020.pdf.
- Dutch Inspectorate of Education (2010), *Opbrengstgericht werken in het basisonderwijs: Een onderzoek naar opbrengstgericht werken bij rekenen-wiskunde in het basisonderwijs*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2013a) *The State of Education in the Netherlands, Summary of the 2011/2012 Education Report*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2013b), *Van Goed, via Differentiatie, naar Beter*, Ambitie-paper van de Inspectie van het Onderwijs, 2013-2017.
- Dutch Ministry of Education, Culture and Science (2012), *Key figures 2007-2011*, Dutch Ministry of Education Culture, and Science, The Hague.
- Hooge, E.H., & Honingh, M.E. (2014), “Are school boards aware of the educational quality of their schools?” *Educational Management Administration & Leadership*, published online before print January 27, 2014, Sage Publications.
- Martin, M. O., et al. (2012), *TIMSS 2011 International Results in Science*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam and TIMSS and PIRLS International Study Center, Boston.
- Mullis, I.V.S., et al. (2012a), *PIRLS 2011 International Results in Reading*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam and TIMSS and PIRLS International Study Center, Boston.
- Mullis, I.V.S., et al. (2012b), *TIMSS 2011 International Results in Mathematics*, International Association for the Evaluation of Educational Achievement (IEA), Amsterdam and TIMSS and PIRLS International Study Center, Boston.
- OECD (2014), *PISA 2012 Results: What Students Know and Can Do (Volume I, Revised edition, February 2014): Student Performance in Mathematics, Reading and Science*, PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264208780-en>.
- OECD (2013a), *International Migration Outlook 2013*, OECD Publishing, Paris, http://dx.doi.org/10.1787/migr_outlook-2013-en.
- OECD (2013b), *Government at a Glance 2013, Country Fact Sheet: Netherlands*, OECD, Paris, www.oecd.org/gov/GAAG2013_CFS_NLD.pdf.

- OECD (2013c), *Education at a Glance 2013: OECD Indicators*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/eag-2013-en>.
- OECD (2013d), *PISA 2012 Results: Excellence through Equity (Volume II): Giving Every Student the Chance to Succeed*, PISA, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264201132-en>.
- OECD (2013e), *PISA 2012 Results: What Makes a School Successful (Volume IV): Resources, Policies and Practices*, PISA, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264201156-en>.
- OECD (2012), *Education at a Glance 2012: OECD Indicators*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/eag-2012-en>.
- OECD, PISA 2012 Database, Interactive Data Selection, available at: <http://pisa2012.acer.edu.au/interactive.php>.
- Scheerens, J. (2013), “Educational Evaluation and Assessment in the Netherlands”, Addendum to the *Country Background Report for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes*, www.oecd.org/education/school/Netherlands%20CBR%20Update.pdf.

Chapter 2

The evaluation and assessment framework

Internationally, the Netherlands stands out for its well-developed evaluation and assessment approaches. Central mechanisms for student assessment, school evaluation and education system evaluation have been in place for several decades, along with requirements for schools to assure their own quality. The Dutch evaluation and assessment system combines a high degree of school autonomy with a set of checks and balances that allow for intervention if schools are found to be at risk of underperformance. The Netherlands is characterised by a complex system of governance with multiple actors shaping evaluation and assessment practices. This multi-actor environment has resulted in a comprehensive and balanced approach to evaluation and assessment, which combines school-based and central elements, quantitative and qualitative approaches, improvement and accountability functions, and vertical and horizontal responsibilities of schools. To further consolidate the evaluation and assessment framework, there is room to embed evaluation and assessment approaches with broader education goals, continue to build on teacher professionalism, and build networks and capacity for effective and forward-looking evaluation and assessment, in particular in the areas of classroom-based student assessment, teacher appraisal and school-self-evaluation.

This chapter looks at the overall framework for evaluation and assessment in the Netherlands, i.e. its various components such as: student assessment, teacher appraisal, school evaluation and system evaluation, the coherence of the whole as well as the articulation among the different components. Following this overview, the succeeding chapters (3-6) will analyse the issues regarding each individual component in more depth.

This report differentiates between the terms “assessment”, “appraisal” and “evaluation”. The term “assessment” is used to refer to judgments on individual student performance and achievement of learning goals. It covers classroom-based assessments as well as large-scale, external tests and examinations. The term “appraisal” is used to refer to judgements on the performance of school-level professionals, i.e. teachers and school leaders. The term “evaluation” is used to refer to judgments on the effectiveness of schools, school systems and policies. This includes school inspections, school self-evaluations, evaluation of school boards, system evaluation and targeted programme evaluations.

Context and features

Key components of the evaluation and assessment framework

The Ministry of Education, Culture and Science has not developed an overall policy or framework for evaluation and assessment, but it sets priorities for policy development through its *Quality Agendas* and *Action Plans*. Current priorities include the enhancement of results-oriented work and teacher professionalism. The Dutch approach consists of the following four components:

- **Student assessment.** Student assessment in the Netherlands is largely the responsibility of schools and classroom teachers, supported by well-developed standardised assessment tools. For formative assessment, nearly all primary schools participate in the LVS (*Leerling Volg Systeem*), a longitudinal student monitoring system developed by the Central Institute for Test Development (*Centraal Instituut voor Toetsontwikkeling*, Cito) covering most subjects. For the first two years of secondary education, there are also a range of monitoring assessments offered by Cito. In their regular classroom assessment, teachers choose assessment tools, typically drawn from tests provided by the particular textbooks or “methods” they use. Summative results are reported to students and parents three times a year on a scale of one to ten. At the end of primary education, schools are required to report on the extent to which their students have reached expected core learning objectives. While schools are free to use different instruments for this purpose, 85% of schools use Cito’s end-of-primary test, which provides information on the school type most suitable for each student in the next phase of education. New laws that will be implemented from the 2014/15 school year make it mandatory for primary schools to administrate regular student monitoring systems as well as a final test at the end of Year 8. Schools can choose to administer the tests developed by Cito or alternative tests provided that they meet central quality requirements. In the secondary sector, school-leaving examinations are administrated at the end of each track. These examinations consist of a central part developed by Cito following guidance from the College for Examinations (*College voor Examen*, CVE), and a school-based part developed by teachers in conformance with the examination syllabuses.

- **Teacher appraisal.** Teacher appraisal is under the jurisdiction of the competent authority of each school. The 2006 Education Professions Act requires school boards to establish human resource policies for their schools, keep competency files for each teacher and ensure that teachers' competencies are maintained. Central regulations specify that schools should have regular performance interviews with all staff, but there is little central guidance on how the performance of individual teachers should be evaluated. As the employing authorities for teachers, school boards are free to develop their own frameworks for teacher appraisal. Many school boards delegate the responsibility for human resource management, including teacher appraisal, to the school leaders, and practices vary from school to school.
- **School evaluation.** There is no legal obligation for schools to implement particular self-evaluation processes, but schools are required to draw up a school prospectus, an annual report and a four-year school plan, which is typically based on an internal review of school quality. Schools benefit from analytical software systems and benchmarked data, and can choose to buy supporting materials and services from different providers. External evaluations are conducted by the Inspectorate at least every four years, with the type and intensity of inspection depending on identified risks in each school. Schools that are considered at risk of underperformance are evaluated more frequently and more thoroughly than others. The initial risk analysis is based on a review of each school's outcomes, annual accounts and "failure signals" such as complaints. For its inspection visits, the Inspectorate uses a detailed framework of quality indicators and a clear set of decision rules. As part of this framework, the Inspectorate of Education also evaluates the internal quality care undertaken by schools. A range of different databases providing school-level information and performance indicators are connected and made available to different audiences through the online information systems Windows for Accountability (*Vensters voor Verantwoording*) and Schools on the Map (*Scholen op de kaart*). School evaluation also uses information recorded in the Basic Register for Education (*Basis Register Onderwijs*, BRON). The Inspectorate publishes its Inspection Reports, as well as School Quality Cards which provide information about the inspection regime schools are assigned to.
- **System evaluation.** The overall performance of the Dutch education system is monitored in several ways. Information on student learning outcomes is collected from international surveys, national monitoring sample surveys (*Periodiek Peilings Onderzoek*, PPO and *Jaarlijks Peilingsonderzoek van het Onderwijsniveau*, JPO), the longitudinal Cohort Survey School Careers (*Cohort Onderzoek Onderwijsloopbanen*, COOL), the standardised test results reported by schools (e.g. results from the LVS or the Cito school leavers test) and the results from the secondary school-leaving examinations. This student performance data is complemented by a wide range of demographic, administrative and contextual data collected by the Ministry of Education, Science and Culture directly from schools. System evaluation also makes use of BRON register data, based on the unique student number. The Inspectorate of Education analyses education system performance based on information aggregated from its inspection activities and undertakes thematic inspections on particular priority themes in samples of schools. In addition, the Ministry regularly commissions a variety of research groups to conduct research studies and programme evaluations. Information on

education system performance is published in several formats, including the Ministry's publications on *Key Figures and Trends in the Picture* and the Inspectorate's *State of Education* report.

Table 2.1 Main instruments used for evaluation and assessment in the Dutch school system (2013)

Level of evaluation	Reference standards	Approaches / Instruments	Reporting
Student	<ul style="list-style-type: none"> Act on Primary Education and Act on Secondary Education (including core learning objectives) Language and Numeracy Act (including reference levels for literacy and numeracy) 	<ul style="list-style-type: none"> School-based assessment including results-oriented work LVS and other student monitoring tools used for formative purposes CITO school leavers test and other examinations used in Year 8 Examinations at the end of each secondary track 	<ul style="list-style-type: none"> Summative reporting to students and parents three times a year on a scale from 1-10 LVS <i>pupil reports</i> and <i>group surveys</i> CITO school leavers test results and advice on the secondary track most suitable for each student Student certificates at the end of secondary school including pass/fail information and marks (1-10) for each subject School-level and national reporting of aggregated student assessment results
Teacher	<ul style="list-style-type: none"> Competency requirements included in Education Professions Act (2006) Inspectorate's indicators for lesson observation 	<ul style="list-style-type: none"> School-based teacher appraisal Voluntary teacher registration 	<ul style="list-style-type: none"> Teacher competency files Teacher register School-level and national reporting on teaching quality
School	<ul style="list-style-type: none"> School goals and strategic plans Inspectorate's Supervision Framework 	<ul style="list-style-type: none"> School internal quality care Quality control by school boards Risk-based inspection 	<ul style="list-style-type: none"> School annual report, school plan and school prospectus Inspection reports School Quality Cards Windows for Accountability website National reporting on school quality
System	<ul style="list-style-type: none"> Act on Primary Education and Act on Secondary Education Language and Numeracy Act Specific policy and programme objectives European Union benchmarks for education (ET 2020) 	<ul style="list-style-type: none"> National sample-based assessments (PPON, JPON) National cohort assessments (COOL) International surveys (TIMSS, PIRLS, PISA, ICILS, ICCS, TALIS) Collection of indicators Inspectorate thematic evaluations Programme evaluations 	<ul style="list-style-type: none"> <i>Key Figures</i> report <i>Trends in the Picture</i> report <i>State of Education</i> report Ministry "Monitors" on specific themes

Source: Authors' own work, based on Scheerens, J., et al. (2012), OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.

Governance

The Dutch approach to evaluation and assessment comprises both central and school-based mechanisms. The central legal framework defines core learning objectives and sets a range of evaluation and assessment requirements for schools. The most important central instruments for evaluation and assessment are the secondary school examinations overseen by the College for Examinations, the external evaluation of individual schools by the Dutch Inspectorate of Education, and the instruments implemented by the Ministry of Education, Culture and Science and the Inspectorate to monitor the performance of the education system. With the implementation of the new laws on student assessment (Chapter 1), the College for Examinations will also oversee the quality of end-of-primary tests administered by schools. The responsibilities of schools and school boards include regular formative and summative assessment of students, performance evaluations of individual teachers, and school-internal quality care.

The Netherlands is characterised by a complex system of governance with multiple actors shaping evaluation and assessment practices. In addition to the three traditional actors (the Ministry, the Inspectorate and the schools) a range of intermediate organisations have an increasing influence on evaluation and assessment in the country. The Primary and Secondary Education Councils play a key role in the collection and dissemination of data on the performance of individual schools, thereby strengthening both horizontal and vertical accountability. The unions and professional organisations of teachers and school leaders, in particular the Education Cooperative (*Onderwijscoöperatie*) created in 2011, have become important actors in supporting the professionalism of school staff. They play a key role in the development of registration, review and feedback systems for teachers and school leaders. Finally, the Education Council, an independent advisory body funded by the government and composed of ten experts, provides independent advice to the government on national education issues, including many topics related to evaluation and assessment.

There is a well-developed infrastructure of private educational providers offering instruments, information systems and support services for evaluation and assessment to schools. The most prominent assessment provider is the Central Institute for Test Development (Cito), which develops standardised tests that are used by the vast majority of Dutch schools for both student assessment and school evaluation. There are also a range of smaller organisations developing tools to support schools in their evaluation and assessment activities. The largest providers of educational support services are the Educational Service Centres, which were privatised and transformed into not-for-profit foundations in 2008. Most educational service providers form part of the umbrella organisation *Edventure*.

Strengths

A long-standing tradition of evaluation and assessment in education

Internationally, the Netherlands stands out for its long history of well-developed evaluation and assessment systems. Central mechanisms for student assessment, school evaluation and education system evaluation have been in place for several decades, along with requirements for schools to assure their own quality. This long-standing focus on developing reliable evaluation and assessment systems has allowed the central agencies to experiment with different approaches and develop and deepen their expertise over time.

Developing reliable measures of student learning outcomes is a priority in assessment frameworks around the world (OECD, 2013). In the Netherlands, the national assessments and examinations currently used by most schools have existed for over 30 years. Cito, created in 1968, was one of the first influential national testing institutes within the OECD and enjoys a high international reputation for its expertise and assessment instruments (Chapter 3). While there are different views as to how the results should be used, the usefulness of standardised assessments to inform certification, accountability and improvement strategies is widely accepted among stakeholders.

The systematic external evaluation of schools has an even longer history in the Netherlands. The Inspectorate of Education, founded in 1801, was one of the earliest Education Inspectorates in the OECD area. As described in Chapter 5, it is a highly structured and professional organisation whose transparent standards and systematic evaluation methods have not only shaped the school evaluation culture in the Netherlands, but also influenced other educational inspectorates in Europe. The Inspectorate emphasises a research-based approach and has repeatedly reinvented its inspection approaches to respond to emerging challenges. Over the past two decades, there has been considerable investment in building each school's capacity for self-evaluation, although this has lessened in recent years (Chapter 5).

The Netherlands has a well-developed structure of research institutes specialising in educational analysis. Programme evaluations have been undertaken in various forms over the past four decades, typically promoted by evaluation committees of the Ministry of Education, which commissions studies to different research groups. The Inspectorate itself also conducts thematic reviews on priority topics and has been involved in coordinating major policy evaluations of primary and lower secondary education. Ensuring that education policy is evidence-based is a key priority of the Ministry. This is reflected by the existence of a dedicated "Knowledge Directorate", responsible for the Ministry's strategic approach to research, evaluation and information management (Chapter 6). The Ministry has set an agenda for effective evaluations of educational policy interventions called *Outlook on Effectiveness* (Scheerens, 2013).

Shared responsibilities

The Dutch evaluation and assessment system combines a high degree of school autonomy with a complex set of checks and balances that allow for intervention if schools are found to be at risk of underperformance. Decisions are typically made closest to where they will have an effect, and the central government takes a subsidiary function performing mainly those tasks that cannot be as effectively dealt with at the school level.

Schools have broad autonomy in the area of evaluation and assessment. They are responsible for choosing approaches to student assessment (except for the central part of secondary school leaving examinations), designing and implementing teacher appraisal processes, conducting their own quality care, and acting upon feedback by the Inspectorate. There is considerable attention on reducing administrative burdens on schools and limiting reporting requirements that generate paperwork but have little bearing on actual practice. As described by Scheerens (2013), there is a tradition of "bottom-up reforms", where schools are offered incentives to join networks exploring new practices, rather than having changes imposed from the top down. Recent policy initiatives continue to emphasise the prime role of schools in their own evaluation and improvement processes. The *Windows for Accountability* (*Vensters voor Verantwoording*) website enhances schools' ownership of their data, allowing them to

enter their own explanation and interpretation of the information provided. The *Schools have the Initiative (School aan Zet)* programme further emphasises the schools' lead role in determining improvement strategies.

While schools have the main responsibility for evaluation and assessment, there are a range of mechanisms in place to ensure that minimum requirements are met in all schools. As some of the stakeholders interviewed by the OECD review team said, the idea is that there is always a “Plan B” in case a school's own quality assurance mechanisms do not work in a satisfying way. The involvement of numerous stakeholders results in a multiplication of evaluation and assessment mechanisms. For example, schools report information to their board, to the *Windows for Accountability* website, to the Inspectorate, and to the Ministry of Education. They receive evaluative feedback from the school boards, the Inspectorate, external consultants of the *Schools have the Initiative* project (for participating schools) and, in many cases, other experts and service organisations hired to support internal quality care and school development.

While there is a risk that these multiple checks and initiatives may lead to inefficiencies and duplication (this challenge will be addressed below), it can be argued that this approach minimises the risk of schools underperforming consistently without weaknesses being picked up by one or the other organisation. It can also be argued that the evaluation approaches of each of the involved groups are slightly different in focus and emphasis, hence justifying the multiplication of initiatives (Scheerens, 2013). This approach is in line with the Dutch tradition of multi-stakeholder involvement and may be described as “strategic redundancy”. It also echoes elements of a governance system often described as the “Polder model”, where each stakeholder group is given room to play its role, and channels for debate and negotiation are kept open, even in periods of major disagreements.

A focus on achieving excellence through differentiation

As part of a general policy emphasising quality and excellence in education, the Dutch government aims to increase the focus of schools on raising student achievement and to encourage school professionals to strive towards excellence (Scheerens et al., 2012). As outlined in Chapter 3, a policy focussed on excellence faces the challenge that it must set high standards for achievement, which not everyone can reach, or at least not at the same speed. In this context, using differentiated approaches to evaluation and assessment can help to take contextual differences into account and provide the right level of support and challenge to those who are evaluated, be it students, teachers or schools.

There are a range of recent evaluation and assessment initiatives in the Netherlands that can be described as pursuing excellence through differentiation. At the level of individual students, the focus on results-oriented education reflects heightened attention to monitoring student progress and providing differentiated feedback for improvement (Chapter 3). At the level of schools, the Ministry of Education, Culture and Science has implemented a prize for excellent schools, and the Inspectorate is developing a differentiated inspection approach with the intention of helping schools that already provide basic quality education to further improve towards excellence (Chapter 5).

A balanced approach to evaluation and assessment

Given the multi-actor environment of Dutch education, the evaluation and assessment system has been shaped by continuous negotiation and debate among different groups promoting a range of competing approaches to evaluation and assessment. This

appears to have resulted in a comprehensive and balanced approach to evaluation and assessment, which combines school-based and central elements, quantitative and qualitative approaches, improvement and accountability functions, and vertical and horizontal responsibilities of schools.

School-based and central elements

There is a clear understanding that shared decision-making and buy-in from schools are essential for successful evaluation and assessment policies. For all key components of the evaluation and assessment system, the Dutch approach combines centralised and school-based elements. This combination appears to have contributed to the development of an advanced evaluation and assessment culture. This has also been supported by the opportunity for schools to draw from tools and expertise offered by a variety of service providers.

Balancing school-based and central elements of evaluation and assessment allows education systems to combine a degree of national consistency along with local diversity. Consistency across schools is important to implement national agendas for improving education, but greater diversity of approaches offers more opportunities for innovation and adaptation to local needs. While standardised evaluation and assessment approaches help provide a reliable external benchmark for schools, authentic, ongoing and school-based approaches are more likely to cover a broad range of quality aspects (OECD, 2013). The balance between school-based and external elements can contribute to strengthening both validity and reliability of measurements, and to hold schools accountable for effective quality assurance, without stifling the commitment and innovation potential of local actors.

Quantitative and qualitative evaluation

Compared to other European education systems, the Dutch approach to evaluation and assessment has a strong quantitative focus. At the same time, and in line with the current focus on promoting excellence and professionalism in schools, there is increasing attention to evaluating the processes and practices in schools that are likely to influence teaching and learning outcomes. It is important to place such qualitative elements alongside quantitative data, to offer a more comprehensive interpretation of what constitutes quality at school level, taking into consideration the school's particular context and circumstances (Santiago et al., 2012).

The availability and richness of quantitative data on pupil performance is an important characteristic of the Dutch evaluation and assessment framework. Schools are expected to use standardised assessment instruments for both formative and summative assessment in order to yield reliable and comparable information on the extent to which national learning objectives are being achieved. The importance of using data from student assessments for accountability and improvement is clearly promoted by current government policy, such as the focus on results-oriented work and the new requirement for schools to use student monitoring systems (Chapter 3). The risk-based approach of the Inspectorate also relies strongly on quantitative indicators, as the type of supervision schools will receive is based primarily on an analysis of school performance and financial data. Also, the emerging approaches to teacher appraisal for accountability are focussed primarily on measurable aspects of teacher quality. The registration system as it is currently conceived does not include a qualitative evaluation of teachers' actual practices,

but it is based on counting formal qualifications and numbers of hours of professional development undertaken (Chapter 4).

On the other hand, there have always been counterbalancing tendencies that promote the use of more qualitative evaluation of school processes. Particularly throughout the early 2000s, there have been major investments in building schools' capacity for evaluating their own processes and outcomes (Chapter 5). While formal school self-evaluation is not a requirement, schools have clear reporting requirements and many schools involve external consultants to help them review their practices. The government's recent (voluntary) school improvement programme *Schools have the Initiative* also comprises elements of qualitative evaluation. As part of the programme, schools receive funding to draw on the services of external experts to support them in school improvement. The priority domains of this programme include process elements such as "human resource management/learning organisation". The external experts typically have backgrounds in either subject matter didactics or organisational management. Their role is to help schools review their own practice, provide external feedback and support schools in developing and implementing improvement strategies.

The Inspectorate's move towards more differentiated inspections and its focus on promoting excellence in schooling is likely to further increase attention on qualitative evaluations. As the Inspectorate will begin to focus more on schools whose students are already achieving average or good results, evaluators will be required to provide more differentiated feedback regarding the processes employed by schools at different stages of development (Chapter 5). In the Inspectorate's Ambition Paper *From Good via Differentiation to Better*, the Inspectorate explains the characteristics of schools at different levels of development (Dutch Inspectorate of Education, 2013b, cited in Scheerens, 2013). Process indicators feature prominently among these, for example leadership, self-regulation, professional culture, teacher peer review and coaching. This suggests that achieving excellence in schooling is conceived by the Inspectorate as partly a matter of improving these organisational dimensions (Scheerens, 2013).

The government's intention to promote teacher professionalism as a key policy priority is well in line with this development. For the Inspectorate, the focus on strengthening teacher professionalism means that it will need to evaluate the degree to which schools foster the development of professional learning communities and use pedagogical and organisational innovations to enhance teacher learning. This is likely to involve a shift in the balance between the current data-driven inspection approach to more qualitative evaluation approaches that focus on harder to measure aspects of quality (Chapter 5). The aim to strengthen teacher professionalism through observation and evaluation of teaching is pursued by the teacher peer review project implemented by the Education Cooperative. This project aims to support mutual observation among teachers in order to identify strengths and challenges in each teacher's practice and to provide qualitative feedback for improvement (Chapter 4).

Improvement and accountability functions

An important priority among OECD countries is to find the right balance between the accountability and improvement functions of evaluation and assessment. Some countries emphasise one function over the other. Countries with a strong focus on accountability typically emphasise: high-stakes standardised assessment of students; teacher appraisal that is linked to decisions regarding career advancement, salary, promotion and dismissal; external reviews or inspections of school quality; publication of school evaluation results

and/or public comparisons of school performance. Countries with a strong focus on development and improvement typically emphasise: formative, low-stakes assessment of students; teacher appraisal that is linked to decisions regarding teacher professional development and learning opportunities; school self-evaluation and external support for organisational learning (OECD, 2013).

With its multiple approaches to evaluation and assessment, the Netherlands appears to strike a relatively good balance between accountability-oriented and development-oriented elements. Student summative assessment results are reported as part of school accountability and system monitoring. At the same time, results from student monitoring systems can be used by students, teachers and school leaders for the improvement of teaching and learning approaches (Chapter 3). At the whole-school level, inspections have a clear focus on accountability and have high stakes for individual schools, including the threat of sanctions for schools that do not improve. However, the Inspectorate also has the mandate to “stimulate improvement” in schools, and a range of other support organisations exercise a developmental function, working with schools to review practices and plan for school development (Chapter 5). The *Schools have the Initiative* programme of the government further enhances the school development and improvement function.

Vertical and horizontal responsibilities of schools

The Dutch evaluation and assessment approach balances different types of school accountability. While vertical accountability of schools towards the Inspectorate has a long tradition in the Netherlands, there has been increased attention to combining this with enhanced horizontal accountability towards the school’s community. The particular responsibility of school boards for internal supervision, quality control and horizontal accountability is emphasised in the 2010 legislation on *Good Education and Good Governance*, part of the Law on Primary and Secondary Education.

Through horizontal accountability, schools provide different stakeholders groups such as parents, feeder schools, local business and social service providers with insights into their processes, choices and results (Faubert, 2009). Horizontal accountability not only serves to enhance transparency about the performance of schools, but the results may also be used by stakeholders to take action, as with parental school choice. Regular interactions and information sharing with local stakeholders also contribute, over time, to building trust and raising awareness for the main concerns of others, thereby enhancing the possibilities for collaboration to support the common goal of enhancing student learning (OECD, 2013).

The Dutch focus on combining vertical and horizontal accountability is best reflected in the recent development of the *Windows for Accountability* online database, which is based on collaboration between the Primary and Secondary Education Councils, the Ministry of Education, the Inspectorate and the schools themselves. This system includes statistical information provided by the central agencies, but it also supports schools in creating their own school-based indicators that are relevant for horizontal accountability, for example information on students’ and parents’ satisfaction with the school. The website comprises a specific interface for parents, the *School Choice Window*, which facilitates the comparison between schools in relation to indicators of interest for school choice (Chapter 5). According to the Secondary Education Council (2011), the purpose of sharing school-level information horizontally with stakeholder groups is the predominant

purpose of the website, while the use of this information by the Inspectorate in its risk analysis is a secondary function.

Challenges

Setting learning goals for the mid-21st century

Across all OECD countries, there is a great deal of discussion about the introduction of so-called 21st century skills into the curriculum, as they are seen as a necessary component of preparation to be economically productive in a global economy (NRC, 2012; OECD, 2013). While there does not exist a consensus on what these skills comprise, most definitions describe a focus on broader learning that comprises a complex integration of knowledge, skills, attitudes and action in order to carry out tasks in real-life contexts and acquire competencies for lifelong learning (OECD, 2013).

Based on an extensive review of the literature, the National Research Council in the United States argued that competencies such as analytic thinking, self-regulation, and productive collaboration, explain a substantial proportion of the variation in valued outcomes in school and in the workplace, even after taking account of the core skills of literacy and numeracy (NRC, 2012). The importance of these skills will only increase with advances in technology and globalisation (Kirsch et al., 2005; Goldin & Katz, 2008, OECD, 2013). With the 2006 *Recommendation on Key Competences*, all EU member states have agreed on a framework of eight key competencies that are seen as necessary for personal fulfilment and development (European Commission, 2011).

There has been a longstanding effort in the Netherlands to introduce 21st century skills into national learning objectives. Under the rubric of ‘general skills’, they have been regarded as part of the core curriculum in the primary sector and the vocational education program. In the early 2000s, innovations in the secondary education curriculum under the title ‘new learning’ also included aspects of 21st century skills, including learning to learn, self-regulated learning, and collaborative problem-solving (Scheerens et al, 2012; Adamson & Darling-Hammond, 2013).

However, an evaluation of the ‘new learning’ initiative in the secondary sector, undertaken by the Parliamentary Committee Educational Innovation (referred to as the “Dijsselbloem Committee” after its chairman) and published in 2008, revealed substantial confusion in the field, as well as some resentment of the top-down nature of the innovations (Scheerens et al, 2012). Although assessment of some of these skills has been incorporated into class assignments and school examinations, results are not reported separately so there is little comparable information on student competencies (Adamson & Darling-Hammond, 2013).

Several groups interviewed by the OECD review team voiced concerns about limited central support and leadership regarding the inclusion of broader competencies in curriculum, teaching and evaluation and assessment. Among others, the Education Council advocates the need to re-focus the education system on broader learning goals, including areas such as citizenship skills, creative thinking and collaboration. Defining such learning goals for the 21st century would allow key stakeholders to engage in reflection and dialogue on how evaluation and assessment should evolve in order to support a future-oriented education system (Chapter 3).

Reflection about national learning goals for 21st century skills needs to take into account the Dutch curriculum landscape. The Netherlands does not have a unified

curriculum to provide clear references for teaching and learning, and evaluation and assessment in education. National “core objectives” specify the knowledge and skills that students should achieve at the end of primary and secondary education, but they are defined in very general terms. Additional reference levels describe in more detail the literacy and numeracy skills to be achieved at the end of each educational stage¹, but they cover only a small subset of student learning objectives. Some of our interviewees expressed concerns that the reference levels also remained too broad and vague to provide clear guidance for evaluation and assessment practices.

Based on the national core objectives and reference levels, a variety of textbooks are available on the market, providing teachers with a choice of different “methods” to organise their subject teaching. The General Institute for Curriculum Development (*Instituut voor Leerplan-Ontwikkeling, SLO*) runs a website providing an overview of available textbooks and teaching materials, but it does not accredit the providers or evaluate the quality of textbooks. The implemented curriculum is decided in the teaching programmes developed by individual schools and teachers. At the school level, teachers in charge of a particular subject typically take a collective decision on the textbook or method to be used for the subject. As a result, there are risks that teaching remains fragmented along subject lines, with cross-curricular skills receiving less attention. These curricular arrangements make it difficult to reach a sense of common purpose and a coherent overall vision of the type of learning to be achieved, both at the school level and the national level.

In the absence of clear and specific central expectations, learning goals are being set with reference to existing assessments, in particular the Cito end-of-primary test and the national examinations at the end of secondary education. Although this approach has some merits, the current examinations do not fully represent the target content, so excessive focus on these can lead to unintended consequences, such as curriculum narrowing (Chapter 3). Research indicates that where standards are poorly designed or not very specific, teachers are more likely to focus on standardised tests, thus narrowing the focus of teaching (Stecher et al., 2000).

There is room to further promote teacher professionalism in evaluation and assessment

The commitment of the government to boost teacher professionalism strongly determines the context in which evaluation and assessment should be conceived in the future. Strengthening teacher professionalism is potentially the most important contextual factor for evaluation and assessment, and there is room to further promote it as a central element in the Dutch evaluation and assessment framework.

Concerning student assessment, although there is a good balance between school-based and standardised student assessments, there is a challenge in the Netherlands to build the competencies of all teachers to fully exploit the potential of assessment to transform and improve classroom practices. Offering opportunities for teachers to build their assessment skills, get more involved in creating their own targeted assessments, evaluate students’ responses in relation to national expectations and reflect on results to adjust teaching processes would be a powerful tool for enhancing teachers’ professionalism and sharpening a common understanding of expected learning goals and quality criteria. Relying primarily on external tests is problematic when these only measure a sub-set of learning goals and do not fully address the information sought by the teacher to identify student learning difficulties. Research also indicates that the data

obtained from such tests is typically not at the level of detail needed to diagnose individual student needs and identify strategies for future instruction (Chapter 3).

In addition, without an increased focus on implementing developmental teacher appraisal in all schools it will be difficult to ensure that all teachers receive adequate support and recognition as professionals. Teachers need feedback on their performance to help them identify how to better shape their teaching practice and, with the support of effective school leadership, to develop schools as professional learning communities. Teacher appraisal, as part of the registration system, could also play a bigger role in recognising and rewarding high quality teaching. It could help facilitate the organisation of schools in ways that are sensitive to individual talent, performance and motivation by allowing teachers to progress in their career and take on new roles and responsibilities based on valid appraisals of their performance (Chapter 4).

Finally, the emphasis on teacher professionalism creates new challenges for the design of school evaluation. The Inspectorate will need to develop new approaches to evaluate the internal capacity of schools foster teacher professionalism, manage human resources and create professional learning communities. In line with the government's focus on promoting excellence, the Inspectorate will need to cover a wider range of quality aspects, especially in those schools that are already achieving acceptable results. This would include identifying and promoting innovation in teaching, and engaging in mutual learning with the country's outstanding teachers and school leaders. In addition, as full quality inspections will continue to be conducted in only a small number of schools every year, there is a need to foster the capacity for schools to conduct their self-evaluations on an ongoing basis, with teachers being fully involved (Chapter 5).

There are some incipient elements and missing links in the evaluation and assessment framework

Several stakeholder groups, including the professional organisations of teachers and school leaders, voiced concerns about the absence of a common vision for schooling in the Netherlands. Evidently, a national consensus on educational goals and priorities is yet to emerge (see above). By extension, there is currently no articulated overarching vision or strategy for how evaluation and assessment should fit within broader educational improvement strategies, and which role they should play in achieving system goals.

While there are a set of central regulations and a well-developed market of sophisticated evaluation and assessment instruments in the Netherlands, the diverse approaches and initiatives are not currently conceived as a coherent whole. The Ministry of Education, Culture and Science regularly launches new priorities or initiatives related to evaluation and assessment, such as the current focus on results-oriented work, the attention to teacher professionalism and the *Schools have the Initiative* programme. However, the national authorities typically do not provide a narrative of how such policy priorities fit together into a coherent plan, or how they are expected to interplay with other parts of the evaluation and assessment framework.

The *OECD Reviews of Evaluation and Assessment in Education* have revealed the challenges faced by all education systems in designing a comprehensive and coherent evaluation and assessment framework. In many systems the different components of evaluation and assessment have developed independently and at different paces, with the result that some components remain incipient and/or not well-connected to other components (OECD, 2013). In the Netherlands, most elements of evaluation and

assessment are well developed. The following areas were identified by the OECD review team as priorities for future policy development.

- **Formative assessment of students.** With its emphasis on results-oriented work and the new laws making the use of student monitoring systems mandatory, the Netherlands has a firm foundation for effective formative assessment. To fully exploit the potential of student monitoring systems, teachers need to further develop skills to interpret results, to understand whether further diagnostic testing of some students may be warranted, and to identify areas where curricular strategies may need adjustment, or where they may invest resources in new programmes to meet student needs. In addition, while the regular use of student monitoring systems is important for identifying areas of need and developing broad teaching strategies, there is evidence that ongoing classroom-based formative assessment and differentiated instruction are key areas deserving further attention in the Netherlands. This includes skills for setting up learning situations, developing sophisticated questions, providing effective feedback and adjusting teaching to meet identified needs. A focus on formative assessment also implies placing students at the centre of the evaluation framework and ensuring that they are engaged in their own learning and assessment (Chapter 3).
- **Teacher appraisal.** As in many other European countries, teacher appraisal is the least developed component of evaluation and assessment in the Netherlands. Teacher appraisal is under the responsibility of each school board and practices vary from school to school. As a result not all teachers receive regular professional feedback or recognition of their competencies (Chapter 4).
- **School self-evaluation.** Although there is no legal requirement for particular school self-evaluation processes in the Netherlands, key expectations in relation to school reporting and quality assurance are well defined. There has been significant investment in developing tools and capacity for self-evaluation and improvement at the school level on which future policies can build. There is room to further clarify respective evaluation responsibilities between school boards and school leadership teams and ensure that both have the necessary competencies (Chapter 5).

An important aspect of designing an effective evaluation and assessment framework is to be strategic in linking the different components in order to generate complementarities and prevent inconsistency of objectives. Without this, the co-existence of different approaches may result in costly duplication and incoherence within the evaluation and assessment system (OECD, 2013). In the Netherlands, there is room to develop better synergies between:

- **Teaching standards and teacher appraisal.** While competency requirements for teachers exist and are widely used in initial teacher education, their use for regular appraisal and professional development in schools appears limited (Chapter 4).
- **Teacher appraisal, professional development and school development.** There is no guarantee that school leaders conduct systematic appraisals of their teachers' classroom practices and that these are followed up with adequate professional development. Teacher appraisal and professional development could also be better articulated with school development priorities (Chapter 4).

- **Teacher registration, appraisal and career development.** Currently, registration does not involve a professional appraisal or attestation of teachers' competencies and it does not correspond to a step within the teacher career (Chapter 4).
- **School self-evaluation and school inspection.** There are no central requirements for school self-evaluation and no official strategies to link internal quality care with external evaluations. As a result, there is a risk that the two processes and the criteria they use are not well aligned and do not mutually reinforce each other (Chapter 5).
- **School evaluation and teacher appraisal.** As there is currently no national system for teacher appraisal, monitoring the quality of school-based human resource management could be a key task of the Inspectorate, but it is not yet fully reflected in the Inspectorate's evaluation criteria (Chapter 5).

The environment of multilevel governance and protected autonomy of schools raises challenges for implementation

Striking the right balance between school/local autonomy and national responsibility for quality is a challenge shared by many education systems (OECD, 2013). In the Netherlands, difficulties typically arise when the central authorities' plans confront the longstanding tradition of "freedom of education" (Chapter 1), which is strongly defended by schools and school boards.

The *Schools have the Initiative* programme offers an example of the challenges in promoting large-scale change. As part of this programme, the Ministry granted schools substantial autonomy both in selecting areas of focus (within a broad framework) and the strategy for implementation (Chapter 1). There are very few reporting obligations for schools. The number of schools currently participating in the programme is somewhat smaller than hoped for and the Ministry is considering different approaches to encourage all schools to take part. At the same time, the Ministry has contracted an evaluation of the program, but this will be very challenging given the range of schools, the variety of initiatives undertaken and the lack of documentation. With plans afoot to initiate a second round of the program, the question is how to use the experiences and results of this first round to improve the effectiveness and accountability of the next round.

Dutch schools also have considerable freedom in deciding whether to participate in reform initiatives or research studies supported by the Ministry. Consequently, it is often problematic to obtain the representative samples of schools and students that are essential to high quality research. This can delay both innovation and a shift to more evidence-based policy making (Scheerens, 2013). It also presents challenges for spreading and sharing good practice across schools and school boards.

While freedom of education enables schools to undertake their own experiments in relation to evaluation and assessment, these are typically not documented and collaboration among schools tends to be local, for example within one school board. A consequence is that the existing knowledge and information on evaluation and assessment may easily be lost, schools may spend unnecessary time on "re-inventing the wheel" and there is little system learning over time. Without a strategy to effectively support, leverage and disseminate local innovations, many promising ideas will remain localised or even fade away for lack of external support (Chapter 3).

There are variations in evaluation and assessment capacity

The Dutch approach to educational governance relies strongly on the capacity of school professionals to implement evaluation and assessment policies and to use the results for improvements at the school and classroom level. Observations from the Dutch Inspectorate of Education indicate that most schools are providing satisfactory education to their students, but many have room for improvement with respect to their evaluation and assessment approaches.

Chapter 3 of this report reveals a set of capacity challenges in the area of classroom-based assessment. Current teacher education programmes do not seem to accord sufficient time to assessment-related competencies, and there is evidence that information from standardised tests is often underutilised in schools (Chapter 3). Many teachers lack sufficient assessment expertise to fully exploit information from formative assessments or to interpret aggregate results with a view to improving their teaching strategies. There also appears to be insufficient emphasis on building teachers' skills to develop their own assessment instruments, and most teachers have not received specific training for their summative assessment and marking of student work, especially in the context of the final examinations in secondary education (Chapter 3).

Chapter 5 of this report also raises concerns about the capacity of school leaders to use evaluation and assessment data for whole-school evaluation and development. Although most schools achieve basic quality outcomes, in its 2011/12 report the Inspectorate highlighted widespread shortcomings, including care and supervision, results-oriented teaching, quality assurance and the quality of examinations (Dutch Inspectorate of Education, 2013). While underperforming schools are typically found to use the feedback from inspection visits for quality improvements, there are indications that schools labelled as achieving basic quality are often less engaged in using evaluative information to gain insights into their own challenges and plan for improvement (Dutch Inspectorate of Education, 2013).

Finally, the role of school boards deserves special attention in the Dutch system of educational governance. Evaluation and assessment in the Netherlands operate in a context where the majority of schools are "owned" by non-public bodies (e.g. religious or professional communities pursuing particular pedagogical philosophies) and even public schools are typically operated by semi-autonomous boards, similar to those in the non-public sector (see Chapter 1). As described above, the size and composition of school boards, as well as the evaluation competencies of school board members, are highly variable. In its 2011/2012 report on the State of Education in the Netherlands, the Inspectorate notes that "quality assurance is an area in which many school boards could improve" (Dutch Inspectorate of Education, 2013, p.27).

Policy recommendations

Education in the Netherlands is at a turning point. Although it has made progress on many fronts and has a high standing on international assessments, there is a general appreciation that the system must strive for the next level, although the nature of that level has not yet been specified. According to various groups interviewed by the OECD review team, it could mean improving the country's standing on international assessments, with particular focus on increasing the proportions of students in the highest category; or enhancing general academic achievement and reducing drop-outs; or turning attention to "21st century skills" such as creativity, collaboration and ICT literacy

(Chapter 3). One of the themes of this report is how evaluation and assessment, broadly conceived, can support the attainment of whatever educational goals are decided upon. Building on the many strengths of the Dutch evaluation and assessment approach, this section proposes the following directions for future policy development:

- Embed the evaluation and assessment framework with broader education goals.
- Further develop and integrate the evaluation and assessment framework.
- Continue to build on teacher professionalism.
- Engage stakeholders and build networks for system-wide learning.
- Build capacity for effective and forward-looking evaluation and assessment.

Embed the evaluation and assessment framework with broader education goals

A crucial aspect for the successful implementation of evaluation and assessment is their alignment with student learning objectives. The core logic of criterion-referenced evaluation and assessment systems rests upon the alignment of goals for student learning, specific content for learning, pedagogical approaches and evaluation and assessment approaches (OECD, 2013). In the Netherlands, building such alignment is challenging, given the lack of a national agreement on specific educational priorities and standards.

Thus, it seems essential to begin as soon as possible a broad-based consultative process to build consensus on a comprehensive set of long-term educational goals for Dutch students that will prepare them well for the mid-21st century. In the Dutch governance context, this will need to involve multiple stakeholders, and negotiations are likely to be difficult given the principle of schools' freedom of education. However, a national conversation on how traditional learning goals should be augmented to meet the challenges of the 21st century will help identify the changes that must be made to the evaluation and assessment framework to support innovative, future-oriented and reflective teaching and learning.

One potential approach is for the Knowledge Directorate of the Ministry to commission white papers that review and summarise the literature on various aspects of 21st century skills, as well as strategies adopted by other education systems. Third parties, such as the Primary and Secondary Education Councils and the Education Cooperation, could then take the lead in convening groups of stakeholders to review and discuss a range of potential strategies. A delicate balance would need to be maintained between the need to move forward in a coherent manner and the prerogatives of governing boards to decide on the direction and pace they wish to follow.

To make the evaluation and assessment system coherent, it is important that the learning goals to be achieved are placed at the centre of the framework and that all other evaluation and assessment activities align to work towards these goals. For example, competency descriptions for school professionals (in relation to which they will be appraised) and quality indicators for school evaluation should reflect the learning goals that the school system is aiming to achieve. This also implies ensuring that the evaluation and assessment framework captures a broad range of student learning objectives. To this end, it would be helpful to initiate research and development to strengthen the range of instruments available to assess, for example, students' broader competencies such as problem solving, reasoning and communication (Chapter 3).

The *OECD Reviews of Evaluation and Assessment in Education* found that such policy development processes are more likely to yield consensus and compromise among parties if the rationale for reform is clearly embedded within a long-term vision of what is to be accomplished for student learning (OECD, 2013). Individuals and groups are more likely to accept changes if they understand the reasons for these changes and can see the role they should play within the broad national strategy. This includes dissemination of the evidence basis underlying the policy diagnosis, research findings on alternative policy options and their likely impact, as well as information on the costs of reform vs. inaction. Such communication and dissemination is critical to gain the support of society at large for educational evaluation reforms, not just the stakeholders with a direct interest (OECD, 2013).

Further develop and integrate the evaluation and assessment framework

As described above, most components of evaluation and assessment in the Netherlands are well developed. The country's longstanding tradition and expertise in educational evaluation and assessment provides the Netherlands with a strong basis to further develop and integrate the evaluation and assessment framework. At the same time, this report finds there is room to further develop approaches to formative assessment of students (Chapter 3), teacher appraisal (Chapter 4) and school self-evaluation (Chapter 5).

Completing and integrating the evaluation and assessment framework could begin by developing a comprehensive overview of all the elements that currently constitute the Dutch approach to evaluation and assessment, including approaches to student assessment, teacher appraisal, school evaluation and system evaluation. Such an evaluation and assessment "map" should include formative and summative elements, qualitative and quantitative approaches and school-based and external components. It should outline the main purposes and functions of each component, clarify responsibilities and include links to relevant goals and reference standards, available measurement instruments and related professional development opportunities for stakeholders at different levels. The process itself would help to identify gaps, missing links and potential imbalances to be addressed in future policy development.

The OECD review team has identified a number of ways to move forward. In the area of student assessment, there is room to go further in promoting classroom-based formative assessment and developing teachers' capacity to use assessment results to inform teaching and learning. Teacher appraisal practices could be strengthened by a more systematic focus on observation of teaching practice and feedback for teachers, which should feed into both professional and school development. School evaluation would benefit from further attention to enhancing school self-evaluation as a process and a product, and linking it with external evaluation. Taken together, these elements would help strengthen the school-based components of evaluation and assessment, which would help maintain a balanced approach to evaluation and assessment in the Netherlands. Such school-based processes involving students, teachers and school leaders have a strong potential to produce results that will be useful to shape future teaching and learning and have an impact on actual classroom practice (OECD, 2013).

Continue to build on teacher professionalism

As the most important school-level factor in student achievement, teachers are key to improving education outcomes. The OECD review team commends the current focus of the Dutch government to build on and further develop teacher professionalism in the

Netherlands. Defining and rethinking the framework for evaluation and assessment also provides an opportunity to place teacher professionalism firmly at the heart of the evaluation and assessment agenda.

Evaluation and assessment have little value if they do not lead to the improvement of student learning, and teachers are central to securing links between the evaluation and assessment framework and teaching and learning practice. This highlights the importance for evaluation and assessment frameworks to draw on the professionalism of teachers in ensuring evaluation and assessment activities result in authentic improvement of classroom practices and student learning (OECD, 2013). Given the focus on teacher autonomy and “professional space” in the Netherlands, this will require further work to customise evaluation and assessment approaches in a way that allows learning from practice-based expertise and supports innovative ways of working towards excellence.

Further work to enhance teacher professionalism can take place at different levels of the education system. Channels that are likely to reinforce the professionalism of teachers and to build links to classroom practice include: an emphasis on teacher appraisal for the continuous improvement of teaching practices; ensuring teaching standards are aligned with student learning objectives; involving teachers in school evaluation, in particular through conceiving school self-evaluation as a collective process with responsibilities for teachers; preparing inspectors to understand the complexity of school-based human resource management and professional learning; ensuring that teachers feel the ownership of student assessment and accept it as an integral part of teaching and learning; building teacher capacity for student formative assessment; and building teachers’ ability to assess students in relation to national goals and reference levels (OECD, 2013).

Engage stakeholders and build networks for system-wide learning

Effective implementation seeks to strike the right balance between top-down and bottom-up initiatives, which is generally believed to foster consensus (Finlay et al., 1998). Given the traditional autonomy accorded to schools, any top-down imposition of innovative evaluation and assessment approaches is likely to be problematic in the Dutch context. It seems more feasible to develop evaluation and assessment policies through the co-operation of different stakeholders towards a common goal.

The *OECD Reviews of Evaluation and Assessment in Education* found that educational evaluation policy has much more to gain from compromises grounded in the productive cross-fertilisation of different perspectives than from their antagonism and the imposition of one view over those of other stakeholder groups (OECD, 2013). For instance, teachers will be more likely to accept a new appraisal process if they are consulted in its design. In addition to taking their interests and concerns into account, the participation of teachers recognises their professionalism, the scarcity of their skills, and the extent of their responsibilities. If teacher appraisal procedures are unilaterally designed at the level of the administrative structure, without addressing and including the core of teaching practice, then there will be a “loose coupling” between administrators and teachers that will both fail to provide public guarantees of quality, and will discourage reflection and review among teachers themselves (Elmore, 2000; Kleinhenz and Ingvarson, 2004).

In more general terms, this calls for practitioners, such as school leaders and teachers, to be engaged in the design, management and analysis of evaluation and assessment policies. Consensus building among stakeholders is all the more important as local actors are often in the best position to foresee unintended consequences and judge what is

feasible in practice (OECD, 2013). New projects to develop innovation in evaluation and assessment should involve partnerships between evaluation and assessment organisations and groups of schools where pilots would take place. Ideally, some professionals of these schools would participate in the research and development aspects and not simply serve to administer new tools (Chapter 3). The Inspectorate should also play a key role in recognising and disseminating promising innovations developed at the school level (Chapters 3 and 5).

In addition, developing more deliberate improvement networks among practitioners can be a powerful organisational tool that embeds reform in interactions of different stakeholders, shares and disperses responsibility, and builds capacity through the production of new knowledge and mutual learning that can feed back into policy and practice (Katz et al., 2009; Chapman and Aspin, 2003). The central authorities can contribute to creating an ambition-friendly and innovation-friendly environment, by providing funding and support for schools and networks of schools to accelerate their work, and to provide regional and national forums where they can showcase their efforts to a broader audience.

Build capacity for effective and forward-looking evaluation and assessment

The effectiveness of the overall evaluation and assessment framework depends to a large extent on whether those who evaluate and those who use evaluation results at the different levels of the system have the appropriate competencies.

In a system where teachers have considerable autonomy in student assessment, strong teacher skills for both formative and summative assessment are essential to monitor progress towards learning goals and ensure that results are used for improvements. Ideally such training should begin during initial teacher education and be carried forward through ongoing professional development to deepen assessment skills. Resources can also be provided locally, online, and through specialised workshops and various networks. Priority areas identified in Chapter 3 include: building teacher competencies to interpret and use assessment results for the improvement of classroom practice and differentiated instruction; enhancing classroom-based assessment, in particular short-cycle formative assessment and feedback; and evaluating and marking complex student work. Serious consideration should also be given to instituting a systematic moderation process to monitor teachers' marking, especially for the final examinations in secondary education (Chapter 3).

It is important that schools have appropriate expertise related to effective teacher appraisal and school self-evaluation. Given the key role of school leadership in the Netherlands' devolved education context, it is difficult to envisage either effective teacher appraisal or productive school self-evaluation without strong leadership capacity. Hence, recruitment, development and support for school leaders are critical in creating and sustaining effective evaluation and assessment cultures within schools. Research internationally has shown that school leadership that focuses on goal-setting, assessment, appraisal and evaluation is positively correlated with teacher and student performance (Pont et al., 2008). To strengthen school leadership in these areas, the Netherlands could consider introducing mandatory training for school leaders, ensuring broader dissemination of resources and training for educational evaluation, and establishing effective performance appraisal for school leaders themselves (Chapters 4 and 5).

Finally, as school boards are key players within the Dutch evaluation and assessment system, developing their evaluation competencies should be a priority. The Primary and

Secondary Education Councils are well placed to support the professional development of school board members and facilitate the sharing of experiences and mutual learning among school boards. Below the national level, the creation of an intermediate regional support structure could help the development of local and regional networks of school boards. Research on the role of school boards, their capacities, shortcomings and potential should also be supported (Chapter 5).

Notes

- ¹ Schools will be required to report on their students' performance in relation to these reference levels from the 2015/16 school year. In the 2014/15 school year, all schools will participate in a pilot study on the coverage of reference levels in end-of-primary tests.

References

- Adamson, F., & Darling-Hammond, L. (2013), Policy Pathways for 21st Century Skills, http://atc21s.org/wp-content/uploads/2013/02/ATC21S_Policy_Paper-20120730.pdf.
- Chapman, J. and D. Aspin (2003), “Networks of learning: A new construct for educational provision and a new strategy for reform”, in B. Davies and J. West-Burnham (eds.), *Handbook of Educational Leadership and Management*, pp. 653-659, Pearson, London.
- Dutch Inspectorate of Education (2013), The State of Education in the Netherlands: Highlights of the 2011/2-12 Education Report, Inspectie van het Onderwijs, Utrecht.
- Elmore, R. (2000), Building a New Structure for School Leadership, Albert Shanker Institute, Washington.
- European Commission (2011), Assessment of Key Competences: Policy Handbook, European Commission, Brussels.
- Faubert, V. (2009), “School Evaluation: Current Practices in OECD Countries and a Literature Review”, *OECD Education Working Papers*, No. 42, OECD Publishing, Paris, <http://dx.doi.org/10.1787/218816547156>.
- Finlay, I., S. Niven and S. Young (1998), “Stakeholders, consensus, participation and democracy”, in I. Finlay, S. Niven and S. Young (eds.), *Changing Vocational Education and Training: An International Comparative Perspective*, Routledge, London.
- Goldin, C., & Katz, L. (2008), *The Race Between Education and Technology*, Harvard University Press, Cambridge, Massachusetts.
- Katz, S., L.M. Earl and S. Ben Jaafar (2009), *Building and Connecting Learning Communities: The Power of Networks for School Improvement*, Corwin, Thousand Oaks, California.
- Kirsch, I., et al. (2005), *America’s Perfect Storm*, Policy Information Center, Educational Testing Service, Princeton, New Jersey.
- Kleinhenz, E. and L. Ingvarson (2004), “Teacher evaluation uncoupled: A discussion of teacher evaluation policies and practices in Australian states and their relation to quality teaching and learning”, *Research Papers in Education*, Vol. 19, No. 1, pp. 31-49, Routledge, London.
- NRC (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. J.W. Pellegrino and M.L. Hilton (eds.), National Research Council, The National Academies Press, Washington DC.
- OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264190658-en>.

- Pont, B., D. Nusche and H. Moorman (2008), *Improving School Leadership, Volume 1: Policy and Practice*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264044715-en>.
- Santiago, P., et al. (2012), *OECD Reviews of Evaluation and Assessment in Education: Mexico*, OECD, Paris.
- Scheerens, J., et al. (2012), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands*, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.
- Scheerens, J. (2013), “Educational Evaluation and Assessment in the Netherlands”, Addendum to the Country Background Report for the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes, www.oecd.org/education/school/Netherlands%20CBR%20Update.pdf.
- Secondary Education Council (2011), *Werkboek Horizontale Verantwoording VO*, VO Raad, Utrecht.
- Stecher, B.M., et al. (2000), *The Effects of the Washington State Education Reform on Schools and Classrooms: Initial Findings*, DB-309-EDU, RAND Santa Monica, California.

Chapter 3

Student assessment

Student assessment in the Netherlands is largely the responsibility of schools and classroom teachers, supported by well-developed standardised assessment tools. The Dutch education system has rich resources and expertise in assessment design, development and administration. There has been strong attention to reaching high quality standards and investing in continuing improvements to central tests. The assessment framework relies on a balance between teacher-based and central assessments, with a recent focus on clarifying reference levels for student learning and strengthening “results-oriented work” in schools. There is an extensive system of formative and summative assessments as well as pockets of promising innovations. To exploit and scale up these innovations, it seems essential to engage a broad-based consultative process to build consensus on the education goals for future generations. It is likely that such a consensus will involve a rethinking of traditional learning goals, as well as the adoption of some of the 21st century skills as important curricular goals. As a consensus on student learning goals begins to emerge, work on a corresponding assessment strategy can begin, including further development of the assessment infrastructure, efforts to strengthen teacher professionalism in assessment and support for innovative assessment practice at the local level. In the short term, there are opportunities to leverage assessment data that is currently being generated and to critically examine current practices that may impede innovation and improvement, with a view to ensuring a balanced use of assessment as, for and of learning.

This chapter focuses on approaches to student assessment within the Dutch evaluation and assessment framework. Student assessment refers to processes in which evidence of learning is collected in a planned and systematic way in order to make a judgment about student learning (EPPI, 2002). This chapter looks at both summative assessment (assessment *of* learning) and formative assessment (assessment *for* learning) of students.

Context and features

References for student assessment

As described in Chapters 1 and 2, there is no national curriculum, but the Ministry of Education, Culture and Science sets core learning objectives that students are expected to achieve by the end of both primary and lower secondary education. For the upper cycle of secondary education, the Ministry of Education, Culture and Science has formulated learning targets, which have been translated into centrally set examination syllabi. These strongly influence the curricula taught in upper secondary education. In addition, more detailed reference levels were set for Dutch language and mathematics. The reference levels are defined for the end of primary education and the end of each of the educational programmes in secondary education. Schools have been required to implement the reference levels since 2010 and they will have to report on their students' performance in relation to these from the 2015/16 school year onwards. To facilitate this, all primary schools will participate in a pilot study during the 2014/15 school year to evaluate the coverage of the reference levels in existing end-of-primary tests.

Approaches to student assessment

Student assessment, in its different guises, plays many important roles in the Dutch education system. These roles can be organised under three main headings, which will be explored further below: monitoring/improvement, decision-making/certification, and school reporting/accountability. However, it should be noted that the purposes for assessments are not always clearly stated and that results from the same assessment are sometimes used for several purposes. New legislation developed in 2013 is likely to strongly influence student assessment policy and practice in the coming years (Box 3.1).

Box 3.1 New Dutch laws on student assessment developed in 2013

At the time of the OECD review visit in June 2013, important changes to the legislation on student assessment were being prepared. In December 2013, the First Chamber of Parliament accepted a law proposal making it mandatory for primary schools to administrate regular student monitoring systems as well as a final summative test at the end of Year 8. In implementing this policy, schools will be allowed to choose between different tests developed by the Central Institute for Test Development (Centraal Instituut voor Toetsontwikkeling, Cito) or other companies, provided that they meet central quality requirements. These laws will be implemented from the 2014/15 school year.

Source: www.eerstekamer.nl/behandeling/20140116/publicatie_wet/document3/f=/vijggcjjiv0zs.pdf.

Assessment for monitoring and improving student learning

Assessment strategies that aim to monitor and improve student learning are broadly referred to as formative assessment. The essence of formative assessment is the collection of evidence regarding the present status of each student in a class and the use of that evidence to adjust the content and/or method of instruction so as to optimise the learning trajectories of all students (Earl, 2003). Formative assessments in Dutch classrooms take many forms, ranging from the informal to the formal. In interviews with the OECD review team, teachers referred to a variety of instruments, some “home-made” and others adopted from instructional packages or purchased directly from vendors. Teachers also use other contextual information (typically only available at the classroom level) in their pedagogical decision-making.

Although formative assessment in the Netherlands is not inscribed in central education laws or regulations (OECD, 2013), it is promoted through other documents and initiatives. Since 2010, there has been considerable policy focus in the Dutch education system on using assessment results for the improvement of student learning. Under the heading “results-oriented work”, the Ministry of Education, Culture and Science and the Inspectorate of Education are encouraging schools to develop systematic, goal-oriented processes to maximise student learning. The policy on results-oriented work emphasises different elements of formative assessment, including working with the national reference levels, using student monitoring systems to measure learning growth, providing feedback to students and differentiating instruction to achieve improvements in learning outcomes (Dutch Inspectorate of Education, 2010).

Several initiatives have been developed since 2010 to facilitate the implementation of results-oriented work in schools. The Dutch Inspectorate of Education (2010) published a report on results-oriented work and the General Institute for Curriculum Development has been developing further guidance regarding the implementation of a results-oriented approach. Results-oriented work also features among the priority domains of the *Schools have the Initiative* programme (Chapter 1). As part of this programme, self-evaluation instruments were developed for schools to monitor their own progress in implementing results-oriented work. In addition, two university research projects (the *Focus* project run by the University of Twente¹ and the *STREEF* project run by the Rijks Universiteit Gronigen²) were developed to train teachers in using assessment results to improve teaching strategies for different sub-groups of students. Since 2010, the *Focus* project has provided training to 45-65 schools each year.

The new laws on student assessment (Box 3.1), which will take effect in the 2014/15 school year, further strengthen the focus on results-oriented work. They mandate that all primary schools implement a student monitoring system to regularly assess their students’ progress in a broad range of subjects (Box 3.1). While primary schools will be required to use a student monitoring system, they will retain the freedom to choose the provider and the frequency of test administration (Scheerens et al., 2012). There are three comprehensive student monitoring systems available to schools for this purpose: LVS (*Leerling Volg Systeem*)³, ParnaSys⁴, and ESIS (*Elektronisch School Informatie Systeem*)⁵. In the primary sector, virtually all schools participate in the LVS developed by the Central Institute for Test Development (*Centraal Instituut voor Toetsontwikkeling, Cito*) (Box 3.2). A range of student monitoring instruments are also offered by Cito for the first two years of secondary education (Box 3.2). In addition, the government commissioned Cito to develop an adaptive test for students with special educational needs (SEN) students.

In line with the focus on results-oriented work, there has been increased interest in measuring students' learning growth. With the introduction of a unique student number (as part of the *Basic Register for Education, BRON*), it is now possible to track students' learning trajectories wherever they attend school. At the time of the OECD review visit, a pilot study regarding student learning growth and value-added measurement was being conducted, in cooperation between the Ministry of Education, Culture and Science, the Inspectorate, educational researchers and practitioners. The final report of this study was published in January 2014 (Janssens et al., 2014). The report argues that learning growth models can be used by schools to analyse the progress of their students, as part of results-based work.

Box 3.2 Dutch student monitoring systems developed by Cito

The LVS (*Leerling Volg Systeem*) is a longitudinal student monitoring system developed by the Central Institute for Test Development (Centraal Instituut voor Toetsontwikkeling, Cito) and offered for Years 1-8 (Scheerens et al., 2012). The tests are taken once or twice a year and are completed by hand or, for some subjects (since 2003), by using computer-based modes. Tests in ordering, language and orientation in space and time are given in Years 1 and 2 only. For Years 3-5, tests are given in several aspects of Dutch language, arithmetic/mathematics, and social and emotional development. These tests are also given in Years 6-8, along with world orientation (geography, history, biology), science and technology, and English (Years 7 and 8 only). The formative/diagnostic function is accomplished through provision of interpretive materials, as well as suggestions for relevant pedagogical strategies.

In a given subject, the tests are vertically linked so that a student's progress across grades can be represented as a series of steps along a single scale. That trajectory is depicted graphically in a student report, along with normative comparisons against five different reference groups. Families of students enrolled in special education schools receive an alternative student report that indicates the level of achievement of the student and provides assistance in comparing the results with children of the same age who are attending mainstream schools. There is also a group survey report that displays the results (trajectories and level scores) for a specific group of students over a number of years.

For the first two years of secondary school, several computer-based, multiple-choice, monitoring assessments are offered by Cito for different tracks, subjects and levels. The tests are administered at the beginning of secondary school and at the end of the first and second years. Tests are offered at three levels of difficulty, corresponding to the three tracks in lower secondary education. For each track, four subjects are tested: Dutch reading comprehension, English reading comprehension, mathematics and study skills. The results are used for didactic purposes and to assist in deciding the appropriate track in upper secondary. Thus, for students in lower secondary education, the monitoring and evaluation system has a dual role of providing pedagogically useful advice, and of directly offering information relevant to the decision of which track to pursue in upper secondary.

Source: Scheerens, J., et al. (2012), OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.

Assessment for decision-making and certification purposes

Assessment for summative (decision-making or certification) purposes typically occurs at the end of a learning unit, school year or educational level. Throughout primary and lower secondary education, summative assessment in the Netherlands is largely based

on teachers' professional judgement and supported by the availability of standardised assessment tools. Assessment tools chosen by teachers vary and students are given various opportunities to demonstrate their ability across a range of assessment contexts. In school visits, educators mentioned developing tests on their own or with their colleagues, using tests provided by the particular "method" they were employing, and purchasing off-the-shelf tests offered by different vendors. Some schools have introduced portfolio assessments to provide a more comprehensive view of student work. Others are employing various systems to aggregate, organise and display student data in order to make it more usable. Student results are reported to students and parents three times a year. They are represented on a scale of one-ten, with six being the pass mark.

In the final phase of primary education, schools are required to report on learning results and clarify the extent to which students have reached the core learning objectives for primary education. While schools are free to use different assessment instruments for this purpose, the vast majority of schools (85%) purchase and use the end-of-primary test developed by Cito. The results from this test provide information on the school type that would be most suitable for each student in the next phase of education. This is a key indicator employed by secondary schools, as well as by students and their parents, in the selection of a secondary school and an appropriate track within the school. At this transition point, the end-of-primary test functions as an external reference for achievement for students, parents and teachers, a validation of the teacher-based school advice and a link between learning targets for primary and secondary education.

The end-of-primary tests are in multiple-choice format, with tested subjects comprising Dutch language, arithmetic/mathematics, and study skills. World orientation is an optional subject. Cito has developed tables that, for almost all secondary tracks, indicate the advice associated with different score bands. There are three different bands: scores in the highest band lead to automatic acceptance, those in the middle band call for consultation with the primary school, and those in the lowest band require more extensive research. The outcome of this test, as well as the recommendation of the teacher (in consultation with the parents) combine to advise parents on the type of secondary school their child should attend.

New laws, that will take effect for the 2014/15 school year (Box 3.1), mandate that all primary schools administer a standardised end-of-primary test. However, schools will retain the right to choose among different examination providers (i.e. alternatives to the Cito test⁶). The draft law also stipulates that the timing of administration will be moved to later in the school year, from February to April. As a result, students will start applying for admission to secondary schools before having received the results of the end-of-primary test. This means that test results will be less useful to the secondary schools in the selection of students and that the recommendation of the school will gain more weight.

In the secondary sector, there is a collection of school-leaving examinations for each track. Typically, in each subject/track, there is a central examination developed by Cito that is aligned with the core learning objectives for that subject and follows the blueprint prepared by the College for Examinations (*College voor Examen, CVE*). Cito constructs the examination based on input from both subject matter experts and teachers. In addition, schools develop their own examination in each subject in general accordance with the syllabus for the subject. School examinations are reviewed and approved by the Inspectorate (Scheerens et al., 2012). Typically, examinations comprise both multiple-choice and open-ended questions.

All students must sit for examinations in Dutch language, a foreign language (usually English) and social sciences. In the pre-vocational streams at the secondary level, students also take examinations in practical and vocational studies. In addition, depending on the track and student interests, they may choose from among mathematics, science, modern foreign language, or other subjects. An examination is first graded by the student's own teacher following guidelines prepared by the CVE. These guidelines include detailed prescriptions including a set of rules for scoring students' responses on each individual item of an examination. A second grading is conducted by a teacher from another school. If there is a disagreement, then there is an attempt to achieve consensus. If that fails, the Inspectorate appoints a third grader whose grade is final (OECD, 2013).

School examination performance is given approximately equal weight to the central examination performance, although some interviewees indicated a shift in importance toward the central examination. Overall performance on the examinations determines the nature of the certificate earned by students and the kind of training or further education they can pursue. Overall performance is reported on a scale of one-ten, with 5 and a half (without rounding) being a passing mark. In the HAVO and VWO tracks, one mark of five in either Dutch language, mathematics or foreign language is allowed for students to pass. The CVE makes some adjustments to the scores in order to take account of variations in difficulty from year to year. Certificates awarded to students at the end of each track in secondary information provide pass/fail information and marks in each subject (OECD, 2013).

Assessment for school self-evaluation, accountability and reporting purposes

With respect to school-level accountability/monitoring, many of the assessments mentioned above are also used to provide information at the school, school board, regional and national levels. For school self-evaluation, results from the LVS by grade/subject within a school can be combined over a number of years (for greater stability) and then compared with the results at the regional or national levels. LVS results can also be aggregated for different segments of the population (e.g. students with special education needs, or students from particular ethnic groups) to enable comparisons: (i) between the group and the national results; (ii) among groups; (iii) within the group over time. These comparisons can be used by schools for school self-evaluation, early identification of risks and reporting to stakeholders.

The Inspectorate emphasises that the LVS is owned by the school and not part of public accountability. The LVS results are not used in the Inspectorate's annual risk analysis of schools. Only in the evaluation of schools considered at risk, inspectors will ask schools to share this data to feed into the inspectors' preparation for their school visit. In schools not at risk, inspectors on site will look into parts of the LVS results to understand learning trends in the school.

Aggregate results from the Cito end-of-primary test are currently used by the Inspectorate (and others) as an indicator of school quality. In addition, the central authorities are proposing to set national targets for performance based on these tests, as a means of encouraging schools to achieve at higher levels. The assessments administered at the end of secondary school are also aggregated to the level of the school (and higher) for monitoring and accountability purposes⁷.

In addition, sample-based national assessments have also been developed specifically for education system evaluation. For further details regarding the use of assessment information for school evaluation and system evaluation, see Chapters 5 and 6.

Responsibilities for student assessment

As with other aspects of education in the Netherlands, responsibility for student assessment involves a balance between central authorities, such as the Ministry of Education and the College for Examinations, and local actors such as governing boards, schools and individual teachers. The nature of the balance depends on the particular function: formative assessments are decided and conducted primarily at the local level, even though schools and teachers are influenced by the results of standardised student monitoring systems; decisions about students' academic tracks or their attainment of a diploma at the end of secondary education are based on information drawn both from the school level and from standardised assessments; and monitoring and accountability is largely conducted by central authorities using evidence provided by both schools and national databases.

The central examinations in secondary education are prepared under the auspices of the College for Examinations, in collaboration with various stakeholders. The test development and administration is carried out by specialised organisations such as Cito. In secondary education the syllabus for a subject allows flexibility for what is tested as part of the school-based component. The extensive involvement of teachers in secondary examinations, both through the school-based component and through the grading of the open-ended questions in the central component, is likely to be a factor in existing support for the system among educators.

Strengths

In the area of student assessment, the Netherlands possesses many strengths. It has rich resources in assessment design, development and administration, with Cito being the best known testing organisation. There has been strong attention to reaching high quality standards and investing in continuing improvements to central tests. The assessment framework relies on a good balance between teacher-based and central assessments, with a recent focus on clarifying reference levels for student learning and strengthening results-oriented work in schools.

Reliable measures of student learning outcomes are available

A clear priority in assessment frameworks is the development of reliable measures of student learning outcomes. The Netherlands stands out internationally with regards to the development of standardised assessments at key stages of education. Major advantages of external standardised assessment include its high reliability and low cost of administration. Standardised central assessment also helps to clarify learning expectations for all schools and motivate teachers and students to work towards high standards and steer their teaching and learning strategies in that direction (OECD, 2013).

Standardised assessment ensures that all students are assessed on the same tasks and that their results are measured by the same standards. The results are made as objective as possible so that they are, within a year, comparable among students, regardless where they go to school. For example, the high-stakes tests developed by Cito for the end of primary and secondary school have excellent psychometric properties as they are highly reliable and carefully equated from year-to-year to preserve scale stability and interpretability over time. They make data on student learning outcomes available, providing a picture of the extent to which student learning objectives are being achieved,

and they grant the opportunity to compare individual student achievements in the tested areas.

High level of expertise in developing standardised assessments

As indicated above, Cito produces tests for summative assessments, as well as tests for monitoring and evaluation. In general, the technical characteristics of the assessments appear to be very strong and throughout the review team's interviews with stakeholder groups, there was little or no criticism of tests on that account. Cito has a large staff with expertise in educational measurement, test design and construction, and test administration. In particular, they employ individuals with specialised expertise in areas such as psychometrics, computer-adaptive testing, and the design and maintenance of the technology infrastructure for large-scale assessment. Cito supplements its in-house staff with consultants drawn from schools and tertiary institutions. Teachers are also heavily involved in the design and quality evaluation of Cito tests employed in the Netherlands. Cito's scientists are well-known contributors to the research literature on assessment and they provide support to the Ministry for the development of innovations in school assessment. Cito's international reputation is attested to by its many partnerships with Ministries of Education in other countries, other testing firms, and its successes in obtaining international contracts.

There are many other entities and vendors providing either stand-alone assessments or tests that complement other educational materials, such as textbooks. In addition, university departments and research institutes undertake studies regarding assessment practices, many of which are commissioned by the Ministry of Education to inform policy decisions.

Continuous development and innovation in student assessment

There is considerable innovation in assessment at the different levels of the system. For example, for some subjects in the vocational education track, the Ministry has introduced a framework for competency-based assessments that is aligned with the nature of proficiency in those domains.

With regard to new modes of test delivery, the monitoring tests for lower secondary education are computer based, and some components of the LVS are administered as computer adaptive tests. Adaptive testing involves tailoring the sequence of items (or sets of items) presented to the student's response pattern. Thus, a student who answers the first few questions correctly is administered more difficult questions, while the student who answers incorrectly is administered easier questions. The result is that the scores for most students are estimated more accurately than would be the case with the same, fixed examination administered to all students. This is especially the case for students at the low and high ends of the score range (Wainer, 2000).⁸ Consequently, if and when the Ministry of Education decides to employ computer administration for the end-of-sector examinations, it can draw on substantial experience with this approach and can rely on some familiarity on the part of schools.

Cito invests in continuing improvements to the secondary school examinations. In the interviews with the OECD review team, Cito representatives mentioned moving to a matrix assessment design in order to broaden coverage of the syllabus, and experimenting with "self-auditing examinations" to detect inappropriate test preparation (Koretz and Beguin, 2010). In a typical matrix design, all students are administered a common block of items, and then there are one or more blocks of items that are administered only to

random sub-samples of students. These latter blocks can be used to: (i) obtain estimates (at a population level) of performance on a larger group of items than would be possible in a single examination, given time constraints; (ii) equate tests from year-to-year; (iii) construct or maintain a vertical scale; (iv) obtain information on the psychometric properties of new items; (v) introduce items from previous examinations (usually several years earlier) to compare the performance of different cohorts and determine if there has been “scale drift”, and/or if students have become adept at answering specific kinds of questions, but have not truly developed the essential competencies the items are intended to test.⁹

At the request of some schools, Cito has also developed a digital portfolio to support more innovative approaches to documenting and evaluating student work. Although this is not currently a high priority, it may become sought after by more schools as greater emphasis is placed on developing and assessing complex competencies, such as problem-solving, creativity, critical thinking and teamwork.

The Ministry has established a unique student number that tracks students as they change schools. This facilitates longitudinal studies of student achievement, and means that fewer students are “lost” during monitoring studies. Analysis is therefore less biased and more reliable, as transient students often perform differently to more stable students.

The newly introduced reference levels provide greater clarity on expectations for student learning

The Ministry has recently published reference levels for Dutch language and mathematics to provide more clarity about instructional objectives at the primary and secondary level, which will result in some modifications to central tests and examinations. This is particularly important in view of the principle of freedom of education that is so prominent in the Netherlands. Greater clarity should help teachers across the country create syllabi at each grade level that better represent national learning goals, as well as develop assessments with improved coverage of those learning goals. Ideally, implementation of the reference levels will result not only in more equality in students’ opportunity to learn, but also in better alignment of instruction across different year levels and sectors of education.

The introduction of more detailed national expectations of what should be taught and assessed in schools has, to varying degrees, been debated and tested in many countries over the last twenty five years. In all student assessment systems, there is a need for expected standards of student performance at different levels of education to have clear external reference points. While it is important to leave sufficient room for teachers’ professional judgements in the classroom, it is necessary to provide clear and visible guidance concerning valued learning outcomes. Such benchmarks are intended to provide consistency and coherence, especially where there is a high degree of local autonomy in the development of curricula, teaching programmes and assessments. They can help bring about equality and fairness in educational opportunities for students across the country.

Recent policy emphasises “results-oriented work” at the classroom and school level

Recent educational policy-making in the Netherlands has had a strong focus on stimulating formative assessment and differentiated instruction through “results-oriented work” at the school level. New laws, to be implemented in the 2014/15 school year, will

require schools to use student monitoring systems for results-oriented work in schools (Box 3.1). Results-oriented work involves helping schools to more fully exploit student monitoring systems and, by analysing the information generated, to design appropriate teaching and learning strategies. Teachers are expected to explicitly define learning targets, regularly assess student performance, adapt teaching and learning to student needs, and intervene rapidly to help those who are falling behind in relation to set targets (Visscher and Ehren, 2011; Scheerens et al., 2012). Such approaches are to be stimulated and monitored by the Inspectorate.

Student monitoring systems, such as the LVS, are important tools that support results-oriented work at the school level. Schools are free to choose in which subjects and at what frequency they use the tests with their students. Such externally designed assessments can provide important signposts for teachers and students by indicating the learning goals that are expected nationally and, by producing timely data that may inform teaching strategies they can offer interesting pedagogical tools for teachers. Frequent use of high-quality monitoring systems can also stimulate teachers' own assessment expertise by providing examples of adequate test items to measure particular learning goals. Positive effects of using the results from such assessments to inform teaching may include: greater differentiation of instruction, greater collaboration among colleagues, an increased sense of efficacy and improved identification of students' learning needs (van Barneveld, 2008).

Because the tests are equated across years and vertically scaled across grades, individual student growth trajectories can be calculated, which provide a longer term perspective on student progress. A recent pilot study (Janssens et al., 2014), conducted in cooperation between the Ministry of Education, Culture and Science, the Inspectorate, educational researchers and practitioners, emphasised the importance of focusing on student progress rather than absolute performance. The report found that measuring student learning growth can make valuable contributions to the further development of results-oriented approaches in schools (Janssens et al., 2014).

The assessment framework relies on a good balance between school-based and standardised assessments

Overall, there is a reasonable balance at key decision points between the use of school-based results and central examination scores. In the transition from primary to secondary education, for example, the school's recommendation is as important as the Cito test results in determining the school and the track most suitable for the student. Although some secondary schools prefer to admit students to certain tracks only if they have sufficiently high Cito scores, a strong argument from the primary school can sway the decision, even if the threshold scores are not achieved. The importance of the school's recommendation may even increase with the proposed later administration time of the Cito test.

As noted earlier, at the end of secondary school, a subject assessment consists of both a centrally prepared examination and a school-developed examination. The fact that schools have some flexibility in deciding the content of their examinations is a mark of educators' professional autonomy, and can lead to improvements in the coverage of the syllabus. For example, in modern languages, the central examination at the end of secondary education focuses on reading skills, which allows school-developed examinations to focus on other skills such as writing, speaking and listening. Moreover, the school tests can be administered over a longer period of time, which can reduce the

pressure on students. A similar division of labour can take place in other subjects, although the actual extent of coverage and the quality of the school-based components is not generally monitored.

This flexibility, together with the integral role that teachers play in the construction and scoring of the central examinations, is likely to account for the general credibility that the secondary examination system has among teachers. This is an important strength of the system as there is less of a disjuncture between what the syllabus mandates and what teachers feel they must do to prepare their students for the examinations. This stands in contrast to some other countries where teachers can be demoralised by having to “teach to the test” in high stakes settings (Madaus et al., 2009).

Although the involvement of teachers in this aspect of the final examinations is both commendable and cost effective, it does raise questions of score comparability across schools. There is also some concern about teachers designing the school-based part of the examination to resemble the central part, which may reduce the scope of material covered in examinations, with potential adverse consequences on teaching and learning. These issues will be examined further under ‘Challenges’.

Challenges

The OECD review team formed the impression that a national consensus on a set of education priorities is some time off (see Chapter 2). Given the structure and traditions of the Dutch system, it is likely to emerge slowly, informed both by the many pockets of innovation arising around the country and by political developments at the national level. The principal challenge for student assessment is how it should evolve in order to best support innovative practices and progress toward new priorities, rather than being an obstacle to their attainment. The difficulty is that assessment development tends to lag behind curricular innovations, so that student results on current assessments may not reflect new or deeper learning. Such a disjuncture can lead to premature and misleading inferences about the utility and efficacy of the innovation.

Developing assessments for the mid-21st century

Across OECD countries, there is a shared ambition to move education systems beyond traditional pedagogical approaches focussing on knowledge transmission and acquisition of basic skills and to promote a broader model of learning that includes competencies to synthesise, transform and apply learning in real-world situations. Such key competencies, or 21st century skills, typically include dimensions such as critical thinking creativity, problem-solving, communication, information and communications technology (ICT) literacy, as well as collaborative, social and citizenship skills (OECD, 2013).

Although some of these skills are already incorporated in national learning goals in the Netherlands, it is likely that a broader set of these skills and competencies will become part of the goals that are set at the national level (Chapter 2). Current paper and pencil tests with their limited item formats will not be able to appropriately assess these skills, neither for formative nor for summative purposes. Thus, there will be a need to develop the expertise and technical capacity to design, develop, deliver and evaluate more complex assessments. Many of these assessments will be integrated with instruction and take place in settings using ICT, with access to the internet, in closed micro-worlds, or in game-like environments (OECD, 2013).

Such developments towards more innovative assessment strategies are likely to raise capacity challenges at the school level. As discussed further below, a significant proportion of teachers and school leaders need to further develop their assessment expertise to fully exploit information from formative assessments or to properly interpret aggregate results (Visscher and Ehren, 2011). This gap will only be exacerbated with the introduction of more challenging learning goals and the corresponding need for new forms of assessment. Efficiently remedying this situation in light of the tradition of school autonomy and budgetary constraints will be a challenge.

Differing views on future directions for assessment

As has been noted earlier in the chapter, the Netherlands has deep resources in many facets of educational measurement that support an extensive system of formative and summative assessments, as well as a range of monitoring assessments. Moreover, there are pockets of promising innovations. Exploiting and scaling these innovations requires a national strategy that is “under development” but has not yet fully crystallised.

As might be expected, various stakeholders have different views on whether the current balance in the assessment framework is a reasonable one and what, if anything, should be done to shift that balance in a specific direction. This was illustrated during interviews as part of the OECD review visit. On the one hand, leaders from the Ministry of Education, Culture and Science asserted the need for a greater role for assessment. The introduction of laws making the use of standardised formative and summative assessment mandatory for schools is one manifestation of this view. They also noted that greater use could be made of the assessment results that are now generated on a regular schedule. On the other hand, some interviewees expressed worries regarding the apparent increase in the importance and influence of standardised examinations, particularly those administered at the end of secondary school. Some interviewees referred to this as a “testing frenzy”, and argued that overreliance on test results was a symptom of a lack of trust in the professionalism of teachers. For example, the syllabi governing the central examinations, and the examinations themselves, are playing an increasingly important role in the lives of students and educators, which leads to tensions with some stakeholder groups (more on this below).

Freedom of education enables schools to undertake their own experiments with new assessments (or other aspects of education) and the Inspectorate can contribute to spreading innovations by drawing attention to schools that have developed innovative projects. But collaboration among schools tends to be local, mostly within the school board. All six schools interviewed by the OECD review team indicated that they participated to some degree in a local network, but only one school mentioned participation in a broader network¹⁰. During OECD interviews, there was no mention of the existence of national networks, or of plans to initiate such networks. The challenge, then, lies in how to effectively support, leverage and disseminate local innovations and best practices. Without a strategy to accomplish this, many promising ideas will remain localised or even fade away for lack of external support.

Building capacity to interpret and use assessment results for improvement

One of the key priorities of the government’s policy on “results-oriented work” is to ensure that student assessment results lead to durable improvements in student learning. While student monitoring systems are widely used in Dutch classrooms, there is evidence that many teachers have difficulties in interpreting and effectively using the information

generated by such assessments to improve teaching and learning (Dutch Inspectorate of Education, 2010; 2013). A study by Visscher and Ehren (2011) found that data from the LVS was generally underutilised. Specifically, they noted the following problems: (i) technical snags in the storage and retrieval of data; (ii) a focus on underperforming students, rather than all students who need assistance; (iii) errors in interpreting the data; (iv) failure to take advantage of the longitudinal nature of the data and to use the data as a basis for collective discussion around student work.

The latest annual report of the Inspectorate notes that many teachers have difficulties in providing effective feedback to students and differentiating instruction to respond to individual student needs. The Inspectorate also finds that, despite improvements in this area, there are still major differences among schools in the degree to which teachers and schools appropriately adapt instruction to students with special education needs, a sub-population that appears to be growing (Dutch Inspectorate of Education, 2013). Shewbridge et al. (2010) noted similar challenges with respect to students who are not native Dutch speakers. It is likely that these challenges are shared by school leadership teams and that the information provided by student monitoring systems could be better utilised for pedagogical decision-making at the classroom and school level (Scheerens et al., 2012). Thus, both teachers and school leaders need more training and experience in using assessment data to inform instruction and resource allocation within the school. This opinion was expressed by nearly all the stakeholders interviewed.

Since 2010, two university research projects were developed to train teachers in using formative assessment results at the University of Twente (Project *Focus*) and the University of Groningen (Project *STREEF*). An analysis of the *STREEF* project (Doolaard, 2013) found that the training led to improved capacity of teachers regarding results-oriented work (compared to non-trained teachers) and that students benefitted from results-oriented approaches. However, the students' learning gains were not as large as expected. One of the reasons identified for the modest improvements in student learning was that many teachers had difficulties in moving from the analysis of student results to the differentiation of teaching practices. This suggests that further training should focus specifically on the interpretation and use of assessment results to adapt teaching strategies.

Strengthening formative assessment in daily classroom practice

Much important evidence on student learning needs can be obtained by teachers through ongoing formative assessment strategies in daily classroom interactions. Classroom-based formative assessment is essentially a pedagogical approach consisting of frequent, interactive checks of student understanding to identify learning needs, provide feedback to students and adapt teaching strategies (OECD, 2005). While medium- and long-term formative assessments (such as the regular use of LVS) are important for identifying areas of need and developing broad teaching strategies, it is short-cycle formative assessment happening in daily classroom interactions, which has the most direct and measurable impact on student achievement (Wiliam, 2006; Looney, 2011).

Although there seemed to be a general recognition of the efficacy of formative assessment in the Netherlands, classroom-based assessments are generally taken “off-the-shelf”, from instructional packages or purchased directly from test providers. The previous section described the strengths of student monitoring systems such as the LVS and their important role in results-oriented work in schools. At the same time, research

from different countries indicates that externally developed tests often do not fully address the information sought by the teacher to identify the learning difficulties of students. They typically can only cover a sub-set of the intended curriculum goals. Also, the data obtained from such assessments is typically not at the level of detail needed to diagnose individual student needs and identify strategies for future instruction (McGehee and Griffith, 2001; Rupp and Leseaux, 2006).

While regular monitoring tests can provide important clues about learning areas that need attention, other assessment strategies may be necessary to identify the causes of poor performance and develop an appropriate instructional intervention (OECD, 2013). In the Dutch context, an analysis of the *Focus* project (see above) found that information from the LVS provided useful feedback to teachers and students, but that additional diagnostic information was necessary to respond to individual student learning needs (Faber et al., 2013). This suggests that teachers could benefit from professional development helping them to select and design adequate diagnostic assessments to respond to the specific learning needs of their students. Getting involved in creating targeted assessments, and evaluating students' responses, can also be a powerful tool for teachers' professional development (Wiliam et al., 2004).

Concerns about the reliability of examination results

In secondary education, the flexibility in formats for the school-based examination component can result in better curriculum coverage, but at the same time, guidelines issued by the Inspectorate introduce other issues. In particular, the Inspectorate requires that school-based examination results be calibrated so that there is, at most, a small average difference (in the aggregate) between those scores and the scores on the central examinations. A large gap is noted as a problem by the Inspectorate. At least one function of this policy is to maintain a rough comparability in school examination results across the country. This is valuable given the flexibility schools and teachers have in developing and scoring these examinations. However, there may be genuine differences in the "objective" quality of students' performances on the two types of examinations that are virtually eliminated by the calibration.

A striking example of this problem can occur with modern languages where students' average productive language skills may differ systematically (higher or lower) from their average reading skills. Thus, schools with similar distributions of reading scores will be forced to have similar total scores, irrespective of meaningful differences in the distributions of scores on other language competencies. In the case of Dutch language, the introduction of reference levels is presumably meant to lead to criteria-referenced scoring, which may clash with the pseudo-norm-referenced scoring imposed by the calibration requirement. Specifically, in evaluating student work in the school-based examination, teachers should be guided by the reference levels that describe the expectations corresponding to different score levels. Thus, the average quality of, for example, the written work produced by a class may correspond to a particular reference level, based on a reasonable interpretation by the teacher of the meaning of the reference levels. However, if that average differs substantially from the class average on the central examination, then the class grades will need to be adjusted to meet the guidelines of the Inspectorate. One obvious consequence is that the adjusted grades will no longer accurately reflect the quality of the work submitted by the class and so interpretations based on these grades will be misleading to some degree.

A more general concern is related to the reliability of teachers' marks. Reliability refers to the consistency of the marks provided by independent graders for a set of papers. The literature is clear that achieving high levels of reliability when evaluating complex student work (e.g. essays) requires substantial training of raters, as well as ongoing monitoring and moderation of the scoring process (Gwet, 2012). In the Netherlands, although scoring guidelines for the secondary school examinations are centrally provided, training of raters (teachers) is done at the school level. Thus, it is important to continually review the reliability of teachers' grading in settings such as end-of-secondary examinations. Since these assessments carry high stakes for students, low reliability of teacher grading would pose challenges to the fairness of assessment.

Standardised assessments strongly influence the teaching programme

Reaching some degree of a consensus on national education goals will not be achieved in the short term. In the interim, education goals are being set with reference to existing assessments. One example is the drive to increase the proportions of students who reach the highest levels in examinations of Dutch language and mathematics. Although this approach certainly has some merit, it should be borne in mind that, in general, the examinations do not fully represent the target content, so that excessive pressure can lead to unintended consequences, such as an unwanted narrowing of the curriculum.

The previous section described the strengths of standardised assessments in the Netherlands, which provide reliable information on student learning in relation to key national learning objectives. At the same time, there is a risk that the high visibility of standardised assessment might lead to distortions in the education process. In the primary sector, the Cito tests use only multiple choice items, with all the accompanying constraints on what skills can be adequately measured. Because of the role of the test results in determining placement in the secondary sector, there is pressure on both students and teachers to concentrate on the specific content and format of the tests, which leads to a narrowing of the delivered curriculum. Excessive time spent on test preparation reduces the amount of content to which students are exposed.

Similar issues pertain to the assessments administered at the end of secondary education where, arguably, the pressures on students and teachers are even greater than in primary education. Although schools have substantial autonomy in developing the school-based part of the examination, teachers interviewed by the OECD review team indicated that they tended to align their own assessments to the formats used in the central examination. Ironically, this phenomenon may be an unintended consequence of the longstanding principle of freedom of education. As there is no national curriculum, the syllabi of the central examinations administered at the end of the primary and secondary education constitute a *de facto* national curriculum and, thus, ensure some degree of comparability across schools. The syllabi, along with features of the central examinations, therefore play a strong role in shaping instruction; particularly in the years when examinations are administered.

Concerns exist that assessment results are increasingly being put to multiple uses. For example, student scores on the tests given at the end of primary and secondary school were originally intended to identify the level of proficiency attained by the student to enter secondary education. However, these scores, aggregated to the school level, are also used as a key indicator of school quality, which increases the stakes associated with these scores for schools. This heightens the pressure on teachers to improve test performance

through a focus on the specific item formats employed by the test, and other strategies that may lack educational value. For example, some schools may adopt strategic approaches to boost the schools' test results, such as making large groups of students repeat the year before the examination year (Dutch Inspectorate of Education, 2013). Moreover, to the extent that these tests suffer from construct-irrelevant variance and/or construct underrepresentation¹¹ (Messick, 1989), their value as monitors of school quality is compromised. That is, there is a risk that the high stakes associated with the test lead to a narrowing of the delivered curriculum and, possibly, a shift of emphasis from other valued outcomes of schooling. The new policy of making the end-of-primary assessments mandatory makes their use for school accountability even more attractive and reinforces the risks of the test being put to multiple uses (more on this in Chapter 5).

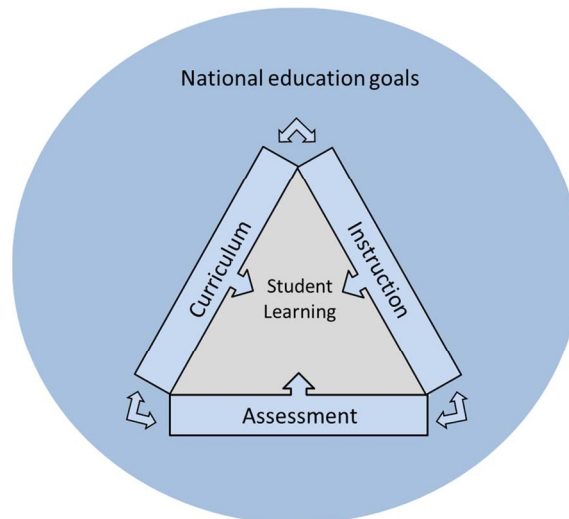
Policy recommendations

Student assessment is a critical component of the Dutch evaluation and assessment framework. The previous sections have outlined the many strengths of the Dutch approach and discussed a number of challenges related to student assessment in the Netherlands. Building on the identified strengths, this section suggests potential directions for further policy development in building a coherent framework for student assessment:

- Build consensus on key learning goals for the 21st century.
- Consider developing learning progressions to complement curriculum goals.
- Develop an assessment strategy corresponding to agreed education goals.
- Further develop the assessment infrastructure.
- Strengthen teacher professionalism for effective development and use of assessment.
- Support innovative assessment practices at the local level.
- Balance the use of assessment for improvement and accountability.
- Critically examine unintended consequences of assessment.

Build consensus on key learning goals for the 21st century

A simple model for education can be represented as a triangle with the vertices representing curriculum, instruction and assessment. At the centre of the triangle is “student learning”, with links to each vertex. A healthy system is one in which there is a both a good balance among the vertices, and meaningful coherence among the activities undertaken under the auspices of each vertex (Figure 3.1). However, when assessment results are used for accountability purposes, they can assume greater importance than originally intended. This is especially the case if curriculum and instruction are not properly grounded and well supported.

Figure 3.1 The “education triangle”

Source: Authors' own work

If student assessment is to be a tool for improving learning, rather than the driver of education in the Netherlands, it is critical that efforts are made to achieve a national consensus on the education goals for future generations. Such a consensus can then inform the work to be done at each vertex of the “education triangle”. It is likely that such a consensus will involve a rethinking of the content and attainment standards of the traditional academic competencies, as well as the adoption of some of the 21st century skills as important curricular goals.

There are a number of different documented approaches to describing 21st century skills (e.g. Rychen and Salganik, 2001; 2003; Partnership for 21st Century Skills, [n.d.]; National Research Council, 2012; Kyllonen, 2012; Kyllonen, forthcoming), for a summary see Annex 3.1 and OECD (2013). What they share is an appreciation of how important it is for students to develop the skills and dispositions that will enable them to both lead economically productive lives in a global economy, and to function as engaged citizens in a democracy. In the United States, for example, new standards for mathematics and English/language arts (Common Core State Standards, n.d.) have been adopted by almost all states, and new standards for science have just been published (Next Generation Science Standards, n.d.). One aspect of all three sets of standards is that they have a strong developmental focus; that is, they are informed by, and clearly reflect, decades of research on how students learn and how to appropriately characterise expertise at different age levels. This developmental focus will have important and, hopefully, beneficial implications for assessment.¹²

In the Netherlands, despite the challenges outlined in Chapter 2, there is continuing interest in the Ministry of Education, Culture and Science to encourage schools to focus on 21st century skills. This broader view of student development is also promoted by the Education Council. With this level of interest, it may be a good time to revisit this domain in a systematic fashion, with due regard to the perspectives of the different stakeholders. A policy proposal to move forward could: (i) argue for deeper learning in traditional academic subjects; (ii) highlight the importance of key skills and dispositions that transcend those subjects; (iii) call for all educators to have sufficient opportunity to

develop the professional competencies needed to support student learning; (iv) celebrate and support innovations in pedagogy and authentic assessment; (v) recognise the importance of engaging all students and having students take greater responsibility for their own learning.

Such a proposal, drawing on the emerging international consensus on 21st century skills, could serve as a starting point for a conversation that would lead the Netherlands to a next-generation education system. It also recognises not only the key role of educators, but also the necessity of providing the sustained support they require to appropriately implement an expanded curriculum and to help students attain higher levels of competence. At the same time, the innovations discussed above hold the promise of increasing student engagement and effort, increasing pedagogical effectiveness and the satisfaction that comes with teaching well.

Consider developing learning progressions to complement curriculum goals

The OECD Reviews of Evaluation and Assessment in Education found that specific learning progressions, describing the way in which students typically move through learning in each subject area, can further help to clarify national curriculum goals (OECD, 2013). Research-based learning progressions can provide a picture from beginning learning to expertise and help provide teachers, parents and other stakeholders with concrete images of what to expect in student learning, with direct links to the final learning objectives and reference levels. Such learning progressions can provide a clear conceptual basis for a coherent assessment framework, along with assessment tools that are aligned to different stages in the progressions.

Teachers can use these learning progressions as roadmaps to identify the set of skills and bodies of enabling knowledge that students must master en route to becoming competent in the more complex and multifaceted learning objectives defined for the end of primary and secondary education. The Ministry of Education, Culture and Science together with stakeholder groups could also facilitate the development of assessment criteria for rating different aspects of performance and exemplars illustrating student performance at different levels of proficiency. These can be used to define what constitutes adequate, good and excellent work and support professionals in clarifying quality definitions and making accurate judgements about student performance and progress in different curriculum areas.

Such materials can be promoted as voluntary resources that teachers use as signposts in their assessment. They can help raise aspirations and communicate a focus on excellence and continuous improvement. In line with the government's focus on teacher professionalism, such guidance could help teachers design their instructional plans and classroom assessment strategies in alignment with national objectives and progressions. Teachers should also be encouraged to share and co-construct intermediate learning goals and assessment criteria with students so that they understand different levels of work quality. Such common work on goals and criteria can promote both student learning and reflective teaching practice (Andrade, 2005; Jonsson and Svingby, 2007).

Develop an assessment strategy corresponding to agreed education goals

As a consensus on learning goals begins to emerge, work on a corresponding assessment strategy can begin. Given the novelty of 21st century skills for most teachers, formative assessment should be the primary focus, as it can contribute directly to improved learning. Ideally, a coherent set of formative assessments (across grades, within

a sector), along with the corresponding scoring rubrics and exemplars of student work, will help to provide illustrations of both the learning goals and the expectations for student performance. As noted by the OECD (2013), such support is essential to the constructive inclusion of broader competencies into the curriculum. Wherever possible, these skills should be incorporated into the existing curriculum so that they enhance the instruction and assessment of the core academic competencies and are not seen simply as an “add-on” to an already crowded schedule¹³. To help this, and where it is feasible and appropriate, teachers should use available technology for both teaching and assessment.

As curriculum changes are introduced, particularly those involving new and/or more advanced skills, they should be accompanied by an assessment strategy that is appropriately matched to the target competencies. An example is offered by the Technasium¹⁴, a grass-roots effort to strengthen the teaching of science through a more engaging, contextualised curriculum that incorporates problem-based learning. Allocating resources to provide assessment support would not only benefit the initiative, but also offer experience in how to develop an assessment system that incorporates more complex challenges. More generally, novel formats for student work may require new psychometric models and analysis strategies.

Exploratory projects can build the expertise needed when versions of these assessments are eventually brought to scale. In the Dutch context, one possibility is to begin by building measures of the “non-cognitive skills” that are relevant to tracking the progress of students with special educational needs (SEN) for whom the development of social skills, for example, may be as important as the development of traditional academic skills. Since most teachers will be unfamiliar with the frameworks for 21st century skills and the ways in which specific skills are defined and operationalised at different levels, they will need extensive professional development with respect to didactics (more on this below).

As mentioned in Chapter 2, the development of a broader set of learning goals, and the accompanying assessments, will have to involve multiple stakeholders. Given the concerns of many stakeholders regarding what they view as the increasing importance of central examinations, too early introduction of these skills into mandated summative assessments could trigger a backlash that would limit progress. Expanding the end-of-sector assessment frameworks to include more of the 21st century skills is likely to be problematic given the time and cost constraints under which they operate, and encouraging their adoption in the school examinations would run counter to the traditional autonomy accorded to schools. Thus, it seems more feasible for the Ministry to provide funding (perhaps through a revised *Schools have the Initiative* programme) for those schools and networks of schools already focussing on the teaching and assessment of 21st century skills to accelerate their work and, equally importantly, to provide regional and national forums where these schools can showcase their efforts to a broader audience. This, in turn, can lead to a national conversation on how traditional learning goals should be augmented to meet the challenges of the 21st century, and what systemic changes must be made to support teaching and learning of this broader set of competencies.

As suggested in Chapter 2, the Ministry of Education could start by commissioning a review of research and of strategies adopted by other education systems. A number of countries (e.g. Singapore, Costa Rica) are moving towards the introduction of problem-solving (in different forms) into national goals and assessment strategies in a systematic fashion, with strong central ministry support. The experiences of Finland and Australia

would also be especially valuable. As with the Netherlands, Finland's education system is highly decentralised and schools have substantial autonomy with respect to pedagogy and student assessment. Australia has a federal system that maintains somewhat greater central control, but has invested considerable resources in the assessment of these skills. Both countries have made considerable progress in raising the focus on 21st century skills (Adamson and Darling-Hammond, 2013).

Further develop the assessment infrastructure

Despite the absence of a new set of education goals, the Netherlands can be proactive in preparing for an uncertain future by building an assessment-related infrastructure that would prove useful in a variety of scenarios. Experimentation with the assembly, delivery and data management of new forms of technology-based assessment will provide the hands-on experience necessary if and when such assessments are introduced at the national level. One option to enhance the assessment technology infrastructure would be to further work on developing data information systems for schools, taking full advantage of the introduction of the unique student number and the requirement for all schools to implement student monitoring systems. Beyond providing real-time access to student data on attendance, enrolment, marks and schedules, such tools could serve as instructional management systems and provide planning tools, instructional materials and national-level performance standards. For example, some education jurisdictions in Australia have developed sophisticated data information systems, such as the School Measurement, Assessment and Reporting Toolkit (SMART) developed by New South Wales (Box 3.3).

Box 3.3 New South Wales, Australia: The School Measurement, Assessment and Reporting Toolkit

SMART provides information on national standardised assessment and state examinations. This information, together with information from school-based assessment activities provides a wealth of objective diagnostic information to which teachers can respond. The SMART package allows educators to identify areas for improvement as well as strengths in student performance. SMART also provides support through specific teaching strategies designed to improve student outcomes. This initiative has the potential to assist teachers in the instruction of their students, provide quick feedback to school agents, serve as a platform to post relevant instructional material to support teachers and improve knowledge management, operate as a network to connect teachers and schools with similar concerns, and create a better data infrastructure for educational research.

Source: Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.

It would also be helpful to initiate work on building expert systems to automatically evaluate some types of student work. Research literature shows that expert systems have been developed to grade essays, short answer questions in various disciplines, mathematical expressions, problem-solving strategies and graphical responses of various types (Williamson et al., 2006).¹⁵ As technology-based assessment becomes more ubiquitous, the capacity to score student work automatically reduces overall costs and shortens the time lag between administration and score reporting. On a deeper level,

designing assessments to facilitate automated scoring can exert a useful discipline that increases comparability and validity (Bejar and Braun, 1994).

In many ways, the Netherlands is in a good position regarding technology-based assessment infrastructure, as most schools already have the capacity to administer tests by computer. Going further, the next steps could involve reflecting about a national strategy to expand and upgrade the infrastructure to support high-speed internet access (see Box 3.4 for an example from the United States). The strategy will have to take into account the rapid development of the internet backbone, as well as the proliferation of devices such as tablet computers. In addition, it will have to establish priorities (e.g. between the primary and secondary sectors) and determine the funding mechanism (e.g. cost-sharing between the national government and local boards). Such a strategy should be led by the Ministry in conjunction with key stakeholders, such as the Primary and Secondary Education Councils. Input from the Ministry of Finance might also be useful. Because of the technical and financial complexities, it would be advisable to have a set of white papers prepared in advance. These papers should describe the different purposes that the next-generation infrastructure would serve, the current state of the infrastructure, the various options for development with their advantages and disadvantages (with respect to cost and timing).

Box 3.4 United States: The National Education Technology Plan

The United States Department of Education recently released a document titled National Education Technology Plan (NETP) that describes goals in five areas: Learning, teaching, assessment, infrastructure and productivity. The plan focuses on how the thoughtful use of technology can contribute to advances in all areas. It builds on earlier initiatives including: (i) The E-Rate programme (www2.ed.gov/about/offices/list/oii/nonpublic/erate.html) that provides eligible schools and libraries with discounted telecommunications charges that enhance internet access and (ii) the State Educational Technology Grants programme (www.ed.gov/edblogs/technology/grants/) that provides funding for various initiatives related to the development and implementation of education technology. Currently, one of the main goals is the provision of technology-related services to support students with disabilities. Another grant programme (<http://nces.ed.gov/programs/slds/>) provides funding for statewide longitudinal data systems that helps states build and expand their data systems to collect, organize and analyse longitudinal student records for policy-related analyses and to facilitate the construction of more sophisticated, test-based indicators for use in school and teacher evaluations.

Source: U.S. Department of Education (2010), *National Education Technology Plan 2010*, available at: www.ed.gov/technology/netp-2010.

Strengthen teacher professionalism for effective development and use of assessment

In addition to further developing the assessment infrastructure, it is equally important to continue to build assessment expertise, including the capacity to use results for improvement, among both teachers and school leaders. The OECD Reviews of Evaluation and Assessment in Education found that teacher professionalism in assessment is key to developing balanced and effective assessment frameworks (OECD, 2013).

Combining teacher-based and external assessments can help ensure maximum validity and reliability in assessment. While learning outcomes that can be readily assessed in external examinations should be covered this way, continuous teacher-based assessment can cover a broader range of complex learning outcomes (Crooks, 2004). Due to its continuous nature, teacher-based assessment allows for important achievements to be measured that are more difficult to capture in an external examination, such as extended projects, practical assignments or oral work. However, in order to reach the full potential of teacher-based assessment, it is important for policy makers and stakeholders to adopt a strategic approach to teacher learning in assessment and invest in professional development opportunities. To be able to assess students' progress in developing complex competencies, it is important that teachers learn to select and/or develop a variety of assessment approaches and understand different aspects of validity, including what different assessments can and cannot reveal about student learning (OECD, 2013).

To this end, assessment capacity, including the capacity to use results for improvement, should be reflected in teacher standards and be addressed in a coherent way across teacher preparation programmes and publicly funded professional development programmes. Eventually, assessment-related competencies should become part of the teacher registration system and teacher appraisal approaches (Chapter 4). The human capital development agenda sketched out below will also require professional development of teacher educators and of providers of in-service teacher training. Fortunately, teacher education programmes can draw on the rapidly expanding resources available internationally. Again, inducements and support from the Ministry will be essential in this regard. The above analysis points to three priority areas for further capacity development: *(i)* classroom-based formative assessment; *(ii)* interpretation and use of assessment results for improvement; *(iii)* reliable summative assessment and marking of examinations.

Capacity for classroom-based formative assessment

The current focus in the Netherlands on using regular student monitoring systems for results-oriented work is commendable. Such medium- and long-term formative uses of results are important for identifying areas for further improvement, developing broad teaching strategies to address needs identified within the student cohort, planning, allocation of resources, and so on. It can also feed into the school-wide coordination of pedagogical support and remediation for students facing learning difficulties.

While medium- and long-term formative assessment strategies are important to ensure consistency of support throughout a student's learning trajectory, research indicates that short-cycle formative assessment – the daily interactions between and among students and teachers – has the most direct and measurable impact on student achievement (Looney, 2011). In short-cycle interactions, formative assessment is part of the classroom culture, and is seen as an integrated part of the teaching and learning process. Teachers systematically incorporate such formative assessment methods in their course planning – for example, in how they intend to develop classroom discussions and design activities to reveal student knowledge and understanding. These interactions encompass effective questioning to uncover student misconceptions and identify patterns in student responses, feedback on student performance and guidance on how to close learning gaps, and student engagement in self- and peer-assessment (OECD, 2013).

The active participation of students in such formative assessment processes has given rise to the term *assessment as learning*, which focuses on students reflecting on and

monitoring their own progress to inform future learning (Earl, 2003). While feedback by teachers and others provides information that can help students improve, it is the students themselves who must make sense of that information, relate it to prior knowledge and take action to close gaps in their own learning. But developing skills for self-assessment and self-regulation takes time and requires structured support by teachers in the classroom. Teachers can use classroom assessment to provide opportunities for students to engage in reflection and critical analysis of their own learning, for example by guiding students in setting learning goals and monitoring their progress towards them; working with them to develop criteria to judge progress; using exemplars and models of good practice and questioning of their own thinking and learning processes. Policy makers can support such practices by developing requirements, guidelines and support regarding learner-centred teaching and assessment (OECD, 2013).

Capacity for interpreting and using assessment results for improvement

If student assessment is to improve future learning, then assessment evidence must be acted upon in subsequent classroom practice, to provide the right levels of support and challenge to each student. In order to successfully implement differentiated instruction, teachers require relevant professional development both to make the best use of the evidence collected, and to better manage a classroom in which multiple learning activities are taking place. The investment appears to be worthwhile as the research literature documents important learning gains of students exposed to high quality differentiated instruction informed by relevant formative assessment (Black et al., 2003; Wiliam et al., 2004).

Competence in designing or selecting an assessment, interpreting the results and using the results effectively should become a key goal of teacher preparation. This can be accomplished, for example, by employing the inquiry cycle¹⁶ as a fundamental didactic method. Undoubtedly, this will require changes in the curriculum of the teacher training programmes (White and Fredericksen, 1998). Currently, according to representatives of teacher education institutions interviewed by the OECD review team, teacher preparation does not allow sufficient time for assessment-related competencies, but there is an effort in some programmes to improve in this area, despite the severe time constraints under which they operate. Practicing teachers should be encouraged to use some of their professional development time to develop a range of assessment skills, appropriate to the subject and grade. Resources can be provided locally, regionally, or online and supported by various networks.

Capacity for reliable summative assessment and marking of examinations

Teachers play an important role in summative assessments. At the end of secondary schooling, not only do teacher-made examinations carry roughly equal weight to the central examinations, but teachers also grade the open-ended responses in the central examinations. Thus, it is important that teachers carrying out these tasks receive sufficient training so that the quality of the teacher-made examinations is high, and that the scores assigned to responses in the central examinations are reliable and valid. Ideally such training should begin during the period of teacher preparation. Exemplars of student work at different levels of proficiency and in different year levels can be used both to provide an example of the expectations at those levels/grades, and to provide opportunities for would-be teachers to develop their own expertise. However, such a redesign of teacher preparation necessarily involve systemic changes that are challenging to implement because of political and bureaucratic obstacles, as well as faculty resistance (Mandinach

and Gummer, 2013). Similar activities can be conducted for in-service teachers through specialised workshops delivered in-person or online.

During the grading of end-of-course or end-of-sector examinations, in addition to the CVE's open telephone line for teachers, consideration should be given to offering online training, and strengthening moderation processes to carry out real-time auditing of teachers' marking. Moderation involves strategies for quality assurance of assessment judgements, such as teachers cross-marking each other's assessments within a school or across schools, teachers discussing samples of student work in groups or in collaboration with experienced graders, or a competent authority or examination board externally checking school-based assessments. The objective is to reduce variations in the ways teachers assess students and set marks, in order to achieve fairness in student assessment and reporting. The Education Cooperative has made some progress in this direction and their work should be supported.

While moderation is a key strategy in validating the consistency of teacher judgement, it also involves professional discussions between teachers about the quality of authentic exemplars of student work, and as such it has the potential to provide a powerful professional learning opportunity for teachers that they can apply directly in their classroom practices. This contributes to improving teachers' professional judgements about student work and developing a shared understanding of marking criteria or standards within schools and between schools (Timperley et al., 2008). Research suggests that collaborative discussions centered on student work can lead to improved pedagogy and greater gains in student achievement (Black et al., 2002, 2003; Hargreaves and Fullan, 2012). Teachers are provided with a chance to reflect on assessment in their subject, with respect to both topics and performance criteria. From a strategic perspective, as teachers develop this expertise they will have a stronger base on which to build the more sophisticated competencies that will be required for next generation assessments. Box 3.5 offers examples on moderation procedures in place in Queensland (Australia) and New Zealand.

Box 3.5 Moderation of examinations in upper secondary education in Queensland, Australia and New Zealand

In **Queensland, Australia** the upper secondary examination system is school-determined and based, but achievement standards and scoring are externally moderated. Moderation processes for the Senior Certificate (Year 12) involve subject-based panels of expert teachers providing advice to schools on the quality of their assessment programme and their judgements of quality of student performance based on sample portfolios. The system involves follow-up with schools where panels identify issues regarding assessment and standards. There is negotiation of the final results to be recorded on the Senior Certificate (Sebba and Maxwell, 2005 in Santiago et al., 2011). Similarly, procedures adopted by educational jurisdictions and particular schools for moderating internal summative teacher judgements (so-called A-E ratings) also facilitate common understanding of year level proficiency standards and foster the development of professional learning communities that can provide crucial support for improving opportunities for student learning and building teacher capacity.

Box 3.5 Moderation of examinations in upper secondary education in Queensland, Australia and New Zealand (*continued*)

In **New Zealand**, an external moderation system is also in place to ensure the dependability of internal assessments in Years 11-13. The New Zealand Qualifications Authority (NZQA) directly checks the quality of internal assessment through a sampling approach. Schools are required to submit ten percentage points of internally assessed student work for NZQA moderation to make sure the assessment is appropriately aligned with standards. The moderation process does not affect the marks assigned to assessment samples by teachers, but is intended to provide feedback to teachers and to inform future assessment policy development at the system level.

Sources: Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.

Nusche, D., et al. (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116917-en>.

Support innovative assessment practices at the local level

As noted in the “Strengths” section, the Netherlands is rich in assessment expertise. That expertise is concentrated on the design, development and analysis of current assessments. However, it is necessary to expand that expertise in directions that will facilitate the introduction of the new forms of assessment demanded by new curricular goals. One approach is to fund small-scale innovation projects (perhaps on a competitive basis) that will lead to pilot administrations of these new assessments.¹⁷ These projects should involve partnerships between testing organisations and groups of schools where the pilots would take place. Ideally, some of the educators in these schools would participate in the research and development effort, and not simply serve as proctors during administration.

Initially, these assessments would generate data for formative purposes and could even be embedded in instruction. For example, in science classes, pedagogy centred on micro-worlds, and other kinds of simulations, could incorporate tasks that require higher order critical thinking, collaboration and effective communication, which are all 21st century skills. Such tasks would challenge both students and teachers, and yield new types of data that could serve as a focus for professional development. The design of these more complex assessments requires more sophisticated design strategies, of which Evidence Centered Design (ECD)¹⁸ is the best known (Mislevy et al., 2006).

In addition, the Inspectorate should consider strategies for accommodating and even rewarding schools’ innovative assessment practices, especially if there is reasonable evidence that they are supporting student learning. Central examinations act as an anchor in the accountability system, which is valuable in a system where educators are accorded so much freedom and flexibility. At the same time, anchors impede movement, so understanding when and how to “raise the anchor” will be important in not stifling innovation. Consequently, the Inspectorate should experiment with different ways to balance central and local results, especially in schools that meet the basic requirements.

More generally, the Inspectorate could favourably recognise schools instituting programmes of professional development that highlight improved assessment practice as a key target. Assessment practice comprises developing tests, grading student work and using the results to inform instructional practice. The Inspectorate has a particular interest in the accuracy of the grading of the end-of-sector examinations, so by recognising those schools that institute some sort of quality control mechanism in the grading process, the Inspectorate can encourage all schools to move in this direction.

The success of the initiatives described above depends, in part, on the existence of networks of educators that can facilitate collaboration and dissemination of new approaches and best practices in assessment (as well as other aspects of pedagogy and didactics). The Ministry should encourage, support and sustain such networks, working through existing organisations (e.g. the Education Council, the College for Examinations, Primary and Secondary Education Councils, the Education Co-operative). Some support could be channelled through the *Schools have the Initiative* programme. Participation would be voluntary, but incentives could be used at the start with the reasonable expectation that the value of being involved in one or more such networks would quickly become evident. Participation by some faculty in each teacher training institution should also be encouraged. Box 3.6 provides an example of a learning network in Norway.

Box 3.6 Norway: Centrally-supported networks on assessment

In Norway, the Ministry of Education and Research and the Directorate for Education and Training identified formative assessment as a priority area for education policy and professional development and launched a range of support programmes and learning networks at the regional, local and school level. For example, the *Assessment for Learning* programme (2010-14) is organised in learning networks at the local and regional level, where practitioners can exchange experiences and create spaces for common reflection on effective practice. Participating municipalities and counties employ a formative assessment contact person who assists in running the project locally. These contact persons attend *Assessment for Learning* workshops run by the national Directorate for Education and Training. The programme also provides online resources including tools and videos on how to enact effective formative assessment in the classroom.

Source: Nusche, D., et al. (2011), OECD Reviews of Evaluation and Assessment in Education: Norway 2011, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, <http://dx.doi.org/10.1787/9789264117006-en>.

Balance the use of assessment for improvement and accountability

Eventually, a coherent system of assessments as, for and of learning would emerge. Some assessment results will be used directly for improving teaching and learning, while other assessment data would contribute indirectly through informing accountability. In principle, these paths should complement one another, although in practice that may not be the case. As Hargreaves and Braun (2013) point out,

Although educational accountability is meant to contribute to improvement, there are often tensions and sometimes direct conflicts between the twin purposes of improvement and accountability. These are most likely to be resolved when there is collaborative involvement in data collection and analysis, collective

responsibility for improvement, and a consensus that the indicators and metrics involved in [data-driven improvement and accountability] are accurate, meaningful, fair, broad and balanced. When these conditions are absent, improvement efforts and outcomes-based accountability can work at cross-purposes, resulting in distraction from core purposes, gaming of the system and even outright corruption and cheating.

In the Netherlands, the Inspectorate must play a critical and, perhaps even a leading, role in harmonising improvement and accountability. In collaboration with the Ministry, the Primary and Secondary Education Councils and the College for Examinations, it should conduct a thorough and critical review of the indicators now employed for school evaluation. The goal should be to determine whether and how current evaluation practices may be impeding innovation, and to consider how to broaden the set of indicators to better reflect school practices that contribute to student development but are not now recognised. For example, schools should be recognised for incorporating 21st century skills into the curriculum, for the effective use of formative assessments, and for accelerating student progress towards competence standards.

Although didactics may lie outside the direct responsibility of the Inspectorate, evaluating the culture of pedagogy and collective responsibility within a school and, where appropriate, within larger school boards, would provide an impetus for school leaders to move in this direction. Assessment vendors such as Cito and others, under the leadership of the College for Examinations, could develop a framework for a coherent system of formative and summative assessments that more fully address the broader range of learning goals. With sufficient foresight and planning it should be possible to devise an assessment strategy that will both support continual improvement in education processes and assure the public that its investments in education are well spent. Box 3.7 provides examples of communication strategies in Canada and New Zealand.

Box 3.7 Defining and communicating balanced assessment strategies in Canada and New Zealand

In **Canada**, the *Principles for Fair Student Assessment Practices for Education in Canada* outline key elements for assessment practice that have served as the foundation for teacher handbooks, board policies and departments of education policy documents on assessment and test development in all Canadian jurisdictions. The Principles were developed in response to assessment practices that were judged as inappropriate for Canadian students. These principles and guidelines, intended for both assessment practitioners and policy makers, identify the issues to be taken into account in order that assessment exercises to be deemed fair and equitable. The text acts both as a set of parameters and a handbook for assessment. The first part deals with developing and choosing methods for assessment, collecting assessment information, judging and scoring student performance, summarising and interpreting results, and reporting assessment findings is directed towards practising teachers and the application of assessment modes in the classroom setting. The second part is aimed at developers of external assessments such as jurisdictional ministry/department personnel, school boards/districts, and commercial test developers. It includes sections dealing with developing and/or selecting methods for assessment, collecting and interpreting assessment information, informing students being assessed, and implementing mandated assessment programmes (for more information, see: www2.education.ualberta.ca/educ/psych/crame/files/eng_prin.pdf).

Box 3.7 Defining and communicating balanced assessment strategies in Canada and New Zealand (continued)

The **New Zealand** Ministry of Education *Position Paper on Assessment* (2010) provides a formal statement of its vision for assessment. It describes what the assessment landscape should look like if assessment is to be used effectively to promote system-wide improvement within, and across, all layers of the schooling system. The paper places assessment firmly at the heart of effective teaching and learning. The key principles highlighted and explained in the paper are: the student is at the centre; the curriculum underpins assessment; building assessment capability is crucial to achieving improvement; an assessment capable system is an accountable system; a range of evidence drawn from multiple sources potentially enables a more accurate response; effective assessment is reliant on quality interactions and relationships. To support effective assessment practice at the school level, the Ministry of Education is also currently conducting an exercise which maps existing student assessment tools. The purpose is to align some of the assessment tools to the National Standards and provide an *Assessment Resource Map* to help school professionals select the appropriate assessment tool to fit their purpose.

Sources: Fournier, G. and D. Milton (forthcoming), OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Canada, prepared for The Council of Ministers of Education, Canada (CMEC).

New Zealand Ministry of Education (2010), Ministry of Education Position Paper: Assessment [Schooling Sector], Wellington.

Nusche, D., et al. (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116917-en>.

Critically examine unintended consequences of assessment

The preceding subsections focused on preparing for the medium-term and long-term future. However, there are opportunities to leverage assessment data that is currently being generated and to critically examine current practices that may impede innovation and improvement. For example, greater effort should be devoted to examining the dynamic responses of the education system to different forms and uses of assessment; that is, “What are the intended and unintended consequences of assessment in the different sectors”? A better understanding of these consequences, and particularly how they vary with context, can help to shape assessment policy for the future (De Wolf and Janssens, 2007).

This brings to the fore the problem of “evidential asymmetry” (Hargreaves and Braun, 2013). This term refers to the fact that the intended consequences of a policy are likely to be related to indicators that are generally collected as a normal output of the system (the obvious examples are test scores and graduation rates). On the other hand, unintended (usually negative) consequences are likely to be related to indicators that are not generally collected (examples are changes to the curriculum, excessive test preparation, and manipulation of student enrolment).

Documenting these consequences, therefore, involves conducting focused studies that require special funding and often take considerable time. Thus, there is a substantial time lag between the reporting of the intended and unintended consequences. Again, conducting such studies during the present regime will help to inform education policy with respect to the proposed new curricula and the accompanying assessment and accountability systems.

Notes

- ¹ <http://project-focus.gw.utwente.nl>
- ² <http://www.project-streef.nl/>
- ³ http://www.cito.nl/nl/onderwijs/primair%20onderwijs/cito_volgsysteem_po.aspx
- ⁴ <http://www.parnassys.nl/>
- ⁵ <http://www.rovict.nl/?main=esiswebbasedinfo>
- ⁶ These alternatives include: Onderwijsadvies.nl, Bureau ICE, A-Vision and Boom test uitgevers.
- ⁷ The results from the end-of-sector examinations constitute one of four indicators for Quality Aspect 1 (“The outcomes are at the level that may be expected on the basis of the characteristics of the student population”) of the inspection framework for secondary education (Dutch Inspectorate of Education, 2009). The four indicators for Quality Aspect 1 are: (1.1) In the lower years, students reach the educational level that may be expected; (1.2) Students need little extra time to complete the second stage of their programme; (1.3) In the national examination, students attain the marks that may be expected; (1.4) The differences between marks obtained in the school examination and those obtained in the national examination are at an acceptable level.
- ⁸ With computer-adaptive testing employing a large pool of items, nearly every student is administered a unique sequence of items. Nonetheless, using item response theory, the results can be put on a common scale just as they would if all students had taken the same test.
- ⁹ Tests incorporating such blocks of items are termed “self-auditing” since they provide evidence as to whether scores are affected by sources of construct-irrelevant variance.
- ¹⁰ It is noteworthy that this network, the Technasium programme, is based on a grass-roots initiative.
- ¹¹ Construct-irrelevant variance refers to sources of differences in student test performance that are unrelated to what the test is attempting to measure. Construct underrepresentation refers to the gap between the intended curriculum (content standards) and what the test actually measures. Test validity is greater as construct-irrelevant variance and construct underrepresentation are reduced.
- ¹² Two multi-state consortia have received funding from the United States Department of Education to build assessment batteries in mathematics and English/Language arts tightly aligned to the Common Core State Standards. The consortia are the Partnership for the Assessment of Readiness for College and Career (PARCC) and the Smarter Balanced Assessment Consortium (SBAC). Two other consortia are charged with the task of developing assessment batteries with especially severe or multiple disabilities.
- ¹³ One approach to such integration has been proposed by Braun (1996).
- ¹⁴ This is a nationally recognised programme created by a private agency and adopted in an increasing number of Dutch secondary schools. See the website of the programme here: <http://www.technasium.nl>.

- ¹⁵ This is an area of ongoing research. For essay scoring see for example <http://www.measurementinc.com/sites/default/files/ASAP-PH1-PressRelease-12April2012.pdf>.
- ¹⁶ There are many versions of an inquiry cycle: They all begin with posing a problem or issue to the student that engages her interest. The student alone, or in a group, explores the background to the problem, investigates possible solutions or explanations and presents a draft of the work. The teacher or teachers examine the work and offer feedback that students use to refine their work. This can be repeated as necessary until a final submission is made. The power of the inquiry cycle lies in the balance of responsibility between student and teacher, the opportunity for teachers to intervene constructively in the process of learning through repeated examination of student work, and how reflecting on these experiences provides teachers with a basis for improving their practice – particularly if this is done in the context of a collaborative community of professionals.
- ¹⁷ Presumably this could be accomplished in the same manner as the current pilot projects on the use of value-added models.
- ¹⁸ ECD is based on the simple premise that the best way to obtain the desired information from a test is to design it from the outset with that focus. Specifically, it begins with a clear articulation of the claims and inferences (about student proficiencies) that are to be made. Those are derived both from the nature of the subject matter and the uses to be made of the results. One then asks, “What sort of evidence is required to make those claims and inferences in a defensible manner”? The succeeding question is, “What sorts of items, tasks, or challenges are needed to generate that evidence”? With the (tentative) answers to those questions in hand, it is possible to draft a preliminary test design that also takes into account the constraints of time and cost. In almost all cases, various compromises are made to achieve a feasible design. Nonetheless, with a clear idea of what evidence may be lost as a result of certain decisions, the final design is usually superior to one built in the traditional manner.

References

- Adamson, F. and L. Darling-Hammond, (2013), *Policy Pathways for 21st Century Skills*, http://atc21s.org/wp-content/uploads/2013/02/ATC21S_Policy_Paper-20120730.pdf.
- Andrade, H. (2005), “Teaching with Rubrics: The Good, the Bad, and the Ugly”, *College Teaching*, Vol. 53, No. 1, pp. 27-30.
- Bejar, I.I. and H.I. Braun, (1994), “On the synergy between assessment and instruction: Early lessons from computer-based simulations”, *Machine Mediated Learning: An International Journal*, 4(1), 5-25.
- Black, P.J., et al. (2002), *Working inside the black box: Assessment for learning in the classroom*, King’s College London School of Education, London.
- Black, P.J., et al. (2003), *Assessment for learning: Putting it into practice*, Open University Press, Buckingham, United Kingdom.
- Braun, H.I. (1996), “A school-based strategy for achieving and assessing work-readiness skills”, in L. B. Resnick and J. G. Wirt (eds.), *Linking School and Work - Roles for Standards and Assessment*, pp. 223-241, Jossey-Bass, San Francisco, California.
- Common Core State Standards Initiative (CCSS) (n.d.), www.corestandards.org.
- Crooks, T.J. (2004), “Tensions between assessment for learning and assessment for qualifications”, paper presented at the Third Conference of the Association of Commonwealth Examinations and Accreditation Bodies (ACEAB), Fiji, March 2004.
- De Wolf, I.F. and F.J.G. Janssens (2007), “Effects and side-effects of inspections and accountability in education: An overview of empirical studies”, *Oxford Review of Education*, pp. 1-18, Routledge, London.
- Doolard, S. (2013), *Effecten van het trainings- en begeleidingstraject ‘Streef: Gebruik maken van opbrengsten’*, Rapportage BOPO-programmalijn Onderwijskwaliteit PO, GION, Gronings Instituut voor Onderzoek van het Onderwijs, Groningen.
- Dutch Inspectorate of Education (2009), *Risk-Based Inspections as of 2009: Primary and Secondary Education*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2010), *Opbrengstgericht werken in het basisonderwijs: Een onderzoek naar opbrengstgericht werken bij rekenen-wiskunde in het basisonderwijs*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2013), *The State of Education in the Netherlands: Highlights of the 2011/2012 Education Report*, Dutch Inspectorate of Education, Utrecht.
- Earl, L. (2003). *Assessments as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, Corwin Press, California.

- EPPI (2002), *A Systematic Review of the Impact of Summative Assessment and Tests on Students' Motivation for Learning*, Review conducted by the Assessment and Learning Synthesis Group, EPPI, London.
- Faber, M., M. van Geel and A. Visscher (2013), *Digitale Leerlingvolgsystemen als basis voor Opbrengstgericht werken in het Primair Onderwijs: Een analyse van de wijze waarop scholen en besturen de mogelijkheden van digitale leerlingvolgsystemen kunnen benutten*, Universiteit Twente, Twente.
- Gwet, K. L. (2012) *Handbook of Inter-rater Reliability* (3rd ed.), Advanced Analytics, LLC, Gaithersburg, Maryland.
- Hargreaves, A. and H.I. Braun (2013), *Data-driven improvement and accountability*, Policy brief, National Education Policy Center, Boulder, Colorado.
- Hargreaves, A. and M. Fullan (2012), *Professional capital: Transforming teaching in every school*, Teachers College Press, New York.
- Janssens, F.J.G., L. Rekers-Mombarg, E. Lacor (2014), *Leerwinst en toegevoegde waarde in het primair onderwijs: Eindrapportage*, available at: http://www.rijksoverheid.nl/ministeries/ocw/documenten-en_publicaties/kamerstukken/2014/03/07/aanbieding-eindrapport-pilot-leerwinst-en-toegevoegde-waarde-primair-onderwijs.html.
- Jonsson, A. and G. Svingby (2007), “The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences”, *Educational Research Review*, 2, pp. 130-144.
- Koretz, D. and A. Beguin (2010), “Self-monitoring assessments for educational accountability systems”, *Measurement: Interdisciplinary Research and Perspectives*, 8(2-3), 92-109, Routledge, London.
- Kyllonen, P.C. (2012), *Measurement of 21st Century Skills within the Common Core State Standards*, Educational Testing Service, Austin, Texas, www.k12center.org/rsc/pdf/session5-kyllonen-paper-tea2012.pdf.
- Kyllonen, P.C. (forthcoming), “Designing tests to measure personal attributes and noncognitive skills”, in S. Lane, M. Raymond and T. Haladyna, (eds.), *Handbook of Test Development*, 2nd edition. London: Routledge.
- Looney, J. W. (2011), “Integrating Formative and Summative Assessment: Progress Toward a Seamless System?”, *OECD Education Working Papers*, No. 58, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5kgx3kbl734-en>.
- Madaus, G., M. Russell and J. Higgins (2009), *The Paradoxes of High Stakes Testing*, Information Age Publishing, Charlotte, North Carolina.
- Mandinach, E.B. and E.S. Gummer (2013), “A Systemic View of Implementing Data Literacy in Educator Preparation”, *Educational Researcher*, Vol. 42 No. 1, pp. 30–37, available at: <http://edr.sagepub.com/content/42/1/30.full.pdf+html>
- McGehee, J.J. and L.K. Griffith (2001), “Large-scale assessments combined with curriculum alignment: Agents of change”, *Theory into Practice*, Vol. 40, pp. 137-144.
- Messick, S. (1989), “Validity”, in R. L. Linn (ed.), *Educational Measurement (third edition)*, American Council on Education and Macmillan Publishing, New York.
- Mislevy, R. J., et al. (2006), “Concepts, terminology, and basic models of evidence-centered design” in D. M. Williamson, I.I. Bejar and R. J. Mislevy (eds.), *Automated*

- scoring of complex tasks in computer-based testing* (pp. 15-47), Erlbaum, Mahwah, New Jersey.
- National Research Council (2012). *Education for Life and Work: Developing Transferable Knowledge and Skills in the 21st Century*. J.W. Pellegrino and M.L. Hilton (eds.), The National Academies Press, Washington DC.
- Next Generation Science Standards (NGSS) (n.d.), Washington DC, www.nextgenscience.org.
- Nusche, D., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Norway 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, <http://dx.doi.org/10.1787/9789264117006-en>.
- Nusche, D., et al. (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116917-en>.
- OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264190658-en>.
- OECD (2005), *Formative Assessment: Improving Learning in Secondary Classrooms*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264007413-en>.
- Partnership for Assessment of Readiness for College and Careers (PARCC) (n.d.) www.parcconline.org.
- Partnership for 21st Century Skills (n.d), *Framework for 21st century learning*, www.p21.org/our-work/p21-framework.
- Rupp, A.A. and N.K. Lesaux (2006), “Meeting expectations?: An empirical investigation of a standards-based assessment of reading comprehension”, *Educational Evaluation and Policy Analysis*, Vol. 28, No. 4, pp. 315-333.
- Rychen, D.S. and L.H. Salganik (eds.) (2001), *Defining and selecting key competencies*, Hogrefe and Huber, Gottingen, Germany.
- Rychen, D.S. and L.H. Salganik (eds.) (2003), *Key competencies for a successful life and a well-functioning society*, Hogrefe and Huber, Gottingen, Germany.
- Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.
- Scheerens, J., et al. (2012), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands*, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.
- Shewbridge, C., et al. (2010), *OECD Reviews of Migrant Education: Netherlands 2010*, OECD Reviews of Migrant Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264086210-en>.
- Smarter Balanced Assessment Consortium (SBAC) (n.d.) www.smarterbalanced.org, accessed 29 January 2014.
- Timperley, H., et al. (2008), *Teacher Professional Learning and Development: Best Evidence Synthesis Iteration (BES)*, New Zealand Ministry of Education, Wellington.

- U.S. Department of Education (2010), National Education Technology Plan 2010, available at: www.ed.gov/technology/netp-2010.
- Van Barneveld, C. (2008), “Using data to improve student achievement”, What Works? Research into Practice, Research Monograph No. 15, The Literacy and Numeracy Secretariat, Ontario, www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/Using_Data.pdf.
- Visscher, A.J. and M. Ehren (2011), De eenvoud en complexiteit van opbrengstgericht werken, Vakgroep Onderwijsorganisatie en –management, Universiteit Twente, Enschede, Netherlands.
- Wainer, H. (ed) (2000), Computer Adaptive Testing: A Primer (2nd edition), Lawrence Erlbaum Associates, Mahwah, New Jersey.
- White, B.Y. and J. Frederiksen (1998), “Inquiry, modeling and metacognition: Making science accessible to all students”, *Cognition and Instruction*, 16(1), 3-118, Routledge, London.
- William, D., et al. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education: Principles, Policy and Practice*, 11(1), 49-65, Routledge, London.
- William, D. (2006), “Formative assessment: Getting the focus right”, *Educational Assessment*, Vol. 11, pp. 283-289.
- Williamson, D.M., I.I. Beja and R.J. Mislevy (eds.) (2006), Automated scoring of complex tasks in computer-based testing (pp. 15-47), Erlbaum, Mahwah, New Jersey.

Annex 3.1

Research on 21st century skills

The volumes edited by Rychen and Salganik (2001; 2003) describe DeSeCo, a multi-year project organised under the auspices of the OECD, that “was initiated to provide solid theoretical and conceptual foundations for the broad range of competencies that are needed to face the challenges of the present and the future” (Rychen and Salganik, 2003, p.vii-viii). Calling on an international group of experts from many different disciplines and perspectives, they developed an overarching tri-partite framework: interacting in socially heterogeneous groups, acting autonomously and using tools interactively. Equally important, the report by Rychen and Salganik (2003) includes discussion of the implications of this framework for building the corresponding assessments and national indicators, as well as “the challenges associated with the development of a coherent, long-term assessment strategy based on the theoretical and conceptual foundations provided by DeSeCo” (Rychen and Salganik, 2003, p. 12).

The Partnership for 21st Century Skills (P21) was founded in 2002 in the U.S. and constitutes a joint effort by the Federal government, many state governments and NGOs, and leading corporations to develop and promote the skills necessary for individual and national success in a global economy. The framework augments the core academic skills with four others: critical thinking and problem solving; communication, collaboration, creativity and innovation. P21 works with educators to incorporate the development of these skills into the curricula of all disciplines. Subsequently, the National Research Council of the National Academy of Sciences (U.S.) convened a task force to clarify and organize the vast literature on 21st century skills, as well as to offer recommendations on how these competencies can be developed in school settings and how student progress can be assessed for both formative and summative purposes. The resulting report (National Research Council, 2012) constitutes a valuable foundational document for any country wishing to explore policy alternatives in this area.

Building on recent progress in defining key competencies, academics and testing organizations have been developing assessments for certain 21st century skills, and refining them to achieve acceptable levels of reliability and validity. The work of Kyllonen and his associates at the Educational Testing Service (Kyllonen, 2012; Kyllonen, in press) exemplifies the high quality of this work and the nature of the information obtained from the administration of these assessments. The crucial point is that there has been substantial progress in this arena and, coupled with the introduction of computer-based assessments, acceleration of this progress can be expected.

Chapter 4

Teacher appraisal

Teacher appraisal in the Netherlands is under the responsibility of the competent authority of each school. Central regulations specify that schools should have regular performance interviews with all staff, but employing authorities are free to develop their own frameworks for teacher appraisal. Many school boards delegate the responsibility for human resource management, including teacher appraisal, to the school leaders, and practices vary from school to school. On a system-wide basis, a register system and a peer review project for teachers have been launched by the Education Cooperative, a teacher professional organisation created in 2011. Improving teaching quality has been a policy priority in the Netherlands in recent years, as evidenced by the introduction of teacher competency requirements, the obligation for school boards to monitor teacher competencies and the Inspectorate's increased focus on monitoring teaching quality in schools. Going further, the teaching profession in the Netherlands could benefit from a revised and refined set of teaching standards; strengthened school-based appraisal processes linked to professional learning opportunities, especially for beginning teachers; and an enhanced registration system that could be linked to teacher career development.

This chapter looks at approaches to teacher appraisal within the Dutch evaluation and assessment framework. Teacher appraisal refers to the evaluation of individual teachers to make a judgement about their performance. Teacher appraisal has typically two major purposes. First, it seeks to improve teachers' own practices by identifying strengths and weaknesses for further professional development – the improvement function. Second, it aims to ensure that teachers perform at their best to enhance student learning – the accountability function (Santiago and Benavides, 2009). An overview of the main features of the teaching profession in the Netherlands is provided in Box 4.1.

Context and features

Teacher appraisal procedures

Teacher appraisal in the Netherlands is under the jurisdiction of the competent authority of each school. National regulations specify that schools should have regular performance interviews with all staff, including teachers, at least once every four years in primary education and once every three years in secondary education (OECD, 2013). The collective agreements between the employers' organisations and the teaching unions also specify that regular teacher appraisal should take place. However, while national requirements state that performance conversations must be implemented, there is little guidance provided nationally on how to evaluate the performance of individual teachers.

As the employing authorities for teachers, school boards are free to establish their own frameworks for teacher appraisal. Many school boards delegate the responsibility for personnel matters, including teacher appraisal, to the school leaders. While practice varies across schools, teacher performance reviews typically take the form of an annual or biannual conversation between the school leader and the individual teacher, in which issues related to teachers' responsibilities, working conditions, career and professional development are discussed. More regular formative observation, feedback and coaching for teachers are typically delegated by the school principal to other members of the school leadership team, department heads or team leaders.

In addition, the Education Professions Act (2006) includes a description of expected teacher competencies and requires school boards to establish human resource policies for their schools and maintain competency files for each teacher. These files should describe the teacher's competencies and how these competencies will be maintained. The intention is to ensure that employers have a clear understanding of the competency mix and professional development needs in their schools. The competency files should also help create greater transparency about each teacher's career development and potential, and ensure that all teachers meet minimum competency requirements. When there are indications that a school's quality of educational provision may be at risk, the Inspectorate examines whether the school board has fulfilled this obligation.

Other forms of feedback for teachers

The Education Cooperative, a teacher professional organisation created in 2011, has recently launched a teacher peer review project, which provides a new form of institutionalised feedback for teachers. Based on the idea that teacher peers are best placed to evaluate teaching practice and provide constructive feedback, the peer review project comprises teams of teachers visiting each other's schools and developing tools to observe and evaluate teaching practice. The intention is to use these collegial visitations to observe teaching practice, discuss issues of concern, draw up an observation report and

provide professional feedback for improvement. The project is subsidised by the Ministry of Education, Culture and Science (De Bruin et al, 2013).

Teaching quality is also monitored through the ongoing self-evaluation activities conducted by schools and the regular school supervision carried out by the Inspectorate. While schools are not legally obliged to conduct self-evaluations, they are required to report on the progress of their students and to produce public accountability information on the school's educational results, quality of education, financial situation and professional governance. This information is used by the Inspectorate. The schools visited by the OECD review team reported that they monitored and evaluated the quality of their teaching and learning to draw up annual reports and plan for pedagogical improvements. However, schools are free to choose their own methods for self-evaluation, and little information is available nationally about the approaches they use to evaluate teaching practice as part of their self-evaluations.

The Inspectorate evaluates teaching practice as part of its school supervision and national monitoring tasks. In full quality inspections (Chapter 5), inspectors collect direct evidence of teaching quality on the basis of classroom observations. The Inspectorate's classroom observation framework specifies that inspectors should observe at least four lessons per school. However, these evaluations focus on the school's overall teaching quality and are not intended as an appraisal of individual teachers. They help the Inspectors form a judgement about whether the school leadership team is giving accurate descriptions of the school's quality. In schools where risks for educational quality are identified, inspectors examine the school's human resource policies and verify teachers' qualification levels in relation to detected risks (Chapter 5). Inspectors also observe classroom practice during specific inspection activities conducted for the purpose of system monitoring. Based on aggregated information from these inspection activities, the Inspectorate provides an overview of teacher functioning across the country in its annual *State of Education* reports (Chapter 6).

Competencies for appraisal

The key role in teacher appraisal is exercised by members of the school leadership team. Most school leaders are experienced teachers who apply for school leadership through open competitions. There are few national eligibility requirements for school leadership. The only formal requirements for individuals to apply for the position of principal or deputy principal at a school is that they hold a certificate of good conduct and a higher education degree. If the principal or deputy principal position involves teaching duties, candidates must also meet the relevant competency standards. In secondary education, candidates are required to hold a teaching certificate qualifying them to teach one of the subjects taught at the school. Additional competency requirements may be set by the school board.

While it is not mandatory for school leaders to undertake any particular professional training, a wide range of leadership training offers are available in the Netherlands. Professional training and development for school leadership is offered by a variety of institutions, including Higher Vocational Education (HBO), the trade unions, professional organisations and a range of private training providers. Courses are tailored to different target groups, including principals with different levels of experience, middle managers in secondary education, teachers aspiring to move up to leadership and individuals from outside the education sector interested in a career change. The type and length of training varies between the different offers and may range from two to three years (with an

average of one day a week of study time) to one-day courses on a specific topic (Bal and De Jonge, 2007).

The Dutch School Leaders Association (*Algemene Vereniging Schoolleiders, AVS*), which has existed since the mid-1990s, also plays a role in supporting school leaders' competency development. The AVS acts both as a trade union and a professional organisation. In addition to its collective bargaining, lobbying and legal support tasks, it is increasingly involved in professional aspects, such as, developing guidance materials, connecting school leaders through networks, and disseminating good leadership practice. The professional development of school leaders is further supported by the Dutch School Leaders Academy (*Nederlandse Schoolleiders Academie, NSA*), an independent institute created in 2002, which has set itself the task of promoting the professional quality of school leadership. It accredits and certifies professional development offers, disseminates information related to leadership development, initiates research on effective leadership, and organises conferences and meetings. The NSA and the AVS jointly developed professional standards for school leadership, and the NSA maintains a registration system that allows school leaders who meet the professional standards to register with the Academy. While registration is currently voluntary, it is scheduled to become mandatory in 2015.

As the employers of teachers, school boards also have a formal role in teacher appraisal. School boards may be constituted by various different groups and there is little information nationally regarding the qualifications of school board staff. As described in Chapter 1, school governors may be volunteers or professionals. They may be parents of students in the school, citizens from the local community, members of a religious or life philosophy community, or professionals with specific expertise such as law, finance, human resource management or education. Hence, the competencies of school board members vary considerably across schools. The involvement of school board members with teacher appraisal is typically limited.

Using appraisal results

Teacher appraisal in the Netherlands is primarily used for formative purposes. Performance reviews are expected to feed into professional development for the teacher, ideally in close linkage to the needs of the school. Teacher appraisal may also have summative consequences for teacher career or salary advancement, but this depends on the internal regulations and practices of each school and school board. If an underperforming teacher is identified, it is expected that the school leader finds a solution. School boards can dismiss a teacher on the grounds of underperformance, or they may delegate this responsibility to the school leader. However, this tends to happen only in rare cases. In order to dismiss a teacher, the school leader needs to prove that the concerned teacher underperformed consistently and did not respond to opportunities for support, coaching or professional development offered by the school.

Box 4.1 The teaching profession in the Netherlands: Main features

Employment status

Teachers in public schools have civil servant status while teachers in private schools have salaried employee status. Teachers may be employed on open-ended or fixed-term contracts (for a maximum duration of three years). The conditions of service and legal status of all school personnel are determined at a decentralised level in sectoral collective agreements.

Prerequisites to become a teacher and teacher recruitment

The main requirements to apply for a job as a teacher are to hold a certificate of good conduct and a teaching certificate for the relevant level of education. There is also the possibility for individuals who are not fully qualified as teachers to be appointed on a temporary basis for a maximum of two years after passing an aptitude test. During these two years, these lateral entrants are given training to gain a full teaching qualification. The Education Professions Act (2006) regulates that teachers can only be appointed if they hold a higher education certificate indicating that they meet the competency standards. School boards are responsible for recruiting and dismissing teachers, but they may delegate this task to the school principal through a management contract. Since 1995, all teachers are employed by the school board rather than by a particular school, which means that they can be more easily transferred to another school governed by the same board.

Teacher registration

The Education Cooperative (*Onderwijscoöperatie*), a teacher professional organisation created in 2011, maintains a voluntary registration system for teachers. To be registered, teachers need to meet criteria regarding the amount and content of professional development they have undertaken. The registration process includes the requirement for teachers to complete 160 hours of professional development in four years in order to maintain and renew their registration. The Education Cooperative has set itself the target to ensure that 40% of teachers are registered by 2015. It is intended that the system gradually becomes mandatory. In a recent document, the Ministry of Education, Culture and Science (2013) states that from 2017, all teachers are to be included in the register, which will then have a formal legal status.

Salary and career structure

The Netherlands has a multilevel career structure for teachers, with two levels in primary education and three levels in secondary education. In 2014, there were 15 salary steps in primary education and 12 salary steps in secondary education. Advancement on the salary scale is based on qualifications, experience, performance reviews and responsibility for additional roles and tasks. The government's "functions mix" policy aims at having a balanced mix of teachers at different career levels within each school.

Initial teacher education

Initial teacher education is offered at institutions for higher professional education (HBO) and universities. These institutions are autonomous in determining the teaching and examination regulations for their programmes. Primary school teacher education is part of higher professional education and is provided at both multi-sectoral HBO institutions and colleges specialising in primary teacher education. There are over 30 HBO institutions offering primary school teacher education. Regular primary school teacher education has a study load of 240 ECTS credits or four years of full-time study. Secondary school teacher education is provided at HBO institutions and universities. HBO teacher education institutions cover both subject training and general pedagogy.

Box 4.1 The teaching profession in the Netherlands: Main features (continued)

Two types of qualifications exist for secondary school teachers. Grade two teachers are qualified to teach all years of pre-vocational secondary education (VMBO), but only the first three years of general secondary education (HAVO) and pre-university education (VWO), whereas grade one teachers are qualified to teach all levels of secondary education. University-based teacher education is offered either as a postgraduate course for university graduates with a Master's degree or as a combination of an educational minor at the Bachelor's level (which leads to a grade two qualification) combined with a Master's degree (which leads to the grade one qualification). There are nine universities providing secondary school teacher education. Practical training is a substantial and compulsory part of teacher education both for the primary and secondary level. Details about the period of teaching practice are set out in the teaching and examination regulations of each teacher education institution. There are no national restrictions or quotas regarding the number of places for teacher education.

Professional development

There are no national regulations regarding the amount and content of professional development to be undertaken by teachers. Schools are autonomous and have their own budget to organise continuous professional development for their teachers. Teacher professional development opportunities are offered by a wide range of public and private institutions including HBO institutions, universities with teacher training departments, school advisory services or experts from within or outside the education system. As part of the so-called Integrated Personnel Policy (*Integraal Personeelsbeleid, IPB*), school leaders are expected to align the competencies and the professional development of teachers to the organisational development and the goals of the school as a whole.

Sources: Scheerens, J., et al. (2012), OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.

Eurypedia (2013), European Encyclopedia on National Education Systems: Netherlands, <https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Netherlands:Overview>

Strengths

Definitions of key competencies for teachers exist

The OECD Reviews of Evaluation and Assessment in Education (OECD, 2013) found that teaching standards or competency frameworks are an important element in any teacher appraisal system, as they provide a clear common reference to make judgements about teacher performance. They support the capacity of school leaders, educational authorities and others to effectively review whether teachers have reached a given level of competency. They also offer the potential to frame and align the organisation of key elements of the teaching profession, such as initial education, registration, professional development, career advancement and teacher appraisal (OECD, 2013).

There has been considerable reflection in the Netherlands around what is considered “quality teaching”. As mentioned above, the Education Professions Act includes a description of teacher competencies, which functions as a professional standard for teachers. The competency requirements comprise seven domains (Box 4.2). For each of these seven domains, the competency requirements provide the following elements:

(i) a description of visible aspects of the competency; (ii) proficiency requirements regarding the type of knowledge and skills a teacher must possess in relation to the competency; (iii) indicators and examples of concrete professional actions that may illustrate this particular competency. The requirements do not describe different levels of performance, but define the minimum level of performance that all teachers should achieve in relation to each competency.

Box 4.2 The seven domains of the Dutch teacher competency requirements

1. Interpersonal competencies.
2. Pedagogical competencies.
3. Subject-specific and didactical competencies.
4. Organisational competencies.
5. Competencies to cooperate with colleagues.
6. Competencies to cooperate with the environment.
7. Self-reflective and developmental competencies.

Source: Website of the Education Cooperation (www.onderwijSCOOPERATIE.nl)

The development of these competency requirements began in 2000, when the Association for Professional Quality of Teachers started to engage with teacher groups to articulate what a “good teacher” should know and be able to do. The intention was to develop competency standards of, by and for teachers. The requirements were finalised in 2004 and included in the Education Professions Act in 2006, along with the obligation for the Association for Professional Quality of Teachers (now the Education Cooperative) to review them every six years. In 2012, the Education Cooperative presented a first proposal for revised professional competency standards, suggesting a re-structuring of the competencies along three perspectives: content, pedagogy and didactics. Among the didactical elements, the revised model highlights the importance of continuously observing, evaluating and improving teaching practice.

The minimum competency requirements are mandatory for initial teacher education institutions and appear to influence the design and orientation of their programmes. They are seen as the basic knowledge and skills that all graduates from teacher education should achieve. Hence, the curricula and examinations of initial teacher education institutions are organised around these requirements. The mandatory competencies are typically complemented with additional skills and specialisations defined by each teacher education institution (Eurypedia, 2013). While the competency requirements are used as a common reference for the graduation of teacher students from their initial training, they seem to have less influence on other aspects of teacher policy and practice. There is no obligation for schools to use the competency requirements as a reference for regular teacher appraisal and professional development.

The Inspectorate plays an increasingly important role in defining teaching quality and providing feedback to schools about the strengths and weaknesses of their teacher’s practices. The Inspectorate’s classroom observation framework provides guidance on aspects of good teaching that can be observed as part of a teacher’s practice in the

classroom. The observation framework that is currently used by the Inspectorate comprises fourteen indicators. Inspectors use a classroom observation form on which they indicate whether the teacher's practice in relation to each indicator is sufficient, insufficient or not measurable in the observed lesson. Several of the indicators refer to differentiated instruction and results-based work of teachers, i.e. teachers adapting instruction to different student needs and using adequate instruments to monitor and analyse the progress of their students (Box 4.3). According to interviews with representatives from the Inspectorate, these indicators are in line with the pedagogical and didactical competencies outlined in the teacher competency requirements.

Box 4.3 The Dutch Inspectorate's indicators for lesson observation

1. Teacher makes efficient use of teaching time.
2. Teacher ensures that pupils interact with each other in a respectful way.
3. Teacher explains things clearly.
4. Teacher explains clearly according to didactical principles (e.g. didactics concerning subject matter).
5. Teacher produces a task-related working atmosphere.
6. Pupils are involved in education activities.
7. Teacher checks if pupils understand explanation and/or exercises.
8. Teacher gives pupils feedback on learning and development process.
9. Teacher adapts instruction to differences in development between pupils.
10. Teacher adapts exercises to differences in development between pupils.
11. Teacher adapts teaching time to differences in development between pupils.
12. Teacher uses a coherent system of standardised instruments to monitor progress and development of pupils.
13. Teacher monitors and analyses progress in the development of pupils systematically.
14. Teacher executes care (for children with special needs) according to the plan (of action).

Source: Dutch Inspectorate of Education (2012), 5. Lesobservatieformulier: toezicht leraarschap P.O, Dutch Inspectorate of Education, Utrecht.

The Education Council has also emphasised the importance of defining teacher professionalism in order to help teachers deal with the complexity and dynamics of teaching practice. In a 2013 publication entitled *Being a Teacher*, the Council explores the concept of 'personal professionalism', i.e. the attitudes, knowledge and practice of effective teachers as shaped by their daily teaching experiences (Dutch Education Council, 2013). According to the Council, 'personal professionalism' is influenced by collective frameworks, such as the competency requirements, but it is also shaped by teachers' personal values and their conception of the teacher they want to be. The Council's advice provides four considerations for developing teachers' personal professionalism. These relate to: (i) being aware of one's own professional values and goals; (ii) developing 'practical wisdom' and the capacity to make quick judgements in complex situations; (iii) using and creating 'professional space' by understanding social

processes and ways to influence them; (iv) maintaining an inquisitive attitude and continually developing one's own capacities.

There is a clear agenda for improving teaching quality

Improving teaching quality has been a clear policy priority in the Netherlands in recent years. The main purpose of the 2006 Education Professions Act is to enhance teacher professionalism and teaching quality. As described above, the Act sets minimum standards of competency for teachers and other educational staff, and introduces an obligation for school boards to ensure that their staff possess the required competencies and are able to maintain them. The development of each school's competency mix should be monitored through the use of teacher competency files.

Another cornerstone of the teacher professionalism agenda is the government's Action Plan *Teaching 2020: A Strong Profession!*, published in May 2011. The Action Plan highlights the importance of increasing teacher professionalism, in particular with regard to results-based work and differentiated instruction (see Chapter 3). Teacher appraisal forms a key component of this agenda. Regular appraisal interviews and the use of competency files at the school level are emphasised as important strategies to provide feedback to teachers and plan for school-wide professional development. The Action Plan also points to the positive results of peer review projects, and suggests further development of such approaches. In addition, it proposes a range of complementary measures, including teacher registration, enhanced personnel policies, opportunities for career development and an extension of the Inspectorate's remit to monitor teaching quality in schools (Box 4.4).

In October 2013, the Ministry of Education, Culture and Science launched a more detailed "Teacher Agenda" (*Lerarenagenda*), outlining key priorities for the period from 2013 to 2020. These include: (i) improving initial teacher education; (ii) offering adequate professional development opportunities; (iii) providing attractive and flexible learning pathways; (iv) ensuring a good start for beginning teachers; (v) developing schools as learning organisations; (vi) helping all teachers maintain and update their competencies; (vii) sustaining a strong professional organisation that represents teachers (Dutch Ministry of Education, Culture and Science, 2013).

Box 4.4 Key elements of the Dutch Action Plan *Teaching 2020: A Strong Profession*

Teacher registration

The Action Plan foresees that within the next few years, all teachers should undergo a professional registration process to ensure that they maintain and develop their competencies. To be able to register, teachers need to complete a defined number of accredited professional development activities. Currently, the expectation is that teachers complete 160 hours of professional development within 4 years in order to maintain their registration status. The register has been launched as a nationwide system in February 2012 by the Education Cooperative (Box 4.1 above). It builds on earlier work by organisations of subject matter teachers and the Association for Professional Quality of Teachers.

Box 4.4 Key elements of the Dutch Action Plan *Teaching 2020: A Strong Profession*
(continued)

Enhanced personnel policies and teacher career development

The Action Plan brings together a number of suggestions to incentivise excellence in individual teacher and team performance. It sets the target of completing the roll-out of the “functions mix” policy by 2020. This policy enables promotion based on differences in teacher competencies and performance. It is expected to support teacher career development and increase the number of teachers in the higher salary scales. The collective labour agreements provide schools with considerable freedom to use their budgets in implementing the functions mix, e.g. to recruit additional teachers or place teachers with specific expertise on a higher salary scale. The government has established performance agreements with the education sector organisations in relation to the performance indicators of the Inspectorate. In a context where school boards are largely autonomous in managing school personnel, this policy aims to ensure that all school boards develop adequate human resource policies.

Extension of the Inspectorate’s remit

Over the past few years, the Inspectorate has paid increasing attention to teacher appraisal and the improvement of teaching quality. The Inspectorate’s responsibilities were extended in 2012 to intensify its focus on teaching quality and teacher professionalism. In line with these extended responsibilities, inspectors are increasingly focussing on how schools safeguard individual teaching skills and enhance teacher professionalism within the school. This includes evaluating the school leader’s quality policy and human resource policy. The Inspectorate is expected to take action if shortcomings in this area are observed, for instance if schools fail to conduct regular appraisal interviews. Further revisions made to the Supervision Framework in 2013 added two additional quality aspects to be evaluated by the Inspectorate: (i) school leaders focussing on teacher development in line with the school’s vision and (ii) teachers using their professional space to deliver good education. At the time of the OECD review visit, the Inspectorate was preparing a thematic review on schools’ human resource and professional development policies.

Source: Dutch Government (2011), *Teaching 2020, A Strong Profession!*, Dutch Government, www.ecbo.nl/ECBO/ReferNet/docs/11-0315_Teacher_2020.pdf.

The teaching profession is taking responsibility for moving the agenda forward

The involvement of teachers and their representative bodies in designing teacher appraisal approaches and wider teacher policy is essential for ensuring that such policies are effective and make sense for the teaching profession. Such participation recognises teachers’ professionalism, the importance of their skills and experience, and the extent of their responsibilities (Hess and West, 2006). If teacher appraisal, and teacher policy more widely, is developed in close cooperation with teachers and their professional organisations, teachers are more likely to feel ownership of the appraisal cycle and be open to receiving feedback and being evaluated.

In the Netherlands, there are a range of teacher representative organisations which have been increasingly involved in shaping teacher policy and the teacher professionalisation agenda. In interviews with the OECD review team, representatives from the main teacher unions reported that the unions have been moving away from an

exclusive focus on labour-related issues towards a stronger content orientation, including a growing emphasis on the professional aspects of teaching. Teachers also shape the educational agenda and professional dialogue through the Education Foundation (*Stichting van het Onderwijs*) which was established jointly by the teachers' unions and employers' organisations in March 2010. The Foundation brings together several organisations from all sectors and levels of education and takes responsibility for the organisation of an annual or semi-annual strategic dialogue between the representatives of the education sector and the government. The foundation provides a platform for structured dialogue with the Ministry of Education regarding broad sectoral issues.

The creation of the Education Cooperative (*Onderwijscoöperatie*) in 2011 epitomises this trend of teacher unions taking increasing responsibility for professional matters. The main focus of the Education Cooperative is on quality and professionalism in education. It has quickly become an important player in the Dutch educational landscape and was given official responsibility to review teacher competency requirements and develop the teacher registration system. In addition, the Cooperative has launched a range of projects contributing to the professional learning of teachers, including the teacher peer review project (see above), a web-based teacher TV with audio-visual materials of teacher practices, an incentive programme to support teacher initiative, and teacher-of-the-year elections intended to boost the image of the teaching profession.

There are formal and informal channels for regular school-based teacher appraisal

Even though there is little national guidance regarding teacher appraisal processes in the Netherlands, there are a range of formal and informal channels through which the majority of teachers receive appraisal and feedback. The mandatory performance interviews provide a structure for teachers to receive occasional feedback from their school leaders. While these conversations are not as yet conducted systematically for all teachers (a challenge that will be addressed below), a 2010 survey among education professionals found that 73% of teachers had participated in performance interviews with their school leader. Key themes addressed in such interviews included observation visits with other teachers, different approaches to keep competencies up to date, participation in coaching, and career and salary development (Bokdam et al., 2011). In the schools visited by the OECD review team, principals typically conducted an annual or biannual performance review with each teacher, and delegated more regular formative observation, feedback and coaching to middle managers, department heads or team leaders.

The organisation of teaching staff in teams and departments provides further opportunities for exchange and peer learning among teachers within a school. The schools visited by the OECD review team all had systems giving team leaders or department heads responsibility for conducting regular classroom observations and performance conversations with teachers, using school-based criteria. The target frequency for such classroom visitations varied from several times a year to once every two years, even though team leaders reported that they often lacked the time to conduct observations systematically for all the teachers in their team or department. These informal observations and feedback sessions were typically intended for formative purposes. Some schools used professional development plans or portfolios to monitor teacher development and several schools also had internal coaching systems, where an experienced teacher or mentor would be available to work with teachers facing difficulties.

In the interviews with the OECD review team, several stakeholders indicated that informal collaboration and feedback within schools was becoming increasingly common in the Netherlands. While Dutch education is characterised by a strong tradition of teacher autonomy, according to the Education Council, there has been a trend in the last few years of teachers collaborating and working in teams more, which provides opportunities for peer review and feedback. Many of the practitioners interviewed by the review team saw student feedback as important information for their self-appraisal and improvement of their practice. The national student organisation (*Landelijk Aktie Komitee Scholieren, LAKS*) indicated that it was common practice in many schools for teachers to request formative feedback from their students through questionnaires.

Challenges

Lack of clarity around teaching standards

While the competency requirements are a mandatory element of initial teacher education, the OECD review team formed the impression that these standards were not systematically carried forward into the practice, appraisal and professional development planning of teachers in schools. It is not mandatory for schools to use the competency requirements in their appraisal processes or planning of continuing professional development, and school-level records of appraisal interviews typically do not provide evidence on whether teachers are meeting the requirements. This risks weakening the alignment between initial teacher education, registration, teacher appraisal, professional development and career development that the common reference standards seek to achieve.

The Education Council have criticised the competency requirements for being too vague, abstract and unspecific to be used for effective teacher appraisal. Several commercial groups have developed simplified web-based instruments to help school leaders assess teacher performance in relation to the competency requirements. But a range of stakeholders interviewed by the OECD review team voiced criticism regarding the use of these instruments by school leaders. According to the Education Cooperative, in the absence of appraisal capacity at the school level, these instruments may lead to appraisal processes based on box-ticking rather than attentive observation and constructive dialogue on teaching practices.

Most reviewers (team leaders, department heads, principals) involved in teacher appraisal have not received any specific training to appraise teachers in relation to the competency requirements, and the requirements provide only limited guidance for appraisal processes. Hence, the point of reference for teacher appraisal tends to be the reviewers' own teaching experience rather than a deep understanding of the level of performance that can be achieved by the most effective teachers in relation to the dimensions set out in the competency requirements. With the competency requirements focusing on minimum standards, there has been less attention nationally to discuss the characteristics of excellent teaching.

The co-existence of several different sets of teaching standards (the competency requirements, the Inspectorate's classroom observation framework, the Education Council's definition of 'personal professionalism'), risks creating confusion and sending conflicting messages about what teachers are expected to know and be able to do. The Inspectorate does not use the competency requirements as a reference for its evaluation of teaching quality in a given school as they are considered too broad to be monitored

through an inspection visit. Within schools, there was often a lack of clarity about what standards are used in teacher appraisal and performance reviews. Some schools have developed their own appraisal criteria, sometimes drawing from the competency requirements, the Inspectorate's observation framework or a mix of both.

The lack of a common framework of references for evaluating teaching quality in the Netherlands is likely to weaken the capacity of schools to appraise teachers effectively. While some schools have developed their own references based on local practice, for teacher appraisal to be effective across the system it would be important that all reviewers have a shared understanding of high quality teaching.

Not all teachers are receiving regular appraisal and feedback

There is an expectation in the Netherlands that all teachers go through processes of regular performance appraisal. However, while most teachers seem to benefit from regular appraisal conversations, there are concerns that not all teachers have opportunities to receive appropriate professional feedback and have their competencies recognised. Inevitably, since school governing boards have flexibility in the design of performance appraisal systems, there is potential for wide variation in the extent and quality of teacher appraisal. According to the Dutch Government (2011), the frequency of appraisal conversations in the education sector is lower than in other public sectors in the Netherlands.

The existing teacher appraisal practices are often the initiative of individual schools (in some cases in the context of requirements established by the school governing board) and largely depend on the leadership style of the principal and the evaluation culture of the school. As mentioned earlier, in some schools teachers receive extensive feedback and support from their immediate supervisors and school leaders, but there is no mechanism in the Netherlands to ensure minimum standards for teacher appraisal in schools and there are no guarantees that every teacher receives proper professional feedback. The Dutch Inspectorate of Education (2013a) found that weaker teachers were often insufficiently aware of their own strengths and weaknesses, possibly due to a lack of professional feedback from colleagues and supervisors. This also means that in those schools where teacher appraisal processes are weak, it might be difficult to identify and address underperformance.

There is also evidence that the use of competency files to monitor teacher competencies has not been widely adopted. Bokdam et al. (2011) found that in 2010 only about 20-30% of teachers were familiar with the new competency regulations and 25% of teachers reported having a competency file. In primary schools, about two-thirds of principals were aware of the competency demands. Many of the stakeholders interviewed by the OECD review team saw the competency files as a mere bureaucratic requirement with little impact on actual practice in schools. A 2010 evaluation of the Education Professions Act confirmed that the use of competency files had not generally become a part of schools' personnel policies or evaluation approaches (Dutch Inspectorate of Education, 2010a). In an exploratory study on the quality of school governance, the Dutch Inspectorate of Education (2013b) found that the mechanisms put in place by school boards to monitor school leader and teacher competencies were often insufficient.

The extent and quality of guidance for beginning teachers varies across schools

While frequent observation, evaluation and feedback can help improve the practice of all teachers, it is particularly important for beginning teachers who have limited

experience in the classroom. In the Netherlands, there is currently no mandatory induction period for new teachers. Schools are expected to organise their own procedures for induction, mentoring and coaching of new teachers, but there is no guarantee that such structures exist in all schools. According to the Dutch Government (2011), 10% of teachers leave the profession after their first year of teaching. This indicates that teachers are facing difficulties in the transition from initial teacher education to actual classroom teaching.

While it is likely that a combination of factors influence beginning teachers' decision to continue in their jobs after the first year, research points to the importance of providing a well-supported working environment, including frequent feedback and mentoring for beginning teachers (OECD, 2010; Jensen et al., 2012). According to the Dutch Education Career Monitor, only 42% of new teachers reported that they were satisfied with the level of guidance they received in schools (Dutch Ministry of Education, Culture and Science, 2009). The Dutch Inspectorate of Education (2010b) evaluated training partnerships in the Netherlands and concluded that a system of regular lesson observation and feedback could help improve the guidance provided to beginning teachers.

Concerns about responsibilities and competencies for teacher appraisal

In the complex governance framework of the Dutch education system, it is important to clearly identify who is responsible for teacher appraisal and whether those in charge possess the required competencies. In the Netherlands, the national authorities monitor teaching quality at the system level but they are not considered responsible for quality at the level of individual professionals. Hence, while the government can provide guidance and reference documents regarding excellence in teaching and effective appraisal practice, it does not have the remit to introduce a centralised teacher appraisal system or to develop a national teacher registration or licencing system. This highlights the important role of teacher professional organisations to take initiatives in this field.

School boards, as the employers of teachers, clearly have a role to play in teacher appraisal, but representatives of the Primary and Secondary Education Councils indicated that many school boards are not taking responsibility for the appraisal for teachers. There are large variations across school boards in terms of the background and competencies of school governors. While school boards should provide feedback to the school leader and ensure that school leaders have functioning personnel and appraisal policies in place, board members often do not have a background in education and may not have the capacity to conduct quality control in a systematic manner. Where this is the case, school governors may lack legitimacy as evaluators in the eyes of teachers and/or they may be reluctant to get involved with the school's approaches to teacher appraisal and feedback.

As mentioned above, most school boards delegate responsibility for the appraisal of teachers to the school leaders. However, given the importance of teacher autonomy in the Netherlands, principals do not typically take a strong role with regard to directly influencing the day-to-day professional practice in their colleagues' classrooms and teachers are generally left to their own devices unless any problems arise. Research by the Dutch Inspectorate of Education (2014) found that school leaders use information on teachers primarily for staff decisions, but less so for planning improvements to the quality of teaching and learning.

As there is no mandatory school leadership training for principals, the preparation and competencies of principals vary across schools. Even though examples of school leaders exemplifying strong pedagogical leadership and human resource management certainly

exist, there is a challenge for the Dutch system in building up the role and capacity of their full cohort of school leaders. Representatives from the Education Cooperation voiced concerns that many school leaders lacked the expertise and tools to identify causes of underperformance and to develop potential remedies, or actions for dismissal where necessary. A recent evaluation of the Education Professions Act confirmed the need for greater clarity regarding the competency requirements for principals and for more consistent professional development for all school leaders in this area (Dutch Government, 2011).

Professional development is not sufficiently informed by teacher appraisal results

For teacher appraisal to have an impact on learning outcomes in the school, it needs to be closely connected to professional development and school development. Without such a link to professional development opportunities the impact of teacher appraisal on teacher performance will be relatively limited (Goe et al., 2012). As a result, the appraisal process may not be taken seriously or encounter mistrust or apathy by the teachers being appraised (Danielson, 2001; Milanowski and Kimball, 2003; Margo et al., 2008). Ideally, teacher appraisal should allow teachers to receive tailored feedback, and such feedback should be followed with opportunities for continuous learning in identified areas through professional development, mentoring and other means (Hill and Herlihy, 2011).

The importance of professional development is clearly recognised in Netherlands and the introduction of the registration system further emphasises the expectation that all teachers engage in ongoing professional learning. Informal mentoring arrangements within schools also appeared to be common practice. However, it was the impression of the OECD review team that teachers' choice of formal professional development was only rarely linked to a thorough evaluation of their strengths and areas for development. The Dutch Inspectorate of Education (2013a) found that schools vary widely regarding the support they provide teachers to facilitate their professional development. The Inspectorate also indicates that training is often too discretionary and lacking in focus on the actual teaching and learning process.

There is scope to better link teacher professional development to school development and improvement. In the interview with the OECD review team, the Education Council voiced concerns about the limited focus on teachers' broader role in school development. The relatively weak linkage between teacher appraisal, teacher professional development and school development is likely related to the fact that individual teacher appraisal is currently not considered a key element of the Dutch evaluation and assessment framework (Chapter 2).

Weak links between teacher appraisal and career development

Providing attractive career pathways for teachers is a challenge in teacher policy around the world. Findings from the OECD's Teaching and Learning International Survey (TALIS) show that in most countries, the link between teacher appraisal and career advancement remains weak. Across TALIS countries, only 16.2% of teachers indicated that the appraisal and/or feedback they received led to a moderate or large change in the likelihood of their career advancement. Only 26.7% reported that it led to changes in work responsibilities that made their job more attractive (OECD, 2009).

In the Netherlands, there does not seem to be a formalised career path for teachers. The competency requirements do not specify skills and competencies required at different

stages of the career in association with roles and responsibilities of teachers in schools. There are some opportunities for teachers to take on more responsibility, including roles such as team leader or head of department. Teachers taking on such responsibilities are typically promoted to ‘senior teacher’ positions, which are connected to a higher salary scale. However, there are few such posts and teachers interviewed by the OECD review team reported frustrations about many teachers waiting for a vacancy to be able to apply for senior teacher positions.

The award of senior teacher positions is not typically linked to an appraisal of teachers’ performance in relation to the competency requirements. School leaders may promote teachers to such positions based on criteria defined at the school level. Criteria typically include length of service, formal qualifications and professional development undertaken, coupled with engagement in improvement activities, positive reviews from peers, students and parents and/or the exercise of leadership responsibilities. School leaders may also use the senior positions to attract and retain teachers in subject areas where the school is facing shortages. It was the impression of the review team that decisions about promotions were disconnected from regular teacher appraisal processes. The teacher performance conversations and reviews were widely seen a routine cycle to validate satisfactory performance of teachers, rather than as a motivating and rewarding system.

At the national level, there is no clearly designed career structure beyond the career step of senior teacher and there are few opportunities for promotion and greater recognition. The organisational structure in schools is typically flat with few promoted posts and few explicit means of giving teachers significant whole-school lead responsibilities. As a result, two major functions of teacher appraisal processes are undermined: *(i)* granting effective teachers opportunities to diversify their careers in response to the roles and tasks performed in schools; *(ii)* providing a means to formally reward teachers for the gained competencies and skills to take on higher responsibilities. This is likely to undermine the potentially powerful links between teacher appraisal, professional development and career development.

The absence of career development opportunities for teachers may be one of the factors contributing to the challenges in attracting young people to the teaching profession in the Netherlands. The Dutch education system is facing difficulties in recruiting and retaining sufficient numbers of effective teachers. It is faced with high drop-out rates in initial teacher training as well as high attrition rates among teachers within the first five years on the job. While the decision to complete initial teacher education and stay on the job is influenced by many factors, well-defined career development opportunities could contribute making the profession more attractive.

In some countries, teacher career advancement is linked to teacher registration and registration renewal. The registration system in the Netherlands is still in the early stages of development and its role in the teacher career has not yet been clearly defined. Currently, registration and registration renewal do not grant teachers access to a higher career step. Teacher registration is contingent on the requirement for teachers to provide proof of professional development undertaken through in-service training courses, but it does not include an appraisal of teachers’ actual practice. The further development of the teacher register may provide valuable opportunities to further develop the teacher career in the Netherlands.

Policy recommendations

This section presents a set of policy options that aim to draw on current strengths in teacher appraisal policies and to address identified challenges:

- Review and refine teaching standards.
- Strengthen school-based appraisal for professional development.
- Focus in particular on beginning teachers.
- Further develop the teacher career structure.
- Build a more elaborate registration system linked to career development.

Review and refine teaching standards

A framework of teaching standards is an important reference point for teacher appraisal. While competency requirements for teachers exist in the Netherlands and are widely used in initial teacher education, their use for regular appraisal and professional development in schools appears limited. To ensure coherence between initial teacher education, registration, appraisal and professional development, it is essential to promote the wider use of the competency requirements as a working document in schools underlying all of these processes.

To this end, further revisions to the competency requirements appear necessary. The current co-existence of several types of references for the evaluation of teaching, in particular the competency requirements and the Inspectorate's classroom observation framework, call for their consolidation into a single set of standards to develop a clear shared understanding of what counts as accomplished teaching in the Netherlands. The consolidated standards should be based on the latest research on teacher effectiveness and give due importance to the links with the student learning objectives that schools are aiming to achieve. This could be done, for example, by including a focus on results-oriented work in the competency requirements, thereby strengthening the coherence of different elements of the Dutch evaluation and assessment framework. It would be important to keep a focus on improving student learning objectives for all students, particularly for groups where there is evidence of underperformance.

The consolidated standards should also build on the practice-based expertise. To this end, it would be helpful to conduct a thematic review on the use of teacher appraisal standards and criteria by schools. Such a review would help to understand how the competency requirements are currently viewed and used, what are seen as most powerful and productive elements, what issues it raises for effective teacher appraisal, what additional checklists and criteria schools have developed themselves, and how the competency requirements might be further developed. The Inspectorate appears well placed to collect such information from schools. It would then be the role of the Education Cooperative to use the results of the review to inform further revisions to the competency requirements in close collaboration with stakeholders in schools.

Another important adjustment could be to develop clearer descriptions of competency requirements for different roles and career steps of teachers, with appraisal criteria specific to distinct career levels. Such a revision of the competency requirements would help recognise the variety of responsibilities in today's schools and the expertise developed while on the job. Defining different competency levels within the requirements

would also help to guide teachers' improvement of skills and competencies, and steer their aspirations to new responsibilities. The description of competencies should be complemented by criteria and illustrations of effective practice, to help make the standards operational for regular use in school-based teacher appraisal. An important contribution in this area is Danielson's Framework for Teaching (1996, 2007), which provides "a road map to guide novice teachers through their initial classroom experiences, a structure to help experienced professionals more effective, and a means to focus improvement efforts".

Strengthen school-based appraisal for teacher professional development

As described above, school-based formative teacher appraisal takes place in many schools across the Netherlands, typically with senior teachers, team leaders or department heads conducting classroom observations, and principals holding performance conversations with their teachers. However, further steps are necessary to ensure that all teachers across the country benefit from meaningful appraisal and feedback, pursue relevant professional development, and are able to implement improvements in the classroom. To make developmental appraisal processes more effective and consistent across the country, the OECD review team recommends that it should be: (i) school-based but underpinned by common reference standards; (ii) firmly rooted in classroom practice; (iii) carried out by qualified internal evaluators; (iv) externally validated by school governing boards and the Inspectorate.

Teacher appraisal for improvement purposes is likely to benefit from a non-threatening evaluation context, simple evaluation instruments and close linkages to school self-evaluation. There are many advantages to having colleagues of the teacher as the evaluators, given their familiarity with the context in which teachers work, their awareness of the school needs, and their ability to provide quick and informed feedback to the teacher. Therefore, developmental appraisal should remain an internal process carried out by line managers, senior peers and the school leader, with a focus on teachers' practices in the classroom. While the process for developmental appraisal should remain school-based, it could be enhanced by a clearer link to common reference standards of "good teaching" (the revised competency requirements, see above). This would allow all school leaders to develop a shared understanding of expected teaching standards and of the level of performance that can be achieved by the most effective teachers.

The competency requirements should also inform the offer of professional development available to teachers. This could be achieved by the Ministry of Education, Culture and Science and/or the Education Cooperative reviewing professional development offers, and, with the competency framework in mind, providing guidance for schools on relevant training offers. For an example from Memphis, Tennessee in the United States, see Box 4.5.

Box 4.5 Memphis, Tennessee, United States: Linking teacher appraisal to professional development

The city of Memphis, Tennessee in the United States has developed a system that explicitly links professional learning to teacher appraisal. In Memphis City Schools, appraisal is based on teaching standards, and professional development is linked to teachers' competence on the standards. Thus, a teacher who has poor performance on a specific indicator on a teaching standard can find professional growth opportunities related to that indicator. Memphis City Schools publishes a professional development guide each year that lists the professional growth offerings by standard and indicator. In addition, most of the professional development courses are taught by Memphis City School teachers, ensuring that the course offerings will be relevant to the contexts in which these teachers work.

Source: Memphis City Schools, www.mcsk12.net.

Effective teacher appraisal should give teachers a choice from a wide range of possible professional learning activities that meet their individual needs in relation to the priorities of the school's overall development plan. In Korea, for example, results of the teacher peer review processes not only feed into teachers' individual professional development plans, but are also used to inform a synthetic report on professional development for the whole school bringing together the results of all appraised teachers (without identifying individual teachers) (Kim et al., 2010).

School-based developmental teacher appraisal can be low-key and low-cost and include a mix of methods appropriate to the school. Some of the elements should be individual goal-setting linked to school goals, self-appraisal, peer appraisal, classroom observation, and structured conversations with the school leader and peers (Santiago and Benavides, 2009). Among these approaches, classroom observation is likely to be the most relevant source of information about professional performance, as most key aspects of teaching are displayed when teachers interact with their students in the classroom. Other proxies of teaching quality, such as lesson plans, are also important pieces of information, but they do not hold the same central position as the observation of classroom teaching. Research suggests that high-quality observations are related to increases in student learning outcome, even though this relationship is highly dependent on having excellent instruments and well-trained observers (Kane and Staiger, 2012; Kane et al., 2010; Milanowski, 2004). Hence, for classroom observations to be useful for professional improvement, each school must have the internal capacity to conduct these accurately.

As described above, effective teacher appraisal depends to a large extent on school leadership capacities within the school. While there are a wide range of offers for Dutch school leaders to develop their capacities, there is no mandatory formal training for school leadership and not all principals have participated in professional learning on pedagogical leadership, human resource management and teacher appraisal. If school leaders are to drive up the quality of outcomes for learners they need to develop the skills, competence and authority to influence practice in this way, and this needs to happen consistently across the system. The following elements could be part of a comprehensive strategy to build capacity in this area: *(i)* considering the implementation of a mandatory training for school leadership (for an example from Norway, see Box 4.5); *(ii)* disseminating resources and training for the direct evaluation of pedagogical practice

including the observation of classroom practice and providing effective feedback for improvement; (iii) ensuring that school leaders themselves receive adequate appraisal and feedback and building the capacities of employers to undertake effective performance review of school leaders; (iv) allowing greater access for school leaders to participate in external reviews and development work with other schools in their areas or elsewhere, for example through school leaders' participation in inspection visits (for an example from Northern Ireland, see Box 4.6)

Box 4.6 Building school leader capacity for teacher appraisal and evaluation in Norway and Northern Ireland, United Kingdom

In **Norway**, where there is little tradition for regular classroom observation by principals, a new national education programme for principals was introduced in 2009. The education programme was initially targeted at newly employed principals who had been in the position for less than two years. It was then extended for more long-standing principals who had not received such an education. The overall aim of this initiative is to better equip principals for their role as leaders, and in particular for taking a stronger role in guiding the teaching and learning processes at school. It is expected that as principals are become better prepared for pedagogical leadership, they will also become more confident in appraising and providing feedback to their teaching staff. It is hoped that this will help increase the acceptance among teachers of school leaders observing classrooms and appraising teaching performance. Among the skills and attitudes principals should be able to master in this area, there are several aspects that relate to appraising and guiding teachers' practices: (1) Setting goals for teaching work; (2) Setting standards for quality in working processes and being able to enforce these; (3) Following up on and giving feedback to individual co-workers; (4) Creating pride, aspirations and a desire to achieve results in teachers; (5) Guiding and giving feedback to teachers; (6) Challenging teachers and setting definite demands on quality.

In **Northern Ireland** (United Kingdom), the Education and Training Inspectorate (ETI) recruits "associate assessors" from among senior staff in schools (e.g. school principals, deputy principals or senior teachers) to participate the external evaluation of individual schools. ETI recruits associate assessors via public advertisement and an interview process. Selected individuals join a pool of associate assessors and can be invited to join an external school evaluation team on an individual school inspection. Normally an individual will not be involved in more than two external school evaluations each year. Associate assessors receive training from the ETI and are introduced to the procedures and performance indicators used in external school evaluation. This strategy has two objectives: first, it is hoped that the experience of involvement in assessing quality in another educational establishment will help to develop the individual's capacity to monitor, evaluate and improve the provision in his/her own school; second, the presence in the team of someone coming directly from the school context adds a dimension which can help to develop the ETI's awareness of the current perspective of schools.

Sources: Nusche, D., L. Earl, W. Maxwell, C. Shewbridge (2011), *OECD Reviews of Evaluation and Assessment in Education: Norway 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264117006->

Department of Education, Northern Ireland (2013), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Northern Ireland*, www.oecd.org/edu/evaluationpolicy.

At the same time, it is also important to recognise that given their wide range of other budgetary, administrative and human resource management tasks, it is challenging for school leaders to make time for the thorough appraisal of each teacher in the school. Therefore, capacity for teacher appraisal and evaluation should be developed not only

among school principals and deputies, but also among other members of school leadership and senior teachers who undertake specific appraisal and evaluation functions in the school. This means that the provision of training opportunities regarding appraisal and evaluation should be scaled up for a wider group of school staff including middle leaders. Such training should include teachers as well, since it is critical for them to understand how their performance will be appraised.

Finally, to ensure that school-based developmental appraisal is systematic and coherent across Dutch schools, it would be important that an external body provides a validation of school level processes for teacher appraisal, holding the school leader accountable as necessary. The school governing boards should ensure that schools develop appraisal processes and could encourage schools to document their practices as part of school self-evaluation.

In addition, the Inspectorate of Education, in its evaluation of the quality of teaching in schools, should also include a review of each school's teacher appraisal processes, holding both the school leader and the governing board accountable. This would help ensure that minimum standards for development teacher appraisal are met and that every teacher receives regular professional feedback, without imposing one particular model of appraisal.

Focus in particular on beginning teachers

The Dutch education system could benefit from the introduction of more systematic induction and feedback systems for new teachers. Research from different countries points to the importance of ensuring that beginning teachers receive adequate guidance (OECD, 2010; Jensen et al., 2012). At this early stage of teachers' career, it is particularly important to ensure teachers can work in a well-supported environment and receive frequent feedback and mentoring. Most high-performing education systems require their beginning teachers to undertake a mandatory period of probation or induction, during which they receive regular support and can confirm their competence to move on to the next stage of the teaching career (OECD, 2010). Box 4.7 provides an example from Northern Ireland in the United Kingdom.

Research indicates that beginning teachers benefit from systematic mentoring programmes as long as mentors are carefully selected, well prepared for their tasks and given adequate time to carry out their mentoring role (Hobson et al., 2009; OECD, 2010; Santiago et al., 2013). However, it is important to note that among TALIS countries there is no quantitatively important relationship between the existence of a formal induction/mentoring process and the frequency of appraisal and teachers in their first two years at school (OECD, 2009). Hence, if the purpose of induction periods is to strengthen observation and feedback mechanisms for beginning teachers, it is important to make such elements an explicit and expected part of the programme.

Box 4.7 Support for beginning teachers in Northern Ireland, United Kingdom

In Northern Ireland, a “career entry profile” is established for each beginning teacher upon completion of initial teacher education. This profile outlines the teacher’s strengths and areas for further development in relation to the Northern Ireland competence model. When taking on a first teaching position, there is a formal one-year induction period to help teachers address the personal and professional needs and objectives identified in their career entry profile. The induction period involves a programme of both centre-based and school-based professional support. The Board of Governors, upon recommendation of the school principal, approves the teacher’s completion of the induction period and the teacher professional organisation (GTCNI) holds a record of completion of induction.

As part of the induction process, teachers then prepare a personal action plan, which forms the basis for a two-year period of Early Professional Development (EPD). This phase involves within-school support by a “teacher tutor” and by the regionally-based Curriculum Advisory and Support Services (CASS). It is aimed at helping beginner teachers further develop and consolidate their competencies. When the beginning teacher and teacher-tutor agree that all the criteria for EPD have been met, they will seek confirmation by the school principal. The Board of Governors approves the completion of EPD, based on the recommendation of the principal and a final reflection document produced by the teacher concerned.

The early teacher education and development phases are further strengthened through the *Teacher Education Partnership Handbook*, which provides guidance to all those involved in the process, including student teachers, beginning teachers, teacher tutors, Education and Library Boards and higher education institutions.

The availability of teacher tutors in each school is an important element in facilitating the transition of teachers from initial education into full-time teaching at a school. Teacher tutors are responsible for placement and care of student teachers in a school. They are typically senior teachers who can draw on their own experience to support beginning teachers through their first years of teaching. The tutors are expected to hold regular meetings with beginning teachers, draw up action plans, assist in lesson planning, observe classroom practice, review progress and provide general support to help the beginning teacher reflect upon his or her practice and improve classroom teaching. Tutors can play a key role in helping beginning teachers understand existing standards, self-appraise their practice and use feedback from others to review and improve their practice.

Source: Shewbridge, C., et al. (2014), (2014), *OECD Reviews of Evaluation and Assessment in Education: Northern Ireland, United Kingdom*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264207707-en>.

Further develop the teacher career structure

There is room to further develop the teacher career structure in the Netherlands in order to recognise and reward teaching excellence and allow teachers to diversify their careers. Schools and teachers are likely to benefit from a more elaborate career structure for teachers, which could comprise a number of key stages. Access to each of the key stages could be associated with a more formal appraisal process, which could potentially be organised through the teacher registration system (more on this below). An important policy objective should be to match the career structure for teachers with the different types and levels of expertise described in the revised teacher competence standards. This would strengthen the incentive for teachers to improve their competences and reinforce

the matching between teachers' competences and the roles that need to be performed in schools to improve student learning.

In the Netherlands, besides the career step of senior teacher, the only possibility for promotion is for teachers to move up to a principal position. But, besides the fact that only few such positions are available, a promotion of outstanding teachers to school leader positions may not respond well to the needs of the teaching profession, for two main reasons. First, a good teacher is not necessarily a good manager or leader and the skills required for teaching a classroom and managing a school are not the same. Second, this practice may have adverse effects on teaching quality within a school because, paradoxically, the best teachers are rewarded by being removed from classroom teaching.

To resolve this dilemma, some education systems have attempted to build career options for excellent teachers who wish to remain in the classroom (Box 4.8 provides examples from Singapore and Australia). When designing a career structure for teachers, education authorities should make sure that career pathways are varied with some teachers moving into leadership roles while others remain predominantly teaching in the classroom.

Box 4.8 Teacher career structures in Australia and Singapore

Australia: Advanced Skills Teaching positions

Teachers in Australia undergo appraisal, on a voluntary basis, to gain promotion positions in schools in recognition of quality teaching performance by applying for Advanced Skills Teaching positions (ASTs). These positions are linked to higher pay and are generally associated with further responsibilities and specific roles in schools. In most cases, teachers do not have to be at the top of the salary scale to apply for these positions which entails a thorough assessment of their performance. Advanced Skills Teaching positions, which exist in almost all educational jurisdictions, for the most part accomplish two important functions: the recognition of advanced teaching skills with a formal position and additional pay; and a better match between teachers' skills and the roles and responsibilities needed in schools through competitions to gain the positions. These have the benefit of rewarding teachers who choose to remain in the classroom rather than to move into management positions.

AST positions embody two key concepts in the teaching profession in Australia. First, they recognise the need to introduce career diversification as a result of the greater variety of roles in schools – e.g. departmental head, team leader, and manager of curriculum development and/or personnel development. Second, they reflect the need to reward teachers for their developing skills, performance and responsibilities, in what constitutes a competency-based professional career ladder. Teachers, as they access AST positions, are expected to have deeper levels of knowledge, demonstrate more sophisticated and effective teaching, take on responsibility for co-curricular aspects of the school, assist colleagues and so on. Access to AST positions involves formal appraisal processes which are more summative in nature.

- New South Wales introduced the Highly Accomplished Teacher (HAT) position in July 2009. The HAT position is an initiative of the Smarter Schools National Partnership on Improving Teacher Quality. A HAT is an excellent teacher who models high-quality teaching for his/her colleagues across the school and leads other teachers in the development and refinement of their teaching practice to improve student learning outcomes. HAT positions are classroom-based positions with a reduced teaching allocation to enable them to mentor other teachers, including student teachers, beginning

Box 4.8 Teacher career structures in Australia and Singapore (*continued*)

and more experienced teachers, work with university partners and take a role in the school's leadership team. HATs are appointed through a merit selection process which requires, as a prerequisite, application to the New South Wales Institute of Teachers for consideration of accreditation at Professional Accomplishment or Professional Leadership. These positions are two-year appointments and are limited to 100 positions over the life of the National Partnerships.

- The **Northern Territory's** Accomplished Teacher status requires applicants to participate in an "inquiry process" over 12 months, based on the Northern Territory Teacher Registration Board Accomplished Standards of Professional Practice for Teaching. The assessment of performance is undertaken by assessment panels and moderation committees and includes the appraisal of teaching modelling and role in curriculum and professional learning. This process was being reviewed in 2011.
- In **Tasmania**, the Advanced Skills Teacher position recognises outstanding classroom teachers and leading staff members. It is targeted at teachers recognised as exemplary practitioners, who are accorded additional responsibilities within their school. It is a promotion available to any permanent teacher who satisfies the application process, operating in a similar way to a salary increment. Positions are advertised by individual schools on a needs basis.
- The **Victorian** school system includes one promotional appointment for those teachers who want to remain in the classroom: Leading Teacher. The programme is intended to serve the dual purpose of recognising outstanding classroom teachers; and providing schools with a human resource to lead various in-school programmes and projects. Schools advertise for Leading Teacher positions on a needs basis – the position is usually associated with a specific anticipated responsibility. The Victorian Department of Education and Early Childhood Development aims to maintain a Leading Teacher profile of 10 to 15% of full-time teaching staff.

Singapore: Linking teacher appraisal to career pathways

The Education Service Professional Development and Career Plan (Edu-Pac) in Singapore recognises that teachers have different interests and aspirations and provides three different career tracks for teachers:

- The **Teaching Track** allows teachers to remain in the classroom and advance to the levels of Senior Teacher, Lead Teacher or Master Teacher. This provides an opportunity for teachers to focus on classroom teaching while obtaining a leadership role along with a senior-level salary.
- The **Leadership Track** provides opportunity for teachers to take on leadership positions within the school or at the Ministry of Education.
- The **Senior Specialist Track** allows teachers to join the Ministry of Education's headquarters and as specialists with particular expertise in specific aspects of education.

The Enhanced Performance Management System (EPMS) serves to support teachers' professional and career development and its results inform promotion decisions as part of Edu-Pac. The EPMS process involves performance planning, performance coaching and performance appraisal. Performance planning involves a teacher self-appraisal and a discussion with the teachers' reporting officer (typically a Head of Department) about target setting and performance benchmarking. Performance coaching is ongoing and includes a formal mid-year review

Box 4.8 Teacher career structures in Australia and Singapore (continued)

between the teacher and the reporting officer. Finally, the performance appraisal at the end of the year includes an appraisal interview and a rating of actual performance against planned performance. Teachers are appraised based on actual achievement as well as potential for future performance. Decisions on the teacher's "current estimated potential" are made in consultation with senior colleagues of the teacher based on observation, dialogue, portfolio evidence and the teacher's contributions to the school and its environment. The final performance grade affects the annual performance bonus received for the year's work as well as promotions to the next level of the career pathway.

Sources: Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.

Lee, C.K. and M.Y. Tan (2010), *Rating Teachers and Rewarding Teacher Performance: The Context of Singapore*, Paper presented at APEC Conference on Replicating Exemplary Practices in Mathematics Education, Koh Samui, Thailand, 7-12 March 2010.

Weinstein, T. L. and K. S. Struthers (2012), "Similar Demands, Different Responses: Teacher Evaluation in the United Kingdom and Singapore", *Educational Policy Analysis and Strategic Research*, Vol. 7, No. 1, International Association of Educators, Urbana, IL.

Build a more elaborate registration system linked to career development

Currently, the registration process appears to serve a limited purpose, as registration is disconnected from teachers' actual classroom practice and performance. Registration does not involve a professional appraisal or attestation of teachers' competencies, and it does not correspond to a step within the teacher career. While this approach has the advantage of emphasising the importance of continuous professional development, it does not make the link between the completion of such courses and actual improvements in classroom practice. Hence, the registration system focuses on recognising formal qualifications more than excellence and improvements in teachers' actual work.

To make registration more meaningful for teachers, its main purpose could be to hold teachers accountable for their practice and determine advancement in the teacher career. This redefinition of teacher registration would convey the message that reaching and demonstrating high standards of competence is the main road to career advancement in the profession. Registration processes that are linked to career development could help provide incentives for teachers to perform at their best, bring recognition to effective teachers, support professional learning, and help recognise and spread good practice more widely. Registration and registration renewal processes could also provide useful information for accountability, hiring and tenure decisions, professional development and promotion opportunities, or, in particular circumstances, responses to underperformance.

One way of re-organising teacher registration along these lines would be to require graduates from initial teacher education to apply to be 'provisionally registered' with the Education Cooperative in order to seek employment as a teacher. Provisionally registered teachers could then apply for full registration upon completion of an induction period, based on an appraisal in relation to the teacher competency requirements. Access to a promotion/higher level of registration could be through a voluntary application process

and teachers should be required to periodically maintain their registration status when not applying to a promotion. See Box 4.9 for an example from Australia.

Box 4.9 Teacher registration in Australia

Registration is a requirement for teachers to teach in Australian schools, regardless of school sector. All states and territories have existing statutory teacher registration authorities responsible for registering teachers as competent for practice. The levels of teaching registration vary according to the jurisdiction. In most jurisdictions, teachers reach the first level of registration from the relevant authority upon graduation from an approved initial teacher education programme. Currently, each teacher registration authority has its own distinct set of standards for registration; however, from 2013 jurisdictions will be progressively introducing the Australian Professional Standards for Teachers (the Standards) which will provide a national measure for teachers' professional practice and knowledge. Advancement to full registration (or professional competence) is achieved after a period of employed teaching practice and, from 2013, an appraisal against the Standards at Proficient level.

In all states and territories, after teachers have initially become registered within their jurisdiction, they must renew their registration. The period of registration varies but is most commonly five years. The main function of the registration process is that of certifying teachers as fit for the profession mainly through the mandatory process of accessing or maintaining "Full/Competence" status – as such, these processes ensure minimum requirements for teaching are met by practising teachers. Registration processes constitute a powerful quality assurance mechanism to ensure that every school in Australia is staffed with teachers with suitable qualifications who meet prescribed standards for teaching practice. At their initial level (provisional/graduate registration), they also provide a policy lever for setting entrance criteria for the teaching profession and, through the accreditation of initial teacher education programmes, strengthen the alignment between initial teacher education and the needs of schools.

Source: Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.

Such appraisal for registration/career advancement would be more summative in nature than the regular appraisal for professional development, and it would need to be ensured that processes are fair and the same standards are applied across schools. While the process should be mostly school-based, led by the school leadership team, there would need to be a stronger component external to the school to validate the process and ensure that practices are consistent across the Netherlands. This element of externality could be introduced via an accredited external evaluator, typically a teacher from another school with expertise in the same area as the teacher being appraised. To this end, the system could build on experience gained from the teacher peer review projects. For example, teachers having participated in the peer review project could be encouraged to apply for the role of external evaluator. External evaluators would need to be accredited by the relevant organisation (possibly the Education Cooperative). They should receive specific training for this function, in particular in standards-based methods for assessing evidence of teacher performance. It would also be desirable to establish moderation processes to ensure consistency of judgements in registration processes.

Given the stakes attached to appraisal for registration, decisions should draw on several types of evidence and encompass the full scope of the work of the teacher. Teacher appraisal for registration could continue to set requirements for continuous

professional development, but should complement this with other instruments to evaluate teacher performance, such as classroom observation and documentation of practices in a simplified portfolio. A portfolio would allow teachers to mention specific ways in which they consider their professional practices are promoting student learning. It could include elements such as: lesson plans and teaching materials, samples of student work and comments on student assessment examples, teachers' self-reported questionnaires and reflection sheets (Isoré, 2009).

Such portfolios could replace the current use of competency files, which are widely considered as a bureaucratic process with little bearing on schools' actual practice. To make teacher portfolios valuable to teachers and schools, it is important that the requirements of a portfolio are closely related to teachers' day-to-day work; the elements required should be a "natural harvest" of teachers' real work rather than something produced in addition to their regular work. In the United States, for example, the National Board for Professional Teaching Standards (NBPTS) offers recognition to teachers who satisfy requirements for a portfolio submission. It is a demanding process for teachers, but those who participate find it to be a rewarding experience because the natural harvest makes the process less burdensome (for more information on the NBPTS, see Box 4.10).

Such a revised system of teacher registration would provide opportunities to recognise and reward teaching competence and performance, which is essential for retaining effective teachers in schools and for making teaching an attractive career choice (OECD, 2005). It would not directly link appraisal results with teacher pay, but instead to career progression (therefore establishing an indirect link with salaries). This is a desirable option as direct links between teacher performance and pay have produced mixed results, according to the research literature (Harvey-Beavis, 2003; OECD, 2005). As such, teacher registration would fulfil the function of recognising formally the knowledge, skills sets and experience acquired in the profession, which presupposes that teachers have access to the related professional development opportunities.

Box 4.10 The National Board for Professional Teaching Standards in the United States

When applying to the National Board for Professional Teaching Standards (NBPTS) (www.nbpts.org), teachers in the United States enter an extensive application process which consists of two major parts: the portfolio of their work including a videotape of a lesson they have taught; and the assessment centre exercises where teachers address a set of questions that relate to the specific content of their field. The assessment is undertaken against detailed teaching standards established by NBPTS. These are based on NBPTS' five core propositions: (i) teachers are committed to students and their learning; (ii) teachers know the subjects they teach and how to teach those subjects to students; (iii) teachers are responsible for managing and monitoring student learning; (iv) teachers think systematically about their practice and learn from experience; (v) teachers are members of learning communities. The standards are developed and reviewed by teachers and other experts.

The NBPTS is designed to consider a wide range of teacher competencies, using videos submitted by the teachers to appraise classroom practice and along with portfolio entries focused on teaching practice and constructed response assessments of content knowledge. Submitted materials are reviewed by trained teachers who are experts in the teachers' content areas. In the United States, the NBPTS has been the chief means of certifying that classroom teachers are performing at high levels. It has been considered as a model for other countries who are interested in standards-based certification systems for teachers (Harris and MacKenzie, 2007; Ingvarson and Hattie, 2008). Nearly all states in the United States allow teachers to take the NBPTS examination as a mechanism for increasing their salary, by tying National Board

Box 4.10 The National Board for Professional Teaching Standards in the United States *(continued)*

Certification to higher salaries. As of October 2012, the National Board had certified 97 000 teachers nationwide, and more than 6 000 became National Board certified in 2011. The Certification is good for ten years and then the teacher must reapply.

Sources: National Board for Professional Teaching Standards website, www.nbpts.org.

Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.

References

- Bal, J. and De Jonge, J. (2007), *Improving School Leadership: Country Background Report for the Netherlands*, available at: www.oecd.org/education/school/38639469.pdf.
- Bokdam, J., et al. (2011), *Tussenmeting Convenant leerkracht 2011*, Regio Plan_Zoetermeer, www.research.nl/files/rvb/reportcenter/Rapporten/B3883/B3883TussenmetingConvenantLeerkracht2011def.PDF.
- Danielson, C. (2001), “New Trends in Teacher Evaluation”, *Educational Leadership*, Vol. 58, No. 5, pp. 12-15.
- Danielson, C. (1996, 2007), *Enhancing Professional Practice: A Framework for Teaching*, 1st and 2nd Editions, Association for Supervision and Curriculum Development (ASCD), Alexandria, Virginia.
- De Bruin, G., S. Van Geel and K. Kans (2013), *Peer Review in de Praktijk, Eindrapport*, commissioned by the Ministry of Education, Culture and Science, Ecorys, Rotterdam.
- Department of Education, Northern Ireland (2013), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Northern Ireland*, www.oecd.org/edu/evaluationpolicy.
- Dutch Education Council (2013), *Being a Teacher*, Dutch Education Council, The Hague.
- Dutch Government (2011), *Teaching 2020, A Strong Profession!*, Dutch Government, www.ecbo.nl/ECBO/ReferNet/docs/11-0315_Teacher_2020.pdf.
- Dutch Inspectorate of Education (2010b), *Begeleiding beginnende leraren*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2012), *5. Lesobservatieformulier: toezicht leraarschap P.O.*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2013a), *The State of Education in the Netherlands: Highlights of the 2011/2-12 Education Report*, Inspectie van het Onderwijs, Utrecht.
- Dutch Inspectorate of Education (2013b), *De kwaliteit van het bestuurlijk handelen in het funderend onderwijs*, Inspectie van het Onderwijs, Utrecht.
- Dutch Inspectorate of Education (2014), *De kwaliteit van schoolleiders in het basisonderwijs, speciaal onderwijs en voortgezet onderwijs*, Dutch Inspectorate of Education, Utrecht.
- Dutch Ministry of Education, Culture and Science (2009), *Loopbaanmonitor Onderwijs*, conducted by Regio Plan.
- Dutch Ministry of Education, Culture and Science (2013), *Lerarenagenda 2013-2020: de leraar maakt het verschil*, Ministerie van Onderwijs, Cultuur en Wetenschap.
- Dutch Inspectorate of Education (2010a), *Bekwaamheidsdossiers in het VO*, Dutch Inspectorate of Education, Utrecht.

- Eurypedia (2013), *European Encyclopedia on National Education Systems: Netherlands*, <https://webgate.ec.europa.eu/fpfis/mwikis/eurydice/index.php/Netherlands:Overview>.
- Goe, L., K. Biggers and A. Croft (2012), *Linking Teacher Evaluation to Professional Development: Focusing on Improving Teaching and Learning*, National Comprehensive Center for Teacher Quality, Washington, DC.
- Harris, W. J. and S.V. MacKenzie (2007), *National Board Teacher Certification in Maine: An Exploratory Study*, Maine Education Policy Research Institute, Orono, Maine.
- Harvey-Beavis, O. (2003), “Performance-based Rewards for Teachers: A Literature Review”, paper distributed at the third workshop of participating countries in the OECD activity “Attracting, Developing and Retaining Effective Teachers”, Athens, 4-5 June, www.oecd.org/edu/teacherpolicy.
- Hess, F. and M. West (2006), *A Better Bargain: Overhauling Teacher Collective Bargaining for the 21st Century*, Program on Education Policy and Governance, Harvard University, Cambridge, Massachusetts.
- Hill, H. and C. Herlihy (2011), “Prioritizing teaching quality in a new system of teacher evaluation”, *Education Outlook*, No. 9, American Enterprise Institute, Washington DC.
- Hobson, A.J., et al. (2009), “Mentoring beginning teachers: What we know and what we don’t”, *Teaching and Teacher Education: An International Journal of Research and Studies*, Vol. 25, No. 1, pp. 207-216.
- Ingvarson, L., and J. Hattie (2008), *Assessing teachers for professional certification: The first decade of the National Board for Professional Teaching Standards*, JAI Press, Oxford.
- Isoré, M. (2009), "Teacher Evaluation: Current Practices in OECD Countries and a Literature Review", *OECD Education Working Papers*, No. 23, OECD Publishing, Paris, <http://dx.doi.org/10.1787/223283631428>.
- Jensen, B., et al. (2012), *The Experience of New Teachers: Results from TALIS 2008*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264120952-en>.
- Kane, T.J., et al. (2010), *Identifying Effective Classroom Practices Using Student Achievement Data*, National Bureau of Economic Research, Cambridge, Massachusetts.
- Kane, T.J. and D.O. Staiger (2012), *Gathering Feedback for Teaching: Combining High-Quality Observations with Student Surveys and Achievement Gains*, MET Project Policy and Practice Brief, Bill and Melinda Gates Foundation, Seattle, Washington.
- Kim, K., et al. (2010), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Korea*, Korean Educational Development Institute (KEDI), Seoul, www.oecd.org/edu/evaluationpolicy.
- Margo, J., et al. (2008), *Those Who Can?*, Institute for Public Policy Research (IPPR) Publications, London.

- Milanowski, A. and S. Kimball (2003), “The framework-based teacher performance assessment systems in Cincinnati and Washoe”, *CPRE Working Paper Series*, TC-03-07, CPRE, Wisconsin.
- Milanowski, A. (2004), “The Relationship between Teacher Performance Evaluation Scores and Student Achievement: Evidence from Cincinnati”, *Peabody Journal of Education*, Vol.79, No. 4, pp.33-53, Routledge, London and Peabody College of Vanderbilt University, Nashville.
- Nusche, D., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Norway 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264117006-en>.
- OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264190658-en>.
- OECD (2010), *Improving Schools: Strategies for Action in Mexico*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264087040-en>.
- OECD (2009), *Creating Effective Teaching and Learning Environments: First Results from TALIS*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264072992-en>.
- OECD (2005), *Teachers Matter: Attracting, Developing and Retaining Effective Teachers*, Education and Training Policy, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264018044-en>.
- Santiago, P. and F. Benavides (2009), *Teacher Evaluation: A Conceptual Framework and Examples of Country Practices*, OECD, Paris, www.oecd.org/edu/evaluationpolicy.
- Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.
- Santiago, P., et al. (2013), *Teacher Evaluation in Chile 2013*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264172616-en>.
- Scheerens, J., et al. (2012), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands*, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.
- Shewbridge, C., et al. (2014), (2014), *OECD Reviews of Evaluation and Assessment in Education: Northern Ireland, United Kingdom*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264207707-en>.

Chapter 5

School evaluation

School evaluation in the Netherlands comprises internal quality care by schools, systematic external evaluation by the Inspectorate of Education, and the publication of school information through online information systems. The school evaluation system is built on the availability of a rich set of data on individual schools and a highly developed system of collecting, managing and analysing school performance information. There has been considerable investment in building capacity and instruments for school self-evaluation, and the Inspectorate also reviews the approaches to internal quality care implemented by schools. The Inspectorate of Inspection operates a risk-based approach, whereby schools at risk of underperformance are evaluated more frequently and more thoroughly than others. There are indications that the risk-based approach has been successful in reducing the number of schools providing weak or unsatisfactory quality of education. To go further in helping all schools improve, the Inspectorate has launched an internal reflection on new approaches to measure the value-added by schools and to develop differentiated inspections for schools already performing at an acceptable level of quality. In developing these approaches, it will be important to continuously review the impact of school development interventions and to further invest in building the professionalism of teachers, school leadership teams and school board members to collect, analyse and use evaluative information for improvement.

This chapter analyses approaches to school evaluation within the Dutch evaluation and assessment framework. School evaluation refers to the evaluation of individual schools as organisations. This chapter covers both internal school evaluation (internal quality care/school self-evaluation) and external school evaluation (inspections), as well as the balance being struck between the two.

Context and features

School evaluation is a particularly important component of the Dutch evaluation and assessment framework. It is strongly connected with the other components of the framework as it uses student assessment results as a key indicator, it comprises important parts of teacher appraisal, and it provides significant inputs to system evaluation. In the decentralised Dutch system, where schools and their owners have extended autonomy, it is generally accepted that the way individual schools are operating has a major impact the quality of student learning outcomes and the overall performance of the education system. The quality of the feedback schools receive and their capacity and willingness to use this feedback to improve their own work is, therefore, of much importance in the Dutch educational context.

Responsibilities for school evaluation

School evaluation in the Dutch system is carried out by several actors and agencies. The school governing boards are key actors in school evaluation. Although the formal owners of schools are either public authorities (such as municipalities) or non-public entities (such as churches or private foundations), they typically do not manage their schools directly but create school boards as intermediary bodies to exercise management responsibilities. Among other management tasks, these school boards are responsible for quality management and the operation of internal accountability functions. In practice, school boards typically delegate most quality assurance tasks to the principals or leadership teams at the school level.

External school evaluation is entrusted to the Dutch Inspectorate of Education. The Inspectorate operates under the supervision of the Ministry of Education, Culture and Science, but is professionally and organisationally independent. Its headquarters are located in Utrecht. The Inspector-General has a mandate for the execution of school supervision. After approval, the Ministry of Education, Culture and Science sends the Inspectorate's annual planning of activities to the Parliament. In 2012, there were 474 full-time equivalent staff (524 persons employed), 400 of which were directly involved in the inspection process. The total budget of the organisation in 2011 was around €56 million (SICI, 2012).

The 2010 legislation *Good Education, Good Governance* has significantly strengthened the role of school boards in quality enhancement (Scheerens et al., 2012). It also led to the strengthening of the role of the national “intermediary” or umbrella organisations: the Primary and Secondary Education Councils (*PO-raad* and *VO-raad*), which are the national associations of school boards. These bodies have become key partners of the Inspectorate in the field of school evaluation and are now providing a number of specific institutional mechanisms and instruments that have become central to the school evaluation process in the Netherlands.

Beyond the Inspectorate and the Primary and Secondary Education Councils, the Netherlands has a well-developed infrastructure for school support services, with most

service providers being private consultancy companies. These service providers offer support for internal school evaluation. They also work on innovation concepts and help address specific challenges faced by schools. Beyond school evaluation in the narrow sense, the different actors described above are typically also involved in related functions such as internal quality management, school improvement, teacher professional development or stakeholder participation. These functions, even if they are not defined as evaluation in the proper sense, are also used for evaluative purposes or provide important inputs to evaluation activities.

Approaches to school evaluation

School evaluation in the Netherlands takes several different forms. As illustrated by Table 5.1, it comprises internal quality care by schools, systematic supervision and risk-based inspection visits by the Inspectorate, as well as the publication of information on schools through school quality cards and a recently launched online information system called *Windows for Accountability*.

Table 5.1 Instruments and mechanisms for school evaluation in the Netherlands

Type of evaluation	Short description	Formal responsibility
School self-evaluation	Internal quality care by schools	Schools
School inspection	Systematic school supervision, using a comprehensive framework of indicators and clear decision rules	Inspectorate of Education
Quality/ supervision cards	Formerly a set of key indicators on school functioning to inform general public and parents. In 2007, quality cards were reduced to an indication of the inspection regime a school has to follow, which is indicative of school performance	Inspectorate of Education
Windows for Accountability	Publicly available information on each school, comprising centrally collected quantitative indicators and qualitative indicators provided by schools	Schoolinfo (a foundation under the responsibility of Councils for Primary and Secondary Education)

Source: Adapted from Scheerens, J. (2013), Addendum to OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands, University of Twente, Netherlands.

School self-evaluation¹

Even though there is no legal requirement prescribing particular school self-evaluation processes, Dutch schools must report on the progress of their students and produce public accountability information on educational results, the quality of education, the financial situation of the school and the arrangements for professional governance. This information is used by the Inspectorate in its external evaluation of schools.

The Law on Primary Education requires schools to produce several strategic documents: (i) an annual report; (ii) a four-yearly school plan; (iii) a school guide (school prospectus). Regulations for secondary schools are similar. These documents make explicit references to quality, performance and strategies for improvement. The regular preparation of these documents, especially that of the school plan which is revised every four years, requires each school to strategically reflect on its strengths and areas for development. School boards are obliged to describe in the school plan how they monitor and improve quality, including how they ensure that school staff maintain and develop

their competencies. Many schools hire external consultants to help them with these processes (Dutch Inspectorate of Education, 2012).

In the early 2000s, the government invested significantly in developing the self-evaluation capacities of schools through two major national development programmes (called “Q5” and “Q Primair”). Although these programmes have been discontinued, there is still significant investment into school-level capacity building in this area. One example is the *Schools have the initiative (School aan Zet)* programme, which was designed to stimulate school improvement. Although self-evaluation is not explicitly mentioned as a priority, the programme has potential to contribute to strengthening quality management and self-evaluation capacities at the school level (more detail further below). Another example is the policy focus on results-oriented work, which also includes a focus on schools using data to improve practices at the school and enhance student learning outcomes.

The Inspectorate has started promoting an approach to quality care in which schools are encouraged to use scores from the LVS student monitoring system to reflect on student growth and learning gains at the school level. The idea is that if schools are given the responsibility for reporting on learning gains, they will develop stronger capacity to analyse their students’ growth patterns and overall improvement trends within the school. For the purpose of external evaluation, the Inspectorate strives to achieve a balance between a range of different indicators, including: (1) schools reporting on the learning gains of their students; (2) performance results adjusted for the impact of student background characteristics (using register data) and adjusted for prior achievement levels in the case of secondary schools; (3) other indicators such as raw test results, study duration and success rates.

School inspection

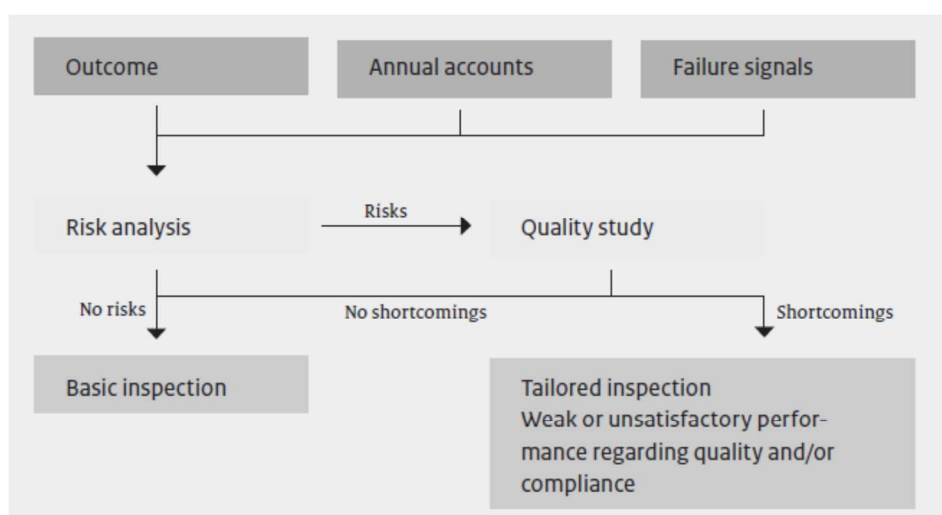
Risk-based inspection

External evaluation in the Netherlands is designed and conducted by the Dutch Inspectorate of Education. School inspection is highly developed and the evaluation approaches of the Dutch Inspectorate are continuously reviewed and revised based on emerging needs. From 2002 to 2007, the Dutch Inspectorate of Education operated a system of “proportional inspection”, in which the extent and frequency of external inspection was to be determined on the basis of the quality of schools’ own self-evaluation. However, the Inspectorate found that the products of school self-evaluation did not meet the quality criteria of the 2002 Supervision Framework.

In its 2005 “vision on supervision”, the Dutch government stressed a more selective approach (Dutch Ministry of the Interior and Kingdom Relations, 2005). Following the 2008 amendment to the requirements on annual reporting for schools (Dutch Ministry of Education, Culture and Science, 2008), the Inspectorate introduced a system of “risk-based inspection”, which can be described as a reinterpretation of the proportionality principle. In this inspection model, the Inspectorate makes a distinction between: (i) schools “at risk”, which receive a full “quality inspection”; (ii) schools “to be trusted”, which are visited only once every four years for a “basic inspection” that checks legal compliance, special needs provision and guidance, but not all aspects of the teaching-learning process. In this approach, the Inspectorate relies on publicly available accountability information, rather than on each school’s own evaluation documents.

The identification of schools “at risk” is based on careful desk-based risk analysis using school-level student performance data, documents submitted by schools to the Inspectorate and “failure signals”, such as complaints or media news. The simplified process of risk-analysis, including its inputs, and the overall scheme of risk-based inspection are presented in Figure 5.1. If the risk analysis is negative, the Inspectorate assumes that that basic quality is achieved and that it can trust the school to assure its own quality (at least until the next risk analysis). In this case, only the regular “basic” inspection is conducted, with a small-scale visit on a four-yearly basis. However, if the school’s quality is judged to be weak or unsatisfactory, the school will receive tailored inspection over the following years until it reaches basic quality. The process for determining whether a schools’ quality is weak or unsatisfactory is further described below.

Figure 5.1 Risk-based inspections in the Netherlands



Source: Dutch Inspectorate of Education (2010), *Risk-based Inspection as of 2009: Primary and secondary education*, Ministry of Education, Culture and Science, www.onderwijsinspectie.nl/binaries/content/assets/Actueel_publicaties/2010/Risk-based+Inspection+as+of+2009+-+printable+version.pdf.

If risks are identified during a primary detection process, the Inspectorate undertakes a quality study or expert analysis. As part of a quality study, the Inspectorate requests additional documentation from the school and conducts a more extensive analysis of available data. There are agreed “signalling values”, or benchmarks for different indicators to signal a potential risk. In high-risk schools, a quality inspection is also conducted, with the focus and evaluation methods depending on the types of risks identified. Quality inspections are typically based on a detailed framework of quality criteria covering the key aspects of pedagogical and organisational processes that may impact on pupil outcomes. The indicator framework comprises five domains, which are broken down into ten quality aspects (see Box 5.1). Each of the quality aspects is further developed into a range of sub-items. The Inspectorate also checks the schools’ compliance with the law and its finances. The Inspectorate has a clear set of criteria and decision rules to determine the benchmarks of basic quality below which a school will be classified as weak or unsatisfactory.

Box 5.1 Quality inspections in the Netherlands: Domains and quality aspects considered in primary education

1. Outcomes

- The outcomes of the pupils are at the level that may be expected on the basis of the characteristics of the pupil population.

2. Teaching-Learning Process

- The curriculum offered prepares pupils for further education and society.
- Teachers allow the pupils sufficient time to master the curriculum.
- The school climate is characterised by safety and respectful interaction.
- Teachers provide clear explanations, organise their educational activities efficiently and keep the pupils involved in their tasks.
- Teachers adapt the curriculum, instruction, time allowed for learning the subject matter and teaching time to accommodate the developmental difference between pupils.

3. Special Needs Provision and Guidance

- The teachers systematically monitor the progress made by the pupils.
- The school guides the pupils in order to allow them to develop according to their capabilities.
- Extra care is provided to pupils who are found to need it.

4. Quality assurance

- The school has a quality assurance system.

5. Statutory Regulations

Source: SICI (2012), *The Inspectorate of Education in the Netherlands: SICI inspectorate profiles*, SICI, Brussels, www.sici-inspectorates.eu/Members/Inspection-Profiles/The-Netherlands.

If the quality inspection reveals no shortcomings, the school can still be re-assigned to a basic inspection regime. If the school shows self-evaluation procedures of high quality and has already started addressing identified risks, a limited external evaluation may suffice. However, if the quality of education is found insufficient, the school will receive a tailored inspection for weak or unsatisfactory quality. If the Inspectorate identifies compliance issues, the school will receive a tailored inspection for compliance with statutory regulations.

Following a tailored inspection, the Inspectorate prepares a detailed report which, after being discussed with the school board, is made public on the Inspectorate's website². The school board is then obliged to prepare an action plan for improvement, which typically involves the launch of an organisational development programme supported by the "flying brigades", a group of school improvement specialists organised within the Education Councils, or other external experts (typically private consultancy agencies). The expected results, areas to be monitored, timeline for improvement and schedule for (potential) interim inspection visits are typically agreed in an inspection plan that the

school board must commit to. The improvement process is completed with an inspection into quality improvement and a final inspection report assigning the school to a new inspection regime.

If weak or unsatisfactory schools do not show improvements, the Inspectorate can exert increasing pressure. It may tighten up the inspection regime and visit the school more frequently and/or issue an official warning to the school. Ultimately, it will report unsatisfactory schools to the Minister, on the basis of which the Minister can impose administrative and/or funding sanctions (Dutch Inspectorate of Education 2009a; 2012).

Introduction of new inspection approaches

A key aspect of school inspection in the Netherlands is the continuous exploration of ways to improve existing practices. At the time of this OECD review, three major ongoing innovation efforts could be observed: *(i)* making inspection more differentiated; *(ii)* integrating separate lines of inspection; *(iii)* strengthening value-added approaches.

Differentiated inspection

While risk-based inspection will remain an important aspect of school evaluation in the Netherlands, the Inspectorate is developing new forms of inspection which will be extended to a wider group of schools that do not fall into the category of weak or unsatisfactory, with the aim of encouraging their excellence. This extension of attention from weak to good schools in order to help them move towards excellence is a key feature of the current government policy and a major contextual factor when thinking about school evaluation. Inspectors will, in the future, make a distinction not only between the categories of basic, weak or unsatisfactory schools, but additional categories of “moderate”, “average” and “good” are also to be used (Dutch Ministry of Education, Culture and Science, 2014a). This policy is announced in the government programme for 2012-2016 and it fits with the ambitions of the Ministry of Education, Culture and Science for achieving excellence both at the student and the school level (Dutch Government, 2012; Dutch Ministry of Education, Culture and Science, 2014b).

Integrated inspection

Until now, financial and pedagogical-didactical inspections have been conducted separately by two different units of the Inspectorate and have been isolated from each other. Recently the Inspectorate has been considering an internal organisational and procedural innovation aimed at integrating these two lines of inspection. This intention is based not only on practical considerations, but also on the growing recognition of substantial linkages between the quality of financial and human resource management at the level of schools and school boards, and the quality of education leading to better learning outcomes. The development towards more integrated inspection was also fuelled a number of cases of financial and organisational mismanagement of schools, which led to calls for stronger supervision of educational governance³.

Value-added modelling

The value-added dimension has been present in the Dutch inspection approach for a long time, but the lack of appropriate techniques has impeded its practical application. One of the bottlenecks in this area has been the fact that although there is a wide range of student performance data available, it is not standardised across schools in a way that would allow the calculation of value-added at school level. At present, the Dutch

government supports a number of pilot studies to develop indicators that will provide useful and credible identification of schools with high relative efficacy. These studies were begun with primary schools in 2011 and in secondary schools two years later. If these efforts are successful, the calculation of value-added might become possible for each school.

Information sources for school evaluation

A key feature of the Dutch system of school evaluation is the availability of data aggregated at school level that can be used as performance indicators. The most important sources of data are the end-of-primary tests and the secondary leaving examinations, but also the outcomes of other regular measurements that are voluntarily administered in the majority of schools (Chapter 3). Most schools use sophisticated data management systems such as the LVS that allow them to not only monitor the progress of each individual pupil, but also to establish school level aggregations, statistical calculations and graphic presentations that can be used for self-evaluation purposes (Chapter 3). These data are typically aggregated at school level and form time series showing not only the current level of school level performance but also trends towards improvement or deterioration. Time series of school level drop-out data are also available for every secondary school. In addition, many schools conduct regular satisfaction surveys among student and parents, which serve as a tool to enhance horizontal accountability and as a potential instrument for self-evaluations.

The different databases are connected and made publicly available through an online information system called Windows for Accountability (*Vensters voor Verantwoording*). This online information system was created at the initiative of the Primary and Secondary Education Councils (*PO-Raad* and *VO-Raad*) to provide information about individual schools. It is subsidised by the Ministry of Education, Culture and Science and managed by the Foundation SchoolInfo (*Stichting SchoolInfo*). *Windows for Accountability* brings together central administrative data, public accountability information and school-based information.

This system has been operational for secondary schools since 2010 (Box 5.2 summarises the indicators included for secondary schools). By mid-2013, 98% of secondary schools had joined the system and it was in the process of being introduced in primary schools. The information for primary schools has since also been made available through the interactive website *Schools on the Map* (*Scholen op de kaart*). *Windows for Accountability* provides a broad range of quantitative and qualitative data on each individual school. Key indicators on individual schools are developed on the basis of information available from the Ministry of Education, Culture and Science (e.g. results in national examinations, number and background of students), the Inspectorate of Education (e.g. type of inspection supervision) and individual schools (e.g. satisfaction of students and parents, school climate). Where possible, results are compared with national averages.

Box 5.2 Windows for Accountability: Indicators for Dutch secondary schools

1. Number of students.
2. Success rates on examinations.
3. Examination marks.
4. Transfer and school leavers.
5. Premature school leaving (drop outs).
6. Co-operation.
7. Student background characteristics (the percentage of students in need of extra care).
8. Choice of school track and subject matter profiles.
9. The so-called “care plan” (for students in need of extra support and care).
10. The satisfaction of students.
11. The satisfaction of parents.
12. Characteristics of schools from which students enter and schools to which students leave.
13. External evaluations.
14. School climate and safety provisions.
15. Market share in the catchment area.
16. Teaching time.
17. Personnel (age composition, teacher absenteeism, professionalisation budgets).
18. Financial indicators (solvability, financial buffers, per student expenditure, school costs).

Source: Scheerens, J., et al. (2012), OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.

The Inspectorate of Education also publishes *School Quality Cards* for individual schools, which include information about the inspection regime schools are assigned to (basic, more intense or extended). In addition, for secondary schools, “attainment cards” are published annually. These contain average marks in national examinations, participation rates across the different education profiles and an indication of student attainment in lower secondary education. It is likely that the *Windows for Accountability* website will replace the *School Quality Cards* in the near future. The Inspectorate has a formal agreement with the providers of *Windows for Accountability* on the usage of the information presented on the website for school inspections, and on inspection results being published through this channel.

Strengths

A balanced approach to school evaluation

The Netherlands operates a highly institutionalised, advanced system of school evaluation based on sophisticated evaluation theories, well elaborated and detailed evaluation criteria, refined instruments and institutional mechanisms and, in particular, the availability of reliable data and data processing mechanisms. This system is based on the cooperation of a number of agencies which, although they differ significantly in goals and approaches, share some basic common principles and are willing and capable to cooperate. The Dutch system of school evaluation has been able to reconcile various approaches which, in other systems, are often in conflict. This appears to result in a good balance between qualitative and quantitative approaches, internal and external evaluation, accountability and improvement functions and vertical and horizontal accountability approaches.

Development of tools and capacity for school self-evaluation

A notable feature of the Dutch school evaluation system is the presence of ambitious programmes aimed at developing the internal quality management and self-evaluation capacities of schools. Although there are indications that not all schools have stable and reliable mechanisms for self-evaluation yet (more on this under ‘Challenges’), there has been significant investment on which future policies can build.

In the early 2000s, the Ministry of Education, Culture and Science, together with the Primary and Secondary Education Councils, the teacher unions and school leaders, initiated two major development projects related to self-evaluation. These programmes were called “Q* Primair” for primary education and “Q5” for secondary education (Scheerens et al., 2012). They supported the development of quality management systems in schools and included components such as the involvement of relevant internal and external stakeholders, the participation in networks for exchanging of information and expertise, and the publishing of school level quality related information.

The current *Schools have the initiative (Scholen aan zet)* programme that is being jointly implemented by the Primary and Secondary Education Councils and the Ministry of Education, Culture and Science, also has components related to school evaluation and quality development. The programme was launched in 2012 with the intention of stimulating school improvement by encouraging schools to enhance the effectiveness of their internal functioning. As part of the programme, the Ministry of Education designated six specific areas of school improvement and encouraged schools to apply for funding to receive (among other things) visits of independent experts as critical friends. Although school self-evaluation is not part of the six designated areas, these development interventions have the potential to help develop capacity within schools to establish mechanisms of internal reflection, self-evaluation and quality management.

Based on the interviews conducted in the Netherlands, the OECD review team formed the impression that schools’ capacity to use data for self-evaluation was variable, but overall there appeared to be considerable attention to continuous data-based analysis and development planning. In many schools, teachers working together as “data teams” regularly use student achievement data to identify problems and find solutions (Schildkamp and Ehren, 2012). As part of the policy on “results-oriented work” at the school level (Chapter 3), this continuous data-based self-diagnosis and use of feedback information is seen as an on-going function of the school that involves a large number of

teachers, rather than an isolated single action conducted by the leadership team at a fixed date. According to Visscher and Ehren (2011), the strong micro-level orientation of results-oriented work draws the attention of school professionals to the improvement of student results and away from accountability uses of performance data.

Another important feature of school evaluation in the Netherlands is the existence of sophisticated data management and evaluation instruments, often developed by private school support agencies to facilitate effective school self-evaluation. According to a report published by the *Standing International Conference of Inspectorates (SICI)* in 2003, there were, at that time, more than 70 different self-evaluation instruments using different background theories and indicators available on the market (SICI, 2003). Some of these instruments offer advanced (IT based) technical solutions for schools to identify their strength and weaknesses. The existence of advanced data management and evaluation instruments, based on agreed and standardised evaluation criteria and indicators, not only creates favourable conditions for the spreading of self-evaluation and internal quality management, but also for making these processes more reliable and more transparent. However, the uneven quality of the available tools also raises a number of challenges (more on this below).

The use and the impact of some of these instruments have been analysed through systematic research (see for example Schildkamp, 2007). The “Q5” and “Q* Primair” provide an evaluation of available data management tools, ranking them according to their reliability, validity, utility, scope of the quality aspects covered, and capacity to generate or enhance school development. This overview likely contributed to supporting the wider use and dissemination of school data management and evaluation instruments.

A well-established approach to external evaluation

The Dutch Inspectorate of Education works as a quasi-independent professional institution that detects quality problems in all parts of the education system and intervenes to improve weak and unsatisfactory schools. It is a highly structured and professionalised institution whose operations follow clear goals and are guided by well elaborated and transparent rules. This is an important strength of the Dutch approach. Clear expectations, norms and standards for external evaluation, and stakeholders who are engaged with and knowledgeable about external evaluation processes appear to have a significant impact on schools (Ehren et al., 2013).

The Inspectorate’s capacity to detect risks and launch effective interventions is well illustrated by the success of the risk-based approach to stimulate improvements in weak and unsatisfactory schools. In its 2011/12 annual report, the Inspectorate reports a significant decrease in the proportion of weak and unsatisfactory schools in almost all sectors and types of education since the introduction of the risk-based approach (Dutch Inspectorate of Education, 2013a). Between 2009 and 2012, the proportion of weak schools decreased from 7.4 to 2.9% in primary education, and from 10.9% to 9.4% in secondary education. Over the same period, the proportion of unsatisfactory schools fell from 1.4 to 0.2% in primary education, and from 1.9 to 0.9% in secondary education (Dutch Inspectorate of Education, 2009b; 2013a). While the overall rate of improvement appears less striking in the secondary than in the primary sector, these figures mask significant differences across educational programmes in secondary education. For although strong improvements could be observed in practical training programmes and basic vocational programmes of the VMBO sector, there were only slight improvements

in pre-university education (VWO) and advanced and mixed/theoretical vocational programmes of the VMBO sector.

The Dutch Inspectorate of Education (2013a) reports that the improvements undertaken by schools labelled as weak or unsatisfactory are often so important that these schools end up reaching above average quality and long-lasting improvements. At the same time, Van Twist et al. (2013) point out that despite the high rate of improvement among unsatisfactory schools, the overall number of unsatisfactory has remained relatively since the introduction of the risk-based approach. This indicates that each year, new schools are joining the ‘unsatisfactory’ category. The authors conclude that while the inspection system appears to be successful in turning around unsatisfactory schools, it seems less successful at preventing new schools from entering this category. Several studies conducted in the Netherlands confirm the positive impact of inspections on weak and unsatisfactory schools (e.g. Scheerens et al., 2005; Ehren et al. 2005, Ehren and Visscher, 2006; Luginbuhl et al., 2009; Twist et al., 2011; Ehren and Honingh, 2012). However, most of them express doubts about the nature and strength of impacts on other schools (this challenge will be discussed below).

Availability of data to support internal and external school evaluation

A major strength of the Dutch school evaluation system is the highly developed system of collecting, managing and analysing pupil performance data and the availability of these data in an aggregated form at school level, including in time series format. The Dutch system is particularly strong in building up large scale datasets that: (i) contain longitudinal micro-level data; (ii) connect various existing databases, and then making these data sets accessible for a wide range of stakeholders in a user-friendly way.

This is well illustrated by the data system *School Drop-out Explorer*⁴ that provides time series of drop-out data at the level of individual schools, cities, regions, and is aggregated at national level. It is publicly available and offers user-friendly graphics and visualisation. It provides access to reliable data on some key performance indicators for each individual school, which creates opportunities to make school evaluation more objective and transparent than in many other systems.

A related significant development has been the creation of the online information system *Windows for Accountability*. The system summarises information about each school, allows comparisons between schools, and grants the opportunity for institutions to share information. The collected data is presented on a multifunctional website providing access to data for different client groups, such as parents, school management teams (including school board members) and national authorities (including the Inspectorate). The Inspectorate, for example, uses the system to access data for its desk-based risk analysis of schools, as well as to publish information about schools generated from its inspection activities. Information is tailored to the type of user, associated with different levels of access, on three different interfaces:

- *The Management Window*: internal information management application for the school, associated with “results-oriented work”, offering the opportunity for comparisons with other schools.
- *The School Window*: the interface for individual schools, which specifically contains information for parents and other stakeholders (e.g. the municipality).

- *The School Choice Window*: the interface for parents, which contains information to facilitate the comparison between schools in relation to indicators of interest for school choice.

The launch of *Windows for Accountability* reflects governance arrangements in Dutch education, whereby school boards play a key role in the management of information alongside the central government. It also highlights the importance of horizontal accountability to local stakeholders such as parents and municipalities, in addition to the existing vertical accountability to educational authorities. *Windows for Accountability* also seeks to reduce the administrative burden on schools relative to the collection and processing of data, and to support the management of individual schools.

This is part of an advanced national education information system that; (i) combines qualitative and quantitative information; (ii) combines central/national data and local data (i.e. less standardised and more differentiated information provided by schools) in order to provide transparent and relevant information for all interested parties.

Progress in the area of value-added measurement

While school-level information on student test results has been publicly available in the Netherlands for a long time, there is a general recognition that evaluating schools on the basis of their pupils' performance can be problematic because it does not take into account their status upon entry. Accordingly, indicators based on students' progress are viewed as offering a valuable complement to indicators based on current performance. Many of the groups interviewed by the OECD review team recognised that league tables based on test results offer very weak signals of school quality and expressed an interest in developing better analyses of school effectiveness in improving student learning. Such studies can generally be classified either as focusing on growth (i.e. learning gains) or as employing value-added models (VAM) – for more information, see Box 5.3.

Box 5.3 Growth models and value-added models

Growth models involve estimating students' test score trajectories, usually drawing on longitudinal data obtained from annual examinations. Typically, these trajectories are estimated for homogeneous groups of students and enable informative comparisons among groups that are useful for both school self-study and national policy planning.⁵ It is important to bear in mind that the utility of the growth metrics depends critically on the quality of the vertical scale from which they are derived.

By contrast, value-added models (VAM) employ some form of statistical regression to establish an expected current test score for a student, based on her previous test scores, her demographic characteristics (possibly) and the characteristics of the class and the school (if available). The difference between the student's actual test score and the expected score is that student's contribution to the school's value-added. When those differences are aggregated over all students, an estimate of the school's value-added is obtained (for that specific grade and subject).⁶ Although this is usually done separately for each grade/subject combination, sometimes these are themselves averaged to obtain a single estimate for each school.

There are many different statistical models used for VAM. Some rely on the existence of a vertical scale, but most do not. It is important to note that for all models, the value-added estimates are normative by construction; that is, the distribution of value-added estimates for the set of schools contributing to the analysis is centred approximately at zero, with zero denoting an "average" school. Schools with positive value-added estimates are considered better than average (in relative efficacy) and schools with negative value-added estimates are considered

Box 5.3 Growth models and value-added models (*continued*)

poorer than average (in relative efficacy). Of course, proper interpretation of these estimates requires consideration of both errors of estimation and the inaccuracy inherent in drawing causal inferences from non-experimental data.

Source: OECD (2008), *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264050259-en>.

As both growth models and VAM require baseline information, the advent of mandatory testing in both the primary and secondary sectors means that the relevant computations could be carried out for all schools (Scheerens, 2013). At the time of the OECD review visit in mid-2013, some small pilot studies of VAM at the school level were being undertaken, comprising 20 primary schools and 29 secondary schools. The Inspectorate was monitoring these studies to carry out a preliminary evaluation of their utility in judging school quality. The pilot studies on learning gains take advantage of the vertical scale on which test results for different grades are reported. They have focused on; (i) identifying schools where students with baseline scores in a particular range have made exceptional progress; (ii) adjusting estimated annual learning gains for summer gains (or losses), thereby taking partial account of extra-school factors; (iii) estimating learning gains in content strands (e.g. spelling) in order to provide a more comprehensive picture of a school's contributions (Scheerens, 2013).

The development of value-added modelling is intended to address concerns regarding the potential misinterpretation of school league tables and to provide a more useful and credible identification of schools with high relative efficacy. It also fits with the Inspectorate's new focus on identifying and promoting "school excellence" and the Ministry of Education's Quality Agenda, which is targeted at enhancing the performance of all schools. In this context, indicators derived from a value-added analysis of students' test score trajectories are seen as a potential way to identify schools that are truly succeeding at supporting student learning (OECD, 2008; NRC, 2010).

The Inspectorate operates as a "knowledge organisation"

The Dutch Inspectorate of Education has many features of an advanced knowledge organisation. It has a dedicated "knowledge directorate" responsible for making relevant data available, supporting analytical work and desk-based research, managing knowledge internally and safeguarding the "organisational memory" necessary for its quality evaluations. The Inspectorate also continuously integrates academic research into its own work. It maintains close communication and cooperation with the Dutch educational research community, for example by inviting doctoral students to study its operation and to publish the results of their research. As a consequence, many publications have been made available in national and international journals about the theoretical basis of its evaluation practice (see for example Ehren, et al., 2005; Ehren and Visscher, 2006; Janssens and De Wolf, 2009; Ehren and Honingh, 2012).

The Inspectorate also appears open to innovations in its own operation, as illustrated by the new approaches to inspection presented in the previous section and recent pilot initiatives to test the applicability of new methodologies, such as value-added models. It

has an internal staff development framework comprising elements such as internal seminars, conferences, team development days, and regular individual performance reviews of inspectors leading to personal development plans. The Inspectorate also has an advanced internal auditing and quality management system that is used to create new knowledge. Thanks to its research-intensive operation, the Dutch Inspectorate of Education seems to be capable of foreseeing a number of future challenges and of identifying appropriate responses to these challenges.

This openness to organisational innovation is further supported by the Inspectorate's close cooperation with other national and international bodies, with a view to improving its own practice. Within the Netherlands, it collaborates with other sectoral inspectorates and plays an active role in the process of state regulative reform (Box 5.4). Beyond national borders, the Inspectorate is an active member of the *Standing International Conference of Inspectorates* (SICI), a European knowledge sharing network of national inspectorates that provides mutual learning opportunities for its members. The Inspectorate has also supported the participation of the Netherlands in a project on governing complex education systems organised by the OECD's *Centre for Educational Research and Innovation*, as part of which the Netherlands School of Public Administration conducted an analytical case study on the Inspectorate's risk-based approach within a complex governance environment (Van Twist et al., 2013).

Box 5.4 Main principles of the Dutch inspection reform

The overall regulatory reform pursued by the Dutch government is an important contextual factor which has significant implications for inspection regimes in every sector of the public service, including education. The most important aim of this reform is lessening administrative burdens and increasing effectiveness. The Inspectorate of Education is a member of the national Inspection Council which defines the general principles of inspection reform to be applied in the various sectors. The general principles of the Dutch inspection reform seem to have a major impact on the way the education inspectorate conceives its own task and defines its own working philosophy. They include the following elements:

- Lessening inspection burden and inconvenience.
- Increasing effectiveness.
- Differentiated inspection.
- Risk based approach.
- Professionalisation of inspectorates and thinking of inspectors as professionals.
- Communicating intensively with partners and the broader society.
- Innovation of inspection methods (including the use of ICT).
- Having an impact on European inspection processes.
- Strategic development of “instrumental frameworks”.
- Collaboration with municipalities and other regulators.
- Supporting impact assessment (“effect measurement”).

Box 5.4 Main principles of the Dutch inspection reform *(continued)*

- Interaction with policy (being involved in policy development at an early stage in order to enhance the enforceability of policies and regulations and to give feedback on their possible effects).

Source: Inspection Council Bureau (2012), Inspection Reform Programme: More effect, less burden!, Inspection Council Bureau, The Hague, www.inspectieloket.nl.

During the interviews in the Netherlands, the OECD review team also learned about the Inspectorate's emerging "environment model", which understands that the impact of inspections on schools depends increasingly on multiple factors in the schools' immediate environment. This means that the Inspectorate has to collaborate with other actors and take their actions into account to achieve synergies between different parallel processes. In this emerging intellectual framework, inspection is seen increasingly as an "intervention" that requires intelligent impact assessment (Inspection Council Bureau, 2012). This focus on impact assessment is reflected in a number of research projects that have aimed to explore the influence of school inspection on school performance. Compared to other countries there is relatively rich research evidence in the Netherlands on the impact and the use of inspection results. Studies on the causal relationships between actions of inspection and performance improvement are also available, some of them analysing underlying impact theories (e.g. Scheerens et al., 2005; Ehren et al. 2005, Ehren and Visscher, 2006; Luginbuhl et al., 2009; Twist et al., 2011; Ehren and Honingh, 2012).

Challenges

The success of school evaluation depends on the capacity and collaboration of multiple actors

An important challenge related to school evaluation in the Netherlands is the involvement of multiple actors. However, this challenge can also be turned into an opportunity and even a strength. Effective school evaluation requires intensive cooperation between the different actors, as well as the continuous development of shared goals and principles to guide the practice of school evaluation. Beyond the school boards and the Inspectorate, there are several further actors and stakeholders whose actions and views have to be taken into account. These are:

- The Netherlands Court of Audit (*Algemene Rekenkamer*) when performing either its regularity or its performance audits might visit school boards and individual schools, and these visits might result in relevant evaluation statements.
- The Education Cooperative (*Onderwijscoöperatie*), created by teacher unions and professional organisations in 2011, has also become a key player not only in fostering teacher professionalism but also in the evaluation of both teachers and schools. This organisation is a strong promoter of the peer review approach, encouraging teams of teachers to mutually evaluate schools, and it helps this process by developing tools for evaluation.

- The National Association of School Leaders (*Algemene Vereniging Schoolleiders*, AVS), is another important organisation that supports school leaders, among others, in developing internal evaluation and quality management systems.
- A particularly important player in the field of school evaluation is the large national network of private or semi-private consultancy agencies, most of them associated in Edventure, an organisation representing approximately 2500 consultants. Among the many forms of service products offered (sold) to school boards and schools there are, as mentioned earlier, many products related to school evaluation. School boards can, for example, buy audit services simulating the formal inspection, which allows them to prepare themselves for the live evaluation by the Inspectorate.⁷
- The Primary and Secondary Education Councils have developed a range of tools and services to support schools and school boards in the evaluation and improvement of schools. In particular, they set up the “flying brigades”, a group of school improvement consultants who provide school development support, typically at the request of school boards following the identification of an unsatisfactory school by the inspectorate.
- Finally, the educational research community is also important as its members appear in many roles in the complex “ecosystem” of individuals and organisations operating together the Dutch system of school evaluation.

Assuring coherence when several actors are pursuing different goals and following different approaches seems to be a major challenge. The dominant professional paradigms, the key instruments, and the way evaluation results are interpreted and used are determined by the complex interplay of all these actors. Although some players are more important and have a more formal role than others, none can fully determine the way schools are evaluated or the feedback schools receive, and none are isolated from the others. Both the risk-based and the differentiated approaches of the Inspectorate imply that the majority of schools do not receive full quality inspections. Many school evaluation activities happen outside the formal activity sphere of the Inspectorate through self-evaluation, supervision by school board members, quality management processes containing evaluation components or through “horizontal accountability” mechanisms to parents and local communities. In this system, the appropriate functioning of school evaluation requires a dense network of communication and cooperation, and the coherence of the various approaches and practices can be assured only if all the relevant actors are able and willing to cooperate.

Building evaluation capacity at the school and school board level

As in-depth external evaluation of schools in the Netherlands is currently limited to a small fraction of schools (those perceived “at risk”), the quality of self-evaluation is crucial to ensure the on-going improvement of all schools. Given the importance of internal quality care in the Dutch school evaluation model, it is key to determine who is responsible for this function and to ensure that those in charge have the necessary competencies. Hence, the way responsibilities for school quality care are shared between the school boards and the principal or school leadership team deserves particular attention.

The policy changes described in the legislation on *Good Education, Good Governance* emphasise the role of school boards in the internal quality assurance of

schools. The revisions made to the Inspectorate's *Supervision Framework* over the past few years also seem to indicate decreasing attention to the role of school leadership and management in quality care. While the 2005 version of the *Supervision Framework* included a number of indicators specifically concerned with the role of school leadership and management in internal quality assurance, these indicators were not maintained in the later versions of the Framework (Scheerens et al., 2012).

The fact that the Inspectorate's attention seems to have shifted away from school leadership and management and more towards school governing boards raises a number of challenges. While it is logical that inspectors provide feedback to the boards as school owners, it seems problematic that the quality of the work of the principal or leadership team might no longer be evaluated systematically by the external evaluators. In fact, most of the substantial work aimed at realising ongoing school self-evaluation activities can be done effectively only by those who are directly responsible for management tasks, that is, school leaders or internal management teams (Pont et al., 2008; OECD, 2013). If the capacity of the school leaders to do this effectively is not evaluated systematically by external evaluators, one of the most important components of external feedback might become neglected.

School boards are expected to conduct regular appraisal of school leaders, with a view to both holding them to account and providing feedback for improvement. However, it is more typical for school boards to rely on the information provided to them by school principals, and there is little evidence that school boards conduct regular appraisal and review of their principals' work (Dutch Inspectorate of Education, 2013b). In a recent study involving 44 school boards from different sectors, the Dutch Inspectorate of Education (2013b) found that it was not common practice for school board members to invite school principals for a review or to visit schools with an evaluative focus.

More broadly, school boards vary with respect to legal status, size and scope of responsibility and they appear to have uneven capacities for assuring quality at the school level. The different entities that set up school boards (e.g. municipalities or religious communities) enjoy large freedom in the way they regulate their remits and operational procedures (Chapter 1). The members of school boards are typically lay persons who may not have received professional training to conduct school evaluation, and the size of their permanent staff is often quite small. While some school boards formulate ambitious performance demands and exert pressure on their schools to improve, others seem to be less committed to quality goals and do not make appropriate efforts to raise quality ambitions.

Articulating internal and external school evaluation

While "quality assurance" is one of the five domains reviewed by the Inspectorate as part of external school evaluations (see Box 5.1 above), the use of school self-evaluation documents by the Inspectorate has been limited over the past decade. As part of the previous "proportional" inspection approach (2002-2008), school self-evaluation documents were to be used as the starting point for external inspections. The idea was that the Inspectorate could use the information provided in school self-evaluations as a basis to form its own judgement about a school's quality (MacBeath, n.d.). However, during the first years of working with the proportional inspection approach, it became apparent that a significant proportion of schools in the Netherlands did not have well-developed self-evaluation approaches. In 2003, only 13% of schools met all of the Inspectorate's criteria regarding internal quality care (Scheerens et al., 2012).

Several research reports also raised concerns about schools' self-evaluation capacities. In a study involving 27 primary schools, Blok et al. (2008) found self-evaluation processes to be of low quality. According to Janssens and De Wolf, in 2009 there was only a small proportion of schools that had the capacity to “produce sound school self-evaluations that provide evaluation results of sufficient quality to be used for accountability purposes” (Janssens and De Wolf, 2009, p. 7). In most cases the schools' self-evaluations were found to be too broad and unspecific to be of use for the Inspectorate. Hence, with the introduction of new requirements for schools' annual reporting in 2008, external school evaluation shifted to using publicly available accountability information rather than each school's own school self-evaluation documents, to determine the intensity of external inspections. Where they exist, school self-evaluation documents continue to inform the Inspectorate's evaluation of the school, but these documents are not used as the basis for determining the scope and intensity of external inspections.

If external evaluation does not appropriately value school self-evaluation and the development of internal quality culture, there is a risk that this practice might start stagnating. In the Netherlands, as school self-evaluation and school inspection are conducted as parallel rather than cooperative activities, there is a risk that the methods and criteria they use are not well aligned, and that the Inspectorate and schools may not use “the same language” in their respective evaluation processes. Livingstone and McCall (2005) argue that where approaches to self-evaluation and external evaluation are consistent, teachers are more likely to see the external feedback in a developmental perspective rather than a judgemental one. As the instruments and the criteria used for self-evaluation in the Netherlands are strongly influenced by the market of school improvement service providers, there might be a need to increase transparency and quality assurance in this market.

Bringing teacher professionalism into the focus of inspections

The intention of the government to promote teacher professionalism as a key policy priority creates new demands for school evaluation and the Inspectorate. According to the government's Action Plan on teacher policy, teacher professionalism is understood, among other things, as a “willingness and ability to reflect upon one's own knowledge, effectiveness and skills, and to do so on an on-going basis” and a “willingness to develop and improve as necessary” (Dutch Government, 2011). The Education Council's publication on *Being a Teacher* further stresses the personal capacity of teachers to make well-grounded and autonomous professional decisions in their daily work in classrooms (Education Council, 2013). In this emerging notion of teacher professionalism, a special emphasis is put on horizontal forms of learning, such as professional networks and learning communities.

This focus on strengthening teacher professionalism and teacher's reflective practices has important implications for the work of the Inspectorate. It means that it will need to evaluate, for example, to what extent horizontal learning opportunities are provided in the school, how far the school operates as a learning organisation that fosters the development of professional learning communities, and how pedagogical and organisational innovations are used to enhance teacher learning. Inspectors will also have to broaden their perspective so that they see the school not only as an isolated entity but as a place that operates in an open environment offering various learning opportunities for teachers. They will have to take into account the impact of environmental factors, such as

potential peer groups outside the school or the interaction of the school with external partners providing pedagogical support.

This development might create a need to redefine the balance between the current data-driven inspection approach and evaluation approaches that focus more on aspects of quality that are not easy to measure. It also raises challenges related to the limited resources of inspection teams (often consisting of only one single person), which makes it difficult to evaluate all relevant processes in schools. The current policy priority of strengthening teacher professionalism might also lead to a further redefinition of the role of inspectors in connection with school-level innovation processes. Increasing teacher professionalism is likely to lead to a broadened space for teachers to initiate changes in their practice, and this will make it necessary for external evaluators to clarify their relationship to innovations initiated by teachers (more on this below).

Supporting innovation in the context of differentiated inspections

Supporting innovation often raises the question of where to draw the line between practices that are commonly seen as resulting in good quality and those which are yet unproven. Adopting innovative practices often means entering into unexplored new areas and taking risks.

An interesting example of this is the Technasium programme⁸, which the OECD review team explored in a secondary school in Rotterdam. The Technasium programme creates a “hybrid learning environment” (Zitter and Hoeve, 2012) in which important parts of learning are realised not in school but in workplaces where students are faced with open technological problems. Students participating in this study line work in teams on research and design projects with external partners from various sectors such as industry, agriculture, health care or education. When evaluating schools that have adopted such innovative solutions, external evaluators need in-depth knowledge or competences (such as understanding the pedagogical processes of innovative learning environments or managing innovation and change in organisations), which might not have been part of their usual evaluation instruments.

This challenge is directly connected with the introduction of “differentiated inspection” and the extension of external evaluation to schools that already offer acceptable quality services. As mentioned above, the Inspectorate will conduct quality inspections (that is, external evaluation covering a wider range of quality aspects) in schools whose students are achieving average or good results. This will make it necessary for the evaluators to elaborate and apply a broader range of evaluation criteria and instruments in order to move towards higher levels of differentiation and lower levels of standardisation. This will create a more complex evaluation environment for inspectors, and require more sophisticated professional competences. This move towards greater differentiation and the stimulation of good schools to become even better seems is a key challenge facing the Inspectorate in the near future.

Concerns that standardised student assessment results are becoming too dominant in school evaluation

It is clear that within the evaluation and assessment framework, the Cito end-of-primary tests and the secondary school-leaving examinations are not only high-stakes for students, but they are becoming increasingly so for individual schools. Both measures are extensively used for school inspection and for school choice by parents through their publication in *Windows for Accountability*. This will become even more the case when

the end-of-primary-education test becomes mandatory and when the timing of test administration is moved from February to April, as foreseen by the new legislation on student assessment (Box 5.5). Inevitably, this risks increasing strategic behaviour on the part of school agents.

The role of end-of-primary tests and secondary school examinations in the evaluation and assessment framework is faced with challenges of two types. First, an assessment, such as the end-of-primary test, is limited in the range of learning outcomes it covers. Not only is the test offered only in Dutch, mathematics and study skills, but its nature, with mostly multiple-choice and open-ended writing tasks, implies that only a subset of student learning objectives are tested. Other curricular areas and their learning objectives are not externally assessed in a way that allows a comprehensive national picture of how students are mastering the entire set of student learning objectives. Similar considerations apply to secondary school-leaving examinations, although here examination tasks are more varied and a broader range of curricular areas is examined.

Second, externally-based student assessments can produce a number of undesired effects, including adverse educational practices, if results are high-stakes for students and schools (as is currently the case). For instance, the publication of results at the school level may lead to a possible narrowing effect on the curriculum due to an overemphasis on aspects that are tested through the assessments, and time may be diverted from the regular curriculum for special assessment preparation. In addition schools that perform satisfactorily may become complacent as the spotlight falls on those schools performing less well comparatively. There may also be negative effects on teacher-based assessments, with teachers trying to “replicate” external assessment formats to prepare students for these high-stakes tests, among others (Rosenkvist, 2010; Morris, 2011; Santiago et al., 2011).

The Education Council has already expressed caution about an excessive focus on “measurement” (Scheerens et al., 2012) in the new requirement for schools to conduct the end-of-primary test, which is perceived as too narrow in what it measures. The Education Council, in its emphasis of the broader developmental aims of schools, also expressed concern about how school autonomy could be affected in realising a broad set of educational and pedagogical objectives. Representatives of the Inspectorate also expressed the need to work with multiple indicators of quality, for example the balanced score card model in which different indicators are opposed in order to obtain a balanced view of the school instead of focussing excessively on one specific indicator. This implies that output results should not be used as the only criterion for quality and that caution is called for when using performance indicators to make decisions about school resource allocations.

A related challenge lies in the trend, in the Netherlands and elsewhere, to publicly report school examination results in the form of league tables. While the Inspectorate of Education does not publish any school rankings, a number of commercial organisations in the Netherlands do. Several of these rankings use statistical adjustments to take account for the impact of the school context (e.g. number of students receiving additional funding) on test results. However, the existing rankings do not take into consideration the students’ prior achievement. Such rankings of schools can be misleading as they ignore differences in initial academic achievement and other contextual elements that are difficult to measure. Schools at the top of the rankings tend to be those that enrol more advantaged students and so are poor indicators of school effectiveness (Raudenbush, 2004; Foley & Goldstein, 2012; National Association of Secondary School Principals, n.d.).

The publication of school performance measures aims to stimulate school efforts to improve performance, provide information to parents for school choice and reduce information asymmetries to inform an effective allocation of resources (Faubert, 2009). However, there is wide consensus in the literature that such league tables are coupled with several methodological challenges (Rosenkvist, 2010). Notably, school performance tables typically are only able to meaningfully distinguish schools performing at the top and bottom ends of the performance distribution whereas differences among the majority of schools are rarely significant (OECD, 2013). Representatives of the Inspectorate expressed the view that the publication of such rankings is not advisable because the data is not sufficiently reliable. Nonetheless, these league tables appear in the media, and some schools are placed undeservedly in a poor light.

Box 5.5 Risks stemming from changes to the Dutch legislation on end-of-primary testing

There are some risks stemming from the new policy of making the assessments at the end of primary school mandatory, but allowing schools to choose the vendor. On the one hand, having test results available from all schools makes their use for school accountability more attractive. On the other hand, with multiple vendors, results from different assessments are likely not to be fully comparable and schools may start “shopping around” for the assessments they believe will give more favourable results. The risk, then, is that the aggregate scores will be less meaningful than would otherwise be the case.

In a letter to the Second Chamber of the Dutch Parliament, the Secretary of State for Education emphasises the need to ensure that the test scores of different end-of-primary tests are made comparable. This is also described in the new laws on testing in primary education. However, research indicates that equating different examinations in order to obtain comparable results is not easily accomplished. The statistical methodology of equating is usually employed to place the scores of different forms of the same test on a common scale. These forms are constructed to be highly similar with respect to both content and psychometric properties, such as the distribution of item difficulties and reliability. As Feuer et al., (1999) point out, equating tests made by different vendors is not recommended, as the differences in psychometric characteristics among them are likely to be much greater than is the case with multiple forms of the same test prepared using a single vendor. Although equating can almost always be carried out as a purely numerical procedure, inappropriate application of this methodology reduces the validity and utility of the resulting scores. In the Dutch setting the different vendors will be employing the same syllabus and reference levels, but there is no guarantee that the test blueprints will agree to the degree required to support equating. Nonetheless, it may be worthwhile to investigate the tests once they are operational to determine whether scores on some or all could be linked to a common reporting scale.

Another potential challenge is the change in the timing of the end-of-primary test from February to April. As the admission procedures for secondary school will begin before the test results become available, the test will be less useful for secondary schools in the selection of students. Only when a student obtains a higher score than the school’s recommendation would suggest, the school will have to reconsider its recommendation and the test results will influence the student’s options for secondary education. Conversely, for students who have already received a recommendation from their school that allows them to enrol in the secondary school of their choice, the stakes in the end-of-primary test will be lowered. Originally, the main purpose of the examination was to provide information to be used in the transition to secondary school, and secondarily as a component of the portfolio used for school accountability. With the change in timing, the first purpose will be less salient for most students, hence lessening the stakes for students but, presumably, not for the primary schools. This reinforces the risks that the

Box 5.5 Risks stemming from changes to the Dutch legislation on end-of-primary testing *(continued)*

test results will be used for purposes other than the summative assessment of individual students (see above).

The difference in stakes for students and schools may also make the test scores less useful for accountability over time. Elmore (2004) notes that in lower-stakes settings, students in different schools differ with respect to their usual level of engagement in test-taking and, therefore, in the level of effort expended. Consequently, in the future, the variation in test results among schools may be due not only to actual differences in proficiency, but also to differences in motivation and effort. There is a longstanding concern with the interpretation and utility of student performances on assessments that have few or no consequences for the students. The problem is that motivation may be low in general and particularly so for certain sub-populations. Thus, in comparing schools on the basis of aggregate student performance, schools with students who have greater intrinsic motivation will be advantaged, and those with students who have lesser intrinsic motivation will be disadvantaged. The validity of using such comparisons for purposes of accountability is therefore reduced.

It has been suggested that employing computerised adaptive testing (CAT) can ameliorate the problem, at least to some extent, by administering questions that are more precisely targeted at a student's level of proficiency. A recent review article (Wise, 2014) concludes that there is some evidence that, in comparison to fixed tests, CATs can induce greater motivation, especially among lower performing students. Student's response records from a CAT can also be used to evaluate differential motivation. This review also suggests strategies for modifying CAT algorithms to enhance the motivational impact.

Source: Authors' own analysis.

Challenges related to the introduction of value-added modelling

Various strategies have been suggested to “level the playing field” in comparing schools on their effectiveness in promoting learning. The most popular approach relies on value-added models (VAM) that use regression analysis in an attempt to adjust for between-school differences (OECD, 2008; NRC, 2010). This takes into account the differences in intake among schools, so as to better “isolate” the contributions of schools to their students' progress.

As described above, there is increasing interest in the Netherlands for value-added analyses, and a number of small pilots have been launched to move forward in this area. Although the use of VAM addresses some of the concerns with conventional league tables, it does raise other issues, for example, the testing regime in use in the Netherlands. Value-added analysis requires prior measures of academic achievement as predictors. In the Netherlands, LVS results could serve this purpose, however, for most children these tests are set by individual schools and so there is no common baseline with which to begin the analysis. This endogeneity problem is difficult to deal with and could lead to strategic behaviour on the part of some schools.

Moreover, it is unlikely that the covariates employed in the regression model fully capture the differences among schools in enrolled students. These covariates are also subject to measurement error. The consequence is that the value-added estimates attached to schools will be biased to some degree, with the amount of bias difficult to determine. In addition, the usual interpretations of the value-added results depend on assumptions

about the psychometric characteristics of the score scale and the relationships among student scores that are difficult to fully justify.⁹ Thus, most methodologists who study VAMs urge caution in their use. If VAM results are used not only for monitoring, but for accountability as well, this may lead to unfair comparisons among schools.

Timmermans (2012) investigated the application of a variety of VAMs in different sectors of the Dutch system with different outcomes. The findings show that different models (e.g. use of different sets of covariates as predictors) produce value-added estimates that are highly correlated, although the classification of schools into categories such as over-performing, average, or under-performing varied with model choice. Overall, it was found that differences in value-added among schools were relatively small and the (estimated) uncertainty associated with the estimates of school value-added were quite large. Thus, value-added analysis is a somewhat crude tool for accountability. Based on the extensive empirical analyses conducted, Timmermans points out a number of difficulties in applying VAM in a school context, but suggests that despite these difficulties, it will become increasingly used in school accountability, as the other indicators of school quality are also problematic. This supports the use of multiple measures in evaluating school quality.

Although this issue did not arise in OECD visits to schools, the introduction of school quality indicators based on value-added analyses is likely to heighten concerns about the multiple use of tests, and particularly that such high-stakes uses will lead to problematic distortions in curriculum and instruction, as well as heightened anxiety for students and educators. Some of our interviewees expressed a concern that excessive reliance on tests for school accountability risked making teaching a low-status job.

Finally, the complexity of the statistical models involved in VAM makes such approaches difficult to understand for teachers, parents and the general public. Unfortunately, the superficial, and often sensationalist, treatment of test results in the media often adds to the confusion. This highlights the challenge in improving the understanding of parents, policy makers and the public at large of assessment results and their implications for education – from that of a particular child, to the cohort in a particular grade in a specific school, to the country overall. The *Windows for Accountability* website could potentially help promote a broader understanding by parents and others of different elements of school quality.

In January 2014, a research project conducted by Janssens et al. (2014) reported its conclusions from a pilot study on learning growth models and value-added measurement in primary education. The study finds that while value-added measurement has the scientific advantage of being able to isolate school effects on student learning, the techniques for calculating these effects are so complex that the interpretation and use of this information may not be straightforward for stakeholders. Learning growth models, on the other hand, can provide helpful information for school self-evaluation and improvement processes, although they do not separate the school effect on student learning (Janssens et al., 2014).

Schools could take advantage of the LVS assessment programme that, in principle, enables the calculation of learning gains in a particular subject at the student level and, when aggregated, at the school level as well. The issue then becomes one of interpretability. What is a reasonable target for a gain? Should it depend on the student's starting point and, if so, in what way? How should a student's gain, in conjunction with her current status, inform instructional decisions? At the school level, what is a

reasonable target for an average gain and what degree of variability in gains is acceptable?

The answers to these questions have both criterion-referenced and norm-referenced aspects. That is, they depend to some extent on substantively grounded interpretations of both gains and status. These, in turn, must rely on the design of the LVS assessment system and the consequences of the vertical scaling that enables the calculation of gains. But the answers also depend on the ability of schools to compare their results with those of other schools, especially those enrolling similar kinds of students. These norm-referenced interpretations require support from a central body to collect, organise and disseminate the results. Although the criterion-referenced interpretations are more directly relevant to pedagogy, in the absence of norm-referenced comparisons they can be both less useful and less effective in promoting need changes.

It is certainly appropriate for schools to assume the responsibility for monitoring their students' progress and for making good use of the evidence derived from analysis of the LVS results. However, there is always the possibility that the aggregated results will be used, if only informally, for some kind of accountability. In that case, the issues already raised concerning the use of VAM are heightened, as raw average gains scores have very problematic properties when used as an indicator of school effectiveness.

Clarifying the role of the Inspectorate in school improvement

Although the Dutch school inspection system already operates as an advanced knowledge organisation capable of using its knowledge capacities to understand future challenges and elaborate appropriate responses to them, in the near future, further efforts might be needed in the organisational domain.

The task of “stimulating” quality improvement, stipulated by the 2002 Law on Inspection, required not only a redefinition of the role of the Inspectorate, but also a renewal of its approach to inspection and its repertoire of evaluation instruments. This has been a major challenge, and without significant internal innovations would not have been possible. Although the Inspectorate does not directly take part in the quality improvement process, it has to cooperate with other actors (such as the “flying brigades” set up by the Primary and Secondary Education Councils) that are directly responsible for quality development. It also has to prove, through impact analysis, that it can have a positive impact on quality. This requires that inspectors understand the dynamics and the complexity of quality improvement processes, which are very different from ensuring compliance with legal requirements (Twist et al., 2011).

The context in which external evaluation will operate in the future will require significant capacity building within the Inspectorate. Inspectors will have to maintain a balance between the current data-driven approach and new trends towards evaluating aspects of quality that are not easy to measure. This implies being aware of the risk that schools focus excessively on improving test results, to the detriment of broader national learning goals. Given that the Inspectorate will have to cooperate with various partners, there is a need to continuously enhance cooperative capacities. Moreover, since inspection is increasingly using data provided by others, including schools uploading data on the *Windows of Accountability* website, there will be a growing need to guarantee the quality of these data. There is also a challenge for the Inspectorate in ensuring that inspection reports are used effectively by stakeholders in order to strengthen horizontal involvement and accountability. According to research cited by Scheerens et al. (2012, pp. 58-59) schools and parents make less use of inspection results than expected.

Policy recommendations

School evaluation is a well-developed component of the Dutch evaluation and assessment framework. Its importance may be reinforced by the emerging focus on teacher professionalism, which is likely to strengthen support for enlarging the autonomy of schools and teacher communities, while reinforcing the role of national authorities in identifying and sanctioning service providers of poor quality. Building on existing strengths, the OECD review team suggests further policy development in the following areas:

- Realise the potential of multi-actor involvement in school evaluation.
- Build the capacity of school boards.
- Enhance school leadership and school self-evaluation capacity.
- Explore new ways to support teacher professionalism through inspection activities.
- Further develop differentiated inspection.
- Further explore the formative use of value-added information for school evaluation.
- Ensure continuous evaluation and improvement of how inspection approaches stimulate quality.

Realise the potential of multi-actor involvement in school evaluation

Given the historical, social and political background of the Netherlands, there is no alternative to the multi-actor nature of the Dutch school evaluation model. In this system the only realistic option to further improve school evaluation is to continuously evaluate the performance of each of the relevant actors, and to improve both the quality of their actions and the way they cooperate with each other. The Netherlands has the potential to benefit from the multi-actor nature of its school evaluation system, turning this into an advantage while avoiding the risks that occur when sharing responsibilities.

Each actor, including the Inspectorate, will need to define its place in a complex ecosystem of actors and partner institutions, and be aware of the behaviour of all other actors who may influence their own potential to have an impact. The Inspectorate should continuously map the environment in which it operates at both a national and local level, and take into account the potential impact of the other partners when defining its way of working and designing its evaluation instruments.

Build the capacity of school boards

The Inspectorate, as the most important actor in school evaluation, is already operating in a highly effective way, although it is facing new challenges in adapting to a changing environment. As the capacity and performance of school boards, which are the second most important actor, seem to be weaker, efforts should be concentrated on improving their capacities and the quality of their work. This is not an easy task as the landscape of school boards is extremely uneven.

The management capacity of school boards should be improved through supporting the professional development of their permanent staff. Research on the operation of school boards that reveals their capacities, possible shortcomings and potential should

also be supported. There is a need to develop a differentiated approach to deal with school boards of different size and capacities, taking into account that some boards are highly professionalised organisations while others are more loosely coupled formations of volunteers.

The Primary and Secondary Education Councils are already playing a key role in this area and seem to be the most influential strategic partners of the government in its efforts to improve the quality of the operation of school boards. The on-going *Schools have the initiative* (*Scholen aan zet*) programme seems to provide a good framework for this as it places the Primary and Secondary Education Councils into a strategic developmental role. The *Windows for Accountability* (*Vensters voor Verantwoording*) platform also seems to be a good framework for strengthening partnerships between the Primary and Secondary Education Councils and the government. The efforts to extend this platform to the primary sector and to ensure that each school is effectively collecting and uploading the relevant qualitative and quantitative data to be used as quality indicators should be continued.

The Primary and Secondary Education Councils are well placed to support stronger cooperation and partnerships between school boards and enhance the sharing of experiences and mutual learning. There are current discussions to develop regional units to provide adapted education for students with special educational needs (SEN) (Chapter 1). This could also provide a model for regional inter-board cooperation in mainstream primary and secondary schools. Box 5.6 provides examples of local and regional networks in Norway, which have been created to support the development of evaluation capacity among school owners.

Box 5.6 Norway: Local and regional networks for school quality improvement

Municipal networks for efficiency and improvement: In Norway, public schools in the primary and lower secondary sector are “owned” by the country’s 430 municipalities. Each school owner is required to establish and maintain a quality framework for its schools. Some of these municipalities are very small and responsible for a handful of schools only. The capacity challenges faced by these school owners are very similar to those of school governing boards in the Netherlands. In 2002, the Association of Local and Regional Authorities (KS), the Ministry of Labour and Government Administration and the Ministry of Local Government and Regional Development set up “municipal networks for efficiency and improvement” that offer quality monitoring tools for local use and provide a platform for school owners to share experience, compare data and evaluate different ways of service delivery in different sectors. For the education sector, an agreement has been established between KS and the Directorate for Education and Training to allow the networks to use results from the national user surveys that are part of the national evaluation and assessment framework. The networks bring together municipal staff and school leaders to discuss school evaluation and assessment issues and engage in benchmarking exercises. Each network meets four or five times and then the opportunity is offered to another group of municipalities.

Regional groups working on external school evaluation: The Norwegian national school improvement project Knowledge Promotion – From Word to Deed (2006-2010) was launched by the Directorate for Education and Training to strengthen the sector’s ability to evaluate its own results and plan improvement in line with the objectives in the Knowledge Promotion reform. One of the outcomes of the project was the establishment of 11 regional groups to continue to work on school evaluation. These groups received training in the programme’s methodology for external school evaluation and have begun to establish local systems for external evaluation of schools.

Box 5.6 Norway: Local and regional networks for school quality improvement
(continued)

Guidance Corps for school improvement: The Norwegian Directorate for Education and Training has also recently established a “Guidance Corps” of exemplary school leaders who make themselves available to work together with school owners that have been targeted as needing help with capacity development.

Sources: Norwegian Directorate for Education and Training (2011), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Norway*, www.oecd.org/edu/evaluationpolicy.

Nusche, D., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Norway 2011*, *OECD Reviews of Evaluation and Assessment in Education*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264117006-en>.

Enhance school leadership and school self-evaluation capacity

As responsibility for internal quality management and self-evaluation is shared between the school boards and school leadership teams, special attention should be paid to whether both of these partners assume their related responsibility appropriately. While it makes sense that the Inspectorate, as the main external evaluator, provides feedback directly to the school boards, the crucial role of the internal management of the school should not be neglected. This is particularly important in light of the ambitious goal set by the development programmes of the early 2000s to raise the proportion of schools with an advanced internal quality care system to 80% (Scheerens et al., 2012).

School boards have the formal responsibility to ensure that their schools having a reliable internal quality management system and only they have the financial capacity and the administrative power to make this happen. However, the actual self-evaluation activities have to be undertaken by the school leaders and the internal management teams. Therefore the capacities and actions of the latter should remain a strategic aspect of external evaluation. Efforts aimed at improving the capacities of school leaders and leadership teams to establish and operate advanced and reliable internal quality management systems must be continued through further targeted development interventions. Box 5.7 provides an example of targeted training on school self-evaluation conducted in New Zealand for both school principals and members of Boards of Trustees.

Box 5.7 Targeted training on school self-evaluation for school principals in New Zealand

As school self-evaluation (self-review) is at the heart of the New Zealand approach, school capacity to conduct self-evaluation is of key importance. There are high expectations on school principals and their organising bodies (Boards of Trustees) and this can be especially challenging for schools in isolated areas or in communities with low socio-economic status. The external school evaluation body (the Education Review Office, ERO) in collaboration with school principal associations delivers workshops on self-evaluation to school principals, their teams and organising bodies and has developed support materials and case studies in good practice in self-evaluation. Such initiatives capitalised on ongoing professional development for

Box 5.7 Targeted training on school self-evaluation for school principals in New Zealand (continued)

external evaluations and so costs were minimal. In 2009, 35 workshops were delivered by a national facilitator and supporting local senior evaluators to over 1 200 participants across New Zealand, including relatively isolated areas. Workshop feedback was positive and external evaluators are reporting improved self-evaluation processes from schools that attended the workshops.

Source: New Zealand Ministry of Education (2010), OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for New Zealand, www.oecd.org/edu/evaluationpolicy.

There is room to further build on the intention of professional organisations to strengthen the role of peer evaluation and professional networks as platforms for both professional learning and evaluation. There is considerable evidence, for example from Finland, Sweden and England in the United Kingdom, that school-school partnerships, clusters and networks can provide mechanisms for sharing effective leadership and practices that contribute to raising the performance of the member schools (Pont et al., 2008). In England, there have been deliberate efforts to give a “systemic” leadership role beyond their own school to the most successful school leaders through specialised training (Hopkins, 2008). And Box 5.8 presents an example of emerging peer review practices in the Flemish Community of Belgium.

Box 5.8 Emerging school peer reviews in the Flemish Community of Belgium

The Ministry of Education and Training has stimulated collaboration among schools by its promotion of “school communities” (*scholengemeenschappen*). Schools in a similar geographical area join a school community on a voluntary basis. However, the Ministry of Education and Training provides incentives for schools to join a school community by offering extra resources (i.e. extra teaching time for primary and secondary schools). In the case of secondary schools, there are also some organisational advantages to joining a school community. These efforts have successfully stimulated further collaboration among schools and virtually all schools offering mainstream primary and secondary education belong to a school community. There are clearly defined responsibilities for schools and belonging to a school community “implies continuous evaluation and adjustment of school policies” and therefore effectively promotes school improvement (Flemish Ministry of Education and Training and the University of Antwerp, Edubron Research Group, 2010).

Although these emerging collegial relationships are at relatively early stages of development, their emergence is a strength in that they are focusing on helping schools develop both their self-evaluation capacity and the potential for critical friendship. The OECD Review visit revealed an example of primary school principals collaborating with colleagues observing teachers in each other’s schools and an inter-schools quality network between secondary school principals focusing on how to stimulate and improve the use of outcomes. Research points out that schools find peer visitation a useful learning experience (Flemish Ministry of Education and Training and the University of Antwerp, Edubron Research Group, 2010). Further, examples of peer visitation include: a project by the umbrella organisation for Provincial education, in which participating schools commit to a code of ethics, visiting teams write a report on findings,

Box 5.8 Emerging school peer reviews in the Flemish Community of Belgium (continued)

strengths, weaknesses and recommendations, and the visited school decides how to address these recommendations; and visits and peer reviews among schools involved in similar innovation projects in the City of Antwerp.

Sources: Flemish Ministry of Education and Training and the University of Antwerp Edubron Research Group (2010), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Flemish Community of Belgium*, www.oecd.org/edu/evaluationpolicy.

Shewbridge, C., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: School Evaluation in the Flemish Community of Belgium 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, <http://dx.doi.org/10.1787/9789264116726-en>

The Netherlands has been one of the leading European countries in the field of research and development in school management and leadership. For example the creation of a European Network for Improving Research and Development in Educational Leadership and Management (ENIRDELM) was a Dutch initiative. This international role, which also provides good opportunities for mutual learning, should be maintained and strengthened.

Development interventions to enhance school self-evaluation capacities should also focus on increasing teacher professionalism. Teachers should be given an increasing role in fostering professional learning communities that conceive schools as intelligent learning organisations. One of the implications of this is that self-evaluation is seen as less of a bureaucratic process focussed on formally regulating internal organisational processes, but instead as a natural “side product” of intensive professional learning, networked knowledge sharing, and knowledge management based on the increasing richness of quantitative and qualitative evidence. Analysis of the use of self-evaluation in Dutch primary schools revealed that schools undertaking this as part of a “learning organisation” perform significantly better in mathematics (Hofman et al., 2009).

The increasing availability of good quality student performance data at the school level increases opportunities for teacher teams to use data in identifying learning problems and elaborating specific school level interventions strategies. Understanding how data can enhance school improvement, and exploiting the widening opportunities created by the new *Windows for Accountability* platform for school level internal diagnosis and development, might require further investigation. School performance feedback systems can be powerful tools that provide timely, high-quality information on performance, which may help schools to identify problems sooner and examine which types of interventions work better in different contexts (Visscher and Coe, 2003). Timely feedback of performance data in an accessible format are important characteristics of data systems. However, of equal importance are characteristics related to the users and the school organisation (see Box 5.9).

Box 5.9 Factors promoting use of assessment data at the school level

The literature identifies three major factors promoting data use in schools:

- Data system characteristics – timeliness of data availability, accuracy, validity, relevance and reliability of data; access to data; tools available to use the data.
- Data user characteristics – whether they believe in the data, have the necessary knowledge, skills and motivation to use them, whether they feel power to make changes (or whether they feel improvement is contingent upon things beyond their control).
- School organisational characteristics – time is allocated for data use, colleague(s) with special role/expertise in data use, training, teacher collaboration, data use is linked to school vision, norms and goals, school principal supports data use.

Source: Schildkamp, K., L.T.M. Rekers-Mombarg and T.J. Harms (2012), “Student group differences in examination results and utilisation for policy and school development”, *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, Vol. 23(2), Routledge, pp. 229-255.

The self-evaluation capacities of schools can also be strengthened by reinforcing interactions with the Inspectorate. The current move of the Inspectorate towards the introduction of differentiated inspections provides a good opportunity to reconsider the relationship between internal and external school evaluation. As the Inspectorate will be focussing increasingly on schools that already achieve basic quality, self-evaluation documents could gain a greater role in school inspections. While challenges remain in developing schools’ self-evaluation capacities, there are indications that they are improving across the Netherlands. In 2010, just over 50% of schools met all of the Inspectorate’s criteria regarding internal quality care, a significant increase from 13% in 2003 (Scheerens et al., 2012). As schools and school boards are developing stronger approaches to self-evaluation, this opens new perspectives to build synergies between internal and external school evaluation.

Explore new ways to support teacher professionalism through inspection activities

If the current policy focus on teacher professionalism continues, evaluating the internal capacity of schools to support teacher professionalism will become an increasingly important component of school evaluation. This will mean that external evaluators (inspectors and other involved partners) will need to understand the complexities of internal human resource management and development practices in schools. They will also have to understand the nature of teachers’ professional competences, and how these can be improved through advanced school level human resource management and development practices. They will need to assess to what extent schools are already using these practices, and what their strengths and weaknesses are in this area.

The Dutch Inspectorate of Education may need to initiate an internal debate on how the strengthening of teacher professionalism may impact on inspection criteria and methods. Further innovations or pilot experiments may be required, for example,

modifying the classroom observation framework or involving teachers as part of inspection visits. The current plans to integrate financial and educational inspection and to enhance the evaluation of school level human resource management practices should be realised in this context, taking stronger teacher professionalism into account. This could also lead to the creation of broader and more diverse inspection teams.

The fact that the Netherlands does not have a national system of teacher appraisal, and does not intend to establish one, has important implications for school evaluation. It particularly means that the quality of internal human resource management needs to be conceived as a strategic component of school evaluation and based on particularly well elaborated and detailed criteria, in accordance with the policy goal of fostering teacher professionalism (see also Chapter 4).

As teacher professionalism is not yet appropriately reflected in the evaluation criteria of the Inspectorate (see Dutch Inspectorate of Education, 2010), special emphasis could be given to this aspect in the next round of revisions. For example, an enrichment of the criteria related to the internal evaluation of the work and performance of teachers, and the quality of the work of school leaders in connection with this, could be considered. This aspect could be strengthened in the professional development of school leaders, and the developers of self-evaluation instruments and their users could also be encouraged to give more attention to this component.

Further develop differentiated inspection

The challenge of policies that strive for excellence is that not everyone can reach the high goals set, or at least not at the same time. This requires the use of an increasingly differentiated approach in evaluation, including the external evaluation of schools. The Inspectorate has already started to move in this direction with the introduction of differentiated inspections.

An example of a way to establish new evaluation approaches and instruments is a system of national competition and quality awards, which has already been initiated by a new committee established by the Ministry of Education, Culture and Science (Scheerens, 2013). Such an award could become a powerful incentive for already “good” schools to make further efforts to improve their internal quality management practices. These practices could be significantly influenced by the selection criteria for the emerging new quality award.

Furthermore, given that full quality inspections, even in the context of a more differentiated approach, will remain restricted to only a small number of schools, the development of effective school self-evaluation should remain a longer-term target for all schools. This will require an enrichment of the repertoire of quality management and quality improvement techniques. Further efforts might be needed to enhance the capacity of schools to adapt these techniques to their specific level of development and needs.

The approach of the *Schools have the initiative* (*Scholen aan zet*) programme, which allows schools or school boards to define areas of development interventions according to their specific needs and conditions, seems to create favourable conditions for improving self-evaluation capacities. Several components of this scheme have the potential to contribute to building the self-evaluation and quality improvement capacities of schools. For example, the *Schools have the initiative* programme includes goals such as enhancing results-oriented work, human resource management and professional learning organisations. These elements can also directly enhance the school evaluation function

and can be conceived as part of the strategy to strengthen school self-evaluation in the Dutch education system.

Differentiated inspection is likely to receive increased attention in international networks dedicated to educational cooperation and knowledge sharing such as SICI, and also in research on inspection. This will offer opportunities for the Netherlands to intensify cooperation with countries that are moving towards more differentiated approaches to inspection.¹⁰ International comparison may help to identify challenges that accompany the shift from a narrower form of risk-based inspection (where full quality inspection is limited only to “weak” and “unsatisfactory” schools) towards a more comprehensive model (where full quality inspection might also be conducted in schools with acceptable performance).

In this respect, the New Zealand or the Victorian (Australia) models of differentiated school reviews may be of interest for the Netherlands. The New Zealand model, introduced in 2009, combines a differentiated approach with a strong reliance on self-evaluation, and with a significant stress on developmental interventions (Nusche et al., 2012). The extension of full inspection to schools that are not classified as weak or unsatisfactory, and the large variety of external evaluation methods, may lead to the need to revalorise the role of self-evaluation in external evaluation. The Victorian model distinguishes four different types of school review (negotiated, continuous improvement, diagnostic, and extended diagnostic) depending on the specific type of risk identified for the given school (Santiago et al., 2011).

Further explore the formative use of value-added information for school evaluation

In the ‘Challenges’ section of this chapter, it was noted that technical problems with league tables, based on the current status of test results, have led to increased interest in alternative approaches, such as growth and value-added models (VAM). However, given the general concerns regarding the use of VAM for accountability (NRC, 2010; Newton et al., 2010), as well as the findings of Timmermans (2012) using Dutch data, the Inspectorate should move cautiously in formally incorporating VAM estimates into the school indicator system. More empirical work is required before a reliable decision can be made on the choice of the statistical model.

As indicated earlier, one difficulty in the Dutch context is obtaining a suitable set of baseline test scores for estimation process, especially in the primary sector. In the secondary sector, it is usually possible to identify one or two baseline test scores. However, measurement error can lead to biased estimates (Lockwood & McCaffrey, 2013).

Further, the complexity of the statistical models employed in value-added analysis results in a lack of transparency that is likely to meet with strong opposition. Drawing on experiences in the United States and England, considerable effort is required in designing reports of the VAM results (and accompanying explanatory materials) that are understandable by teachers, school leaders, school board members, parents and the general public. Consequently, the current focus should be on helping participating schools to use this information as a component of a more comprehensive self-evaluation effort, while enlisting a broader and much larger sample of schools to participate in future pilot studies.

In moving forward, the OECD (2008) report on value-added analysis provides helpful guidance, and, among other topics, an extensive discussion of issues related to the implementation of VAM. In England, the experience of Ofsted with value-added analysis is also relevant, particularly the substantial investment in developing on-line reports and explanatory materials (Ray, 2006). In both England and the United States, the reaction to the use of VAM results for accountability has been decidedly mixed, with both enthusiastic proponents and vigorous critics in evidence. Again, this suggests the Netherlands should proceed slowly and thoughtfully with VAM.

Ensure continuous evaluation and improvement of how inspection approaches stimulate quality

In the Dutch context, education policy aims to raise already high standards to an even higher level, and school evaluation policy is driven by the intention to strive for improvement and excellence throughout the entire education system. This is increasing the complexity of tasks of the Inspectorate, and it will cope with this only if it enhances internal learning processes through high quality evaluations of its own practices and effective internal knowledge management. The Dutch system of school evaluation seems to be at a stage where there is large potential for further development by focussing on the evaluation and continuous improvement of the evaluators.

To fulfil the Inspectorate's role of "stimulating quality", the efforts to help inspectors better understand the complex dynamic of school improvement should be continued. An effective way of doing this is to demand that external evaluators not only provide static evaluations of the state of quality in schools, but also to review the impact of development interventions and change processes that might lead to improved quality.

At the same time, the intelligent balance that has been struck between different actors and evaluation approaches should be maintained. In particular, the intended separation between accountability-oriented evaluation approaches on the one hand, and more formative school development functions on the other, is a valuable feature of the Dutch school evaluation framework. While the Inspectorate is in charge of guaranteeing basic quality across schools, support for school development is offered by educational consultancy organisation, for example through the *Schools have the initiative* programme. The Inspectorate aims to stimulate quality improvement but to avoid giving advice that is so specific that it would put inspectors in a situation where they must evaluate their own advice (Janssens and van Amelsvoort, 2008). It appears more effective to keep the specialised functions of school evaluation and school development separate and ensure that they are closely connected through regular cooperation between schools, school support organisations and the Inspectorate.

This may require a further enrichment of the already rich theoretical basis of school inspection in the Netherlands, and a further focus on impact evaluation. As the Inspectorate also has a key role in system evaluation, a good synergy could be created between the macro and the micro level perspectives of evaluation. There seems to be potential in the capacity of the Dutch inspection system to apply advanced programme theories when it comes to evaluating either the impact of system level policies, or school level quality improvement interventions.

This implies an even more intensive interaction with those who design and implement either system level or school level development interventions, without blurring the lines that separate them from those who are responsible for evaluation. This could raise the need, among other things, to explore how the quality of feedback given to schools by

inspectors could be improved so that beyond the clear accountability purposes, it could enhance school improvement efforts, in accordance with the explicit role of inspectors to stimulate quality improvement processes.

The Dutch Inspectorate of Education has been particularly active in international cooperation at the European level as one of the founding members of SICI. This strong international engagement in mutual learning and knowledge sharing between national inspectorates should be maintained and reinforced. In this network, there is an on-going reflection on how inspection could be made more effective, and this provides an excellent platform not only to learn from others, but also to share the Dutch experiences with others. For example the debate within SICI on “impact assessment” and how it applies to inspection¹¹ has already led to internal discussions within the Dutch Inspectorate of Education.

Notes

- ¹ The OECD (2013) broadly defines school self-evaluation as an evaluation or review conducted by members of the school to assess the effectiveness of structures and processes in place and the quality of student learning outcomes. In the Dutch context, the ongoing internal processes conducted by schools to review and evaluate the quality of their processes and outcomes are referred to as “quality care” or “quality assurance”. Throughout this report, these terms are used interchangeably.
- ² <http://www.onderwijsinspectie.nl>.
- ³ The need to strengthen educational governance is expressed in the Minister’s letter to the Second Chamber of the Parliament (*Kamerbrief versterking bestuurskracht onderwijs*) of 19 April 2013, available at: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2013/04/20/kamerbrief-versterking-bestuurskracht-onderwijs.html>.
- ⁴ The English version of the *Drop-out Explorer* website is available at: <http://www.vsvverkenner.nl/english/>.
- ⁵ In the Dutch context, data from the LVS or from cohort studies such as COOL could be employed for growth modelling. However, since participation in cohort studies is voluntary, the results from growth modelling may not be representative of the nation as a whole.
- ⁶ This description is necessarily simplified. In practice, various adjustments are made to increase the reliability and stability of the estimates derived from VAM.
- ⁷ See the advertisement of this product (“Audit gebaseerd op het toezichtkader van de inspectie” meaning “Audit based on the supervisory framework of inspection”) on the website of “Onderwijsexpertise.nl” here: <http://www.onderwijsexpertise.nl/dienst/858>.
- ⁸ This is a nationally recognised programme created by a private agency and adopted in an increasing number of Dutch secondary schools. See the website of the program here: <http://www.technasium.nl>.
- ⁹ For further information on these assumptions and their role in value-added analysis, see Reardon and Raudenbush (2009) and Timmermans (2012).
- ¹⁰ For a relevant comparative analysis within the SICI community see van Bruggen (2010).
- ¹¹ See the summary of the SICI Workshop on 23-25 November in Stockholm on the impact of inspection: <http://www.skolinspektionen.se/Documents/Om-oss/report-SICI-workshop-in-stockholm-nov-2011.pdf>.

References

- Blok, H., Slegers, P., & Karsten, S. (2008). Looking for a Balance between Internal and External Evaluation of School Quality: Evaluation of the SVI model. *Journal of Education Policy*, 23(4), 379-395.
- Bruggen, Johan C. van (2010), *Inspectorates of Education in Europe; some comparative remarks about their tasks and work*, SICI, www.aqs.rlp.de/fileadmin/aqs/Service/Inspectorates_in_Europe.pdf.
- Corcoran, S. (2010), *Can teachers be evaluated by their students' test scores? Should they be? The use of value-added measures of teacher effectiveness in policy and practice*, Annenberg Institute for School Reform, Providence, <http://annenberginstitute.org/pdf/valueaddedreport.pdf>.
- Dutch Government (2011), *Teaching 2020, A Strong Profession!*, Dutch Government, www.ecbo.nl/ECBO/ReferNet/docs/11-0315_Teacher_2020.pdf.
- Dutch Government (2012), *Bruggen slaan, Regeerakkoord VVD – PvdA*, available at: <http://www.rijksoverheid.nl/ministeries/ocw/nieuws/2014/03/09/uitdaging-onderwijs-voor-toptalenten.html>.
- Dutch Inspectorate of Education (2009a), *Risk-Based Inspections as of 2009: Primary and Secondary Education*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2009b), *De Staat van het Onderwijs: Onderwijsverslag 2007/2008*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2010), *Risk-based Inspection as of 2009: Primary and secondary education*, Dutch Inspectorate of Education, Utrecht, www.onderwijsinspectie.nl/binaries/content/assets/Actueel_publicaties/2010/Risk-based+Inspection+as+of+2009+-+printable+version.pdf.
- Dutch Inspectorate of Education (2013a), *The State of Education in the Netherlands: Highlights of the 2011/2012 Education Report*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2013b), *De Kwaliteit van het bestuurlijk handelen in het funderend Onderwijs*, Dutch Inspectorate of Education, Utrecht.
- Dutch Inspectorate of Education (2012), *Summary Profile of the Inspectorate of Education in the Netherlands*, Dutch Inspectorate of Education, Utrecht.
- Dutch Ministry of Education, Culture and Science (2008), *Richtlijn jaarverslag onderwijs*, The Hague.
- Dutch Ministry of Education, Culture and Science (2014a), *Toezicht in transitie*, available at: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2014/03/28/kamerbrief-over-toezicht-in-transitie.html>.

- Dutch Ministry of Education, Culture and Science (2014b), *Plan van aanpak toptalenten 2014 – 2018*, available at: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2014/03/10/plan-van-aanpak-toptalenten-2014-2018.html>
- Dutch Ministry of the Interior and Kingdom Relations (2005), *Minder last, meer effect. Kaderstellende visie op toezicht*, The Hague.
- Education Council (2013), *Being a Teacher. More focus on personal professionalism*, Education Council, The Hague, www.onderwijsraad.nl/publicaties/2013/leraar-zijn/item5072.
- Ehren, Melanie C., et al. (2005), “On the Impact of the Dutch Educational Supervision Act Analyzing Assumptions Concerning the Inspection of Primary Education”, *American Journal of Evaluation*, Sage Publications, Vol. 26 No. 1, pp. 60-76.
- Ehren, Melanie C. and A.J. Visscher (2006), “Towards a Theory on the Impact of School Inspections”, *British Journal of Educational Studies*, Routledge, London and Society Educational Studies, Birmingham, Vol. 54, No. 1., pp. 51-72.
- Ehren, Melanie C. and Honingh, M. (2012), “Risk-based school inspections in the Netherlands: A critical reflection on intended effects and causal mechanisms”, *Studies in Educational Evaluation*, Elsevier, Philadelphia, 37(4), pp. 239-248.
- Ehren, Melanie C., et al. (2013), “School Inspections and School Improvement: Testing Assumptions on Causal Mechanisms”, *Oxford Review of Education* (submitted and under review).
- Foley, B. & Goldstein, H. (2012), *Measuring success: League tables in the public sector*, The British Academy, London.
- Hofman, R.H., N.J. Dukstra and W.H.A. Hofman (2009), “School self-evaluation and student achievement”, *School Effectiveness and School Improvement*, Routledge, London, Vol. 20, No. 1, pp. 47-68.
- Hopkins, D. (2008), "Realising the potential of system leadership", in Pont, Beatriz, D. Nusche, Hopkins, David (eds.), *Improving School Leadership, Volume 2: Case Studies on System Leadership*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264039551-3-en> pp. 21-36.
- Janssens, F. and G. van Amelsvoort (2008), “School self-evaluations and school inspections in Europe: An exploratory study”, *Studies in Educational Evaluation*, No. 34, Elsevier, pp. 15-23.
- Janssens, Frans J. G. and I. de Wolf (2009), “Analyzing the Assumptions of a Policy Program An Ex-ante Evaluation of ‘Educational Governance’ in the Netherlands”, *American Journal of Evaluation*, Sage Publications, Vol. 30(3), pp. 411-425.
- Janssens, F.J.G., L. Rekers-Mombarg and E. Lacor (2014), *Leerwinst en toegevoegde waarde in het primair onderwijs: Eindrapportage*, available at: <http://www.rijksoverheid.nl/ministeries/ocw/documenten-en-publicaties/kamerstukken/2014/03/07/aanbieding-eindrapport-pilot-leerwinst-en-toegevoegde-waarde-primair-onderwijs.html>.
- Livingstone, K. and J. McCall (2005), “Evaluation: Judgmental or Developmental?”, *The European Journal of Teacher Education*, Routledge, London, Vol. 28, No. 2.

- Lockwood, J.R. and McCaffrey, D. (2013), *Correcting for test score measurement error in ANCOVA models for estimating treatment effects*, Journal of Educational and Behavioral Statistics, Sage Publications, vol. 39 no. 1 22-52.
- Luginbuhl, Rob, D. Webbink and I. de Wolf (2009), “Do Inspections Improve Primary School Performance?”, *Educational Evaluation and Policy Analysis*, Sage Journals, Vol. 31(3), pp. 221-237
- MacBeath (n.d.), *School self-evaluation: What we are learning from other countries*. <https://www.teachers.org.uk/files/word/MacBeath-nutcountries.doc>
- Morris, A. (2011), “Student standardised testing: Current practices in OECD countries and a literature review”, *OECD Education Working Papers*, No. 65, OECD Publishing, Paris, www.oecd.org/edu/workingpapers.
- National Association of Secondary School Principals (n.d.). *Rank Foolishness*, accessed October 11, 2013 from www.principals.org/Content.aspx?topic=56684.
- Newton, X., et al. (2010), “Value-added modeling of teacher effectiveness: An exploration of stability across models and contexts” in *Education Policy Analysis Archives*, EPAA, 18(23), <http://epaa.asu.edu/ojs/article/view/810>.
- NRC (National Research Council) (2010), *Getting value out of value-added: Report of a workshop*, H.I. Braun, N. Chudowsky, & J. Koenig, (eds.), The National Academies Press, Washington DC, <http://216.78.200.159/Documents/RandD/Other/Getting%20Value%20out%20of%20Value-Added.pdf>.
- Nusche, D., et al. (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116917-en>.
- Nusche, D., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Norway 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264117006-en>.
- OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264190658-en>.
- OECD (2008), *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-Added of Schools*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264050259-en>.
- Pont, B., D. Nusche and H. Moorman (2008), *Improving School Leadership, Volume 1: Policy and Practice*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264044715-en>.
- Ray, A. (2006). *School Value Added Measures in England: A Background Report for the OECD Project on the Development of Value-Added Models in Education Systems*, www.dcsf.gov.uk/rsgateway/DB/RRP/u015013/index.shtml.
- Raudenbush, S. (2004), *Schooling, statistics, and poverty: Can we measure school improvement?* The Ninth Annual William H. Angoff Memorial Lecture, Educational Testing Service, Princeton, New Jersey, www.ets.org/Media/Education_Topics/pdf/angoff9.pdf.

- Reardon, S. & Raudenbush, S. (2009), *Assumptions of value-added models for estimating school effects*, *Education Finance and Policy*, MIT Press Journals, 4(4), 492-519.
- Rosenkvist, M.A. (2010), “Using student test results for accountability and improvement: A literature review”, *OECD Education Working Papers*, No. 54, OECD Publishing, Paris, oecd.org/edu/workingpapers.
- Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, *OECD Reviews of Evaluation and Assessment in Education*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.
- Scheerens, J., et al. (2005), *Positioning and validating the supervision framework*, University of Twente, Netherlands.
- Scheerens, J., et al. (2012), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands*, University of Twente, Netherlands, www.oecd.org/edu/evaluationpolicy.
- Scheerens, J. (2013), Addendum to *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands*, University of Twente, Netherlands.
- Shewbridge, C, et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: School Evaluation in the Flemish Community of Belgium 2011*, *OECD Reviews of Evaluation and Assessment in Education*, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116726-en>.
- Schildkamp, Kim (2007), *The utilisation of a self-evaluation instrument for primary education*, Thesis, University of Twente, Netherlands, <http://doc.utwente.nl/57803>.
- Schildkamp, Kim and M. Ehren (2012), “Intuition” to “Data”-based Decision Making in Dutch Secondary Schools’ in Schildkamp, K., Lai, M.K., & Earl (eds.), *Data-based Decision Making in Education: Challenges and Opportunities*, Springer Publishing, Dordrecht, pp. 49-68.
- SICI (2003), *Effective school self-evaluation (ESSE) project*, SICI, Brussels, www.edubcn.cat/rcs_gene/extra/05_pla_de_formacio/direccions/primaria/bloc1/1_ava luacio/plugin-essereport.pdf.
- SICI (2012), *The Inspectorate of Education in the Netherlands: SICI inspectorate profiles*, SICI, Brussels, www.sici-inspectorates.eu/Members/Inspection-Profiles/The-Netherlands.
- Timmermans, A. (2012), *Value-added in educational accountability: Possible, fair and useful?* Ph.d. thesis, University of Groningen, Netherlands.
- Twist, Mark van, M. van der Steen and M. Kleiboer (2011), *Coping With Very Weak Primary Schools: Towards ‘Smart Intervention’ in Dutch Education Policy*, Netherlands School of Public Administration (NSOB), Paper presented at the OECD conference: “Effective Governance From the Centre”, First GCES Thematic Conference, The Hague, Netherlands, 21-22 November 2011.
- Visscher, A.J. and R. Coe (2003), “School performance feedback systems: Conceptualisation, analysis and reflection”, *School Effectiveness and School Improvement: An International Journal of Research, Policy and Practice*, Vol. 14:3, pp. 321-349, Routledge.

- Visscher, A.J. and M. Ehren (2011), *De eenvoud en complexiteit van opbrengstgericht werken*, Vakgroep Onderwijsorganisatie en –management, Universiteit Twente, Enschede, Netherlands.
- Wise, S.L. (2014), “The utility of adaptive testing in addressing the problem of unmotivated examinees”, *Journal of Computerized Adaptive Testing*, 2(1).
- Wright, P., et al. (2010), *SAS® EVAAS® statistical models*, www.sas.com/resources/asset/SAS-EVAAS-Statistical-Models.pdf.
- Zitter, I. and A. Hoeve (2012), “Hybrid Learning Environments: Merging Learning and Work Processes to Facilitate Knowledge Integration and Transitions”, *OECD Education Working Papers*, No. 81, OECD Publishing, Paris, <http://dx.doi.org/10.1787/5k97785xwdvf-en>.

Chapter 6

Education system evaluation

The Dutch approach to evaluation and assessment gives a strategic role to education system evaluation as part of policy planning and development. This is reflected in the establishment of comprehensive information systems, sample-based national assessments, longitudinal surveys and programme evaluations. It is also visible in the transparent reporting of indicators and education outcome measures in online databases and annual publications on the state of the education system. There has been increased attention to disseminating information widely and in different formats and there is wide acceptance of the principle that education policies and programmes should be continuously monitored in order to inform future policy development and educational planning. In addition, education system evaluation could benefit from a clearer and more comprehensive set of educational goals and priorities, which could contribute to reminding all stakeholders of the full spectrum of national priorities, while also communicating that not all of these are currently measured or measurable. Going further, consideration should be given to monitoring a broader range of valued student competencies, devoting more attention to the analysis of learning outcomes across student groups, and incorporating information on school professionals, practices and perceptions into the system evaluation framework.

This chapter looks at system evaluation within the Dutch evaluation and assessment framework. System evaluation refers to approaches to monitor and evaluate the performance of the education system as a whole. The main aims of system evaluation are to provide accountability information to the public and to improve educational policies and outcomes.

Context and features

Responsibilities for evaluation of the Dutch education system

The responsibility for the evaluation of the Dutch education system is essentially shared between the Ministry of Education, Culture and Science and the Inspectorate of Education. The Ministry's main functions in education system evaluation are: the development of tools to monitor the performance of the education system (e.g. indicators framework, national student assessments, cohort studies), the promotion of evaluative studies of particular aspects of the education system (e.g. policy and programme evaluation), and the use of evaluation results for decision-making and policy development.

The Inspectorate of Education assumes the major responsibility for monitoring the quality of education. The Dutch Constitution entrusts the Inspectorate of Education with the preparation of an annual report on the State of Education in the Netherlands. Overall, the Inspectorate is responsible for reporting publicly on the education system as a whole, providing information for policy development, and supplying reliable information on education. In consultation with the Ministry of Education, Culture and Science, it also engages in policy evaluations, and contracts research and analysis on specific aspects of the education system.

The Ministry of Education, Culture and Science and the Inspectorate of Education are supported by a range of organisations in the planning and execution of system evaluation procedures, and the associated analysis. For instance, the Central Institute for Test Development (Cito) is responsible for the periodic assessment project (PPON) and the Cohort studies are undertaken by a consortium bringing together educational research institutes, Cito and the National Bureau of Statistics (CBS). Typically, the Inspectorate of Education, research institutions, Cito and organisations such as the National Bureau of Statistics and the Netherlands Bureau for Economic Policy Analysis (*Centraal Planbureau, CPB*¹) take responsibility together for empirical studies. The planning of research and evaluative studies by the Ministry of Education, Culture and Science, conducted within its Knowledge Directorate generally involves external advice given by organisations such as the Organisation for Scientific Research (NWO) (e.g. monitoring studies, Cohort studies).

In addition, the Education Council (*Onderwijsraad*), an independent advisory body funded by the government and composed of ten experts, provides independent advice to the government on national education issues. These are typically at the request of the Minister of Education or the Parliament (e.g. review of draft laws). The Education Council has also the right to initiate reviews, investigate specific themes and issue its positions. Recent themes addressed by the Education Council include value-added measurement, compulsory testing at the end of primary education, the governance of school boards, the teacher career and small schools. The Education Council is scientific in nature, it does not represent a particular stakeholder group. It seeks to provide a strong empirical basis to its advice.

The Netherlands has a well-developed structure of institutes specialised in educational research. Some are connected to universities and some are private companies. Typically a number of research institutes compete for obtaining a contract to carry out specific evaluation projects in the educational sector. For example, the “Monitors” promoted by the Ministry of Education, Culture and Science (small-scale studies about specific education policy issues) are carried out by educational research institutes. In 2010, the National Education Plan (*Nationaal Plan Onderwijs/Leerwetenschappen*)² laid out a strategy to reorganise education research to make it more responsive to the needs of policy-makers and practitioners and to facilitate dissemination. In line with these plans, a new national coordinating body, the Netherlands Initiative for Education Research (*Nationaal Regieorgaan Onderwijsonderzoek, NRO*) was created within the NWO in 2013 in order to allocate funding for all types of education research including fundamental, policy-oriented and practice-oriented research³. The role of both the Primary Education Council (*PO-Raad*) and the Secondary Education Council (*VO-Raad*) in education system evaluation consists mostly of the provision of information at the school level, in particular through the development of the school information system “Windows for Accountability” (see Chapter 5).

Further monitoring of the education system is provided from outside the education sector by the Netherlands Court of Audit (*Algemene Rekenkamer*), which conducts performance audits of public spending. Recent topics for performance audits in education include “needs-based” education, special needs in primary and secondary education and practical training in vocational secondary education.

Goals for education system evaluation

Objectives and goals for the Dutch education system are not explicitly stated in a single official document but are to be inferred from current educational policies and programmes (Scheerens et al., 2012). The definition of student learning objectives at the central level, with due account of the high levels of autonomy and the “freedom of education” of Dutch schools, is elaborated in broad core learning objectives for primary and lower secondary education outlined in the Act on Primary Education and the Act on Secondary Education. For the upper cycle of secondary education, learning targets have been translated into centrally set examination programmes (Chapter 1). More recently, through the 2010 “Language and Numeracy Act”, the Ministry of Education, Culture and Science developed additional reference levels for literacy and numeracy (*referentieniveaus*). These function as “benchmarks” or attainment targets for the knowledge and skills that students are expected to acquire in literacy and numeracy at the end of primary education, and at the end of each of the educational tracks in secondary education⁴ (indicating a “fundamental” and an “advanced” level).

The Ministry of Education, Culture and Science is increasingly using Action Plans to express policy agendas. These typically contain specific lines of action and explicit attainment targets, which can be used as further references for education system evaluation. Recent Action Plans include: the Action Plan to improve primary education (“Basis for Performing”), the Action Plan to improve secondary education (“Better Performance”), the Action Plan to address drop out (“Action Plan Youth”) and the Action Plan to improve the quality of teachers (“Teaching 2020”). Examples of explicit attainment targets to be achieved at a particular point in time (e.g. 2015 and/or 2018) include: specific PISA scores or end-of-primary-education Cito test scores; a given percentage of schools engaged in results-oriented work; a specific decrease of the number of underperforming schools; and achieving the membership of all teachers in the Register

for teachers by 2018. The main attainment targets for each of the two main Action Plans are described in Table 6.1 below.

Table 6.1 Main attainment targets for specific Action Plans

Action Plan	Main attainment targets
Action Plan to Improve Primary Education (“Basis for Performing”)	<p>Better performance</p> <ul style="list-style-type: none"> The average score in the end-of-primary-education Cito test should increase to 537 by 2015 (from 535.4 in 2010) The score corresponding to the 20th percentile (top performers) in the end-of-primary-education Cito test should increase to 545 by 2015 (from a level between 543 and 544 in 2010) <p>Reduce number of underperforming schools</p> <ul style="list-style-type: none"> The number of primary schools considered as “unsatisfactory” is reduced in half by 2015 (from 57 to 35) The number of primary schools considered as “weak” is reduced in half by 2015 (from 238 to 119). <p>Generalise results-oriented work</p> <ul style="list-style-type: none"> Ensure that by 2015 60% of primary schools engage in results-oriented work and that by 2018 90% of them do so.
Action Plan to Improve Secondary Education (“Better Performance”)	<p>Better performance</p> <ul style="list-style-type: none"> PISA mathematics score to increase to 536 in 2015 and 541 in 2018 (from 526 in 2009) PISA reading score to increase to 516 in 2015 and 520 in 2018 (from 508 in 2009) PISA science score to increase to 526 in 2015 and 528 in 2018 (from 522 in 2009) <p>Greater focus on top performers and gifted students</p> <ul style="list-style-type: none"> Increase the performance of the 20% top students <p>Generalise results-oriented work</p> <ul style="list-style-type: none"> Ensure that by 2015 at least half of secondary schools engage in results-oriented work and that by 2018 90% of them do so.

Sources: Dutch Ministry of Education, Culture and Science (2011a), *Actieplan po Basis voor presteren*, www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/05/23/actieplan-po-basis-voor-presteren.html

Dutch Ministry of Education, Culture and Science (2011b), *Actieplan vo Beter presteren*, www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/05/23/actieplan-vo-beter-presteren.html

Major tools to monitor performance of the education system

National assessments of student performance

The Netherlands collects a range of data on student performance to monitor learning outcomes at the system level. Information on student learning outcomes is collected from monitoring sample surveys, standardised student assessment and national examinations at the end of secondary education.

The Periodical Survey of Education (PPON) is a sample-based assessment (sample of schools) that has been administered periodically in different disciplines since 1987, typically to Year 8 students and, in some cases, Year 4 students. It monitors skills in Dutch and mathematics on a five-year cycle. Other curriculum areas that are monitored as part of the PPON include: world orientation (social science), history, geography, biology, English, visual arts, music and physical education (Cito, 2008). A school questionnaire is also administered to assess subject-matter coverage of the school curriculum. The design of the PPON aims to provide robust measures of changes over time, covering large amounts of the curriculum.

A new monitoring survey, the Annual Survey of Educational Levels (JPON), was introduced in 2008 to specifically monitor progress on the roll out of the Ministry for Education, Culture and Science’s quality agenda “Schools for Tomorrow”, and monitors student mastery of Dutch language and mathematics at two points in primary education (Years 4 and 8). The design of JPON aims to provide more regular and timely feedback on a narrower area corresponding to the national reform agenda in primary education.

Both the PPON and JPON are implemented by Cito and use matrix sampling (dividing items into different versions of the test) to cover a broad range of items without overburdening individual students. The surveys use Item Response Theory and therefore allow reporting of what students can or cannot typically do against defined performance standards. Both the PPON and JPON identify and report trends in student achievement and provide information for policy makers, educational practitioners and the general public.

As mentioned in Chapter 3, education system evaluation also relies on the assessments that were initially designed for individual student assessment. For example, the central authorities are setting national targets for performance based on the Cito end-of-primary test. This is a high-stakes test administered by about 85% of schools, primarily to advise students on the appropriate secondary education track for them. The Cito test is used by the Inspectorate of Education to compare performance across schools and its aggregate results are often also used to evaluate the impact of a specific policy or programme. As of 2014, the administration of an end-of-primary standardised test aligned to the reference levels will become mandatory, but schools will be free to choose from among different tests that comply with centrally set quality criteria. Results from the secondary school leaving examinations are also aggregated to the school, regional and national level for accountability and monitoring purposes.

As of 2014, primary schools will also be required to implement a student monitoring system that consists of formative standardised assessment, such as the LVS (*Leerling Volg Systeem*, see Chapter 3). Schools will be free to select from among different monitoring systems that comply with centrally set quality criteria. Schools will be required to use these systems to quantitatively assess the progress of their students towards the reference levels. Following the implementation of the unique student identifier, student progress can now be analysed as results in LVS tests, Cito test and national examinations will be registered for each student. In particular, scores in LVS tests are vertically equated, which allows the calculation of student growth trajectories in primary education. While there are no intentions to use the LVS results for accountability, the student number can facilitate the longitudinal analysis of student assessment results.

Longitudinal surveys

The Netherlands has a long tradition of longitudinal research programmes. Initially, two different cohort studies in primary education (PRIMA) and secondary education (VOCL) were developed. Essentially, both comprised: achievement testing of students in language and mathematics/arithmetic at various stages during the school career; psychological testing; the recording of data on students’ progress throughout the school programme (class repetition, drop out, transfer to another school type, and examination results); and the collection of school background data by means of questionnaires to parents, teachers and school leaders.

From 2007, the primary and secondary cohort studies were integrated into the Cohort Survey School Careers (*Cohort Onderzoek Onderwijsloopbanen*, COOL). The COOL

study comprises data collection in primary (Years 2, 5 and 8), general secondary (third year) and vocational schools (second year). COOL surveys have been planned until 2015. Student achievement is tested in language and mathematics/arithmetic. The achievement tests make use of the Cito pupil monitoring system LVS. In addition, short questionnaires are administered to students.

COOL is jointly financed by the Ministry of Education, Culture and Science and the Foundation for Scientific Research (NWO). It is undertaken by a consortium that brings together educational research institutes, Cito and the National Bureau of Statistics (CBS). It is a data source for analysis by a range of groups such as the Ministry of Education, Culture and Science, the Education Council and research groups. The COOL study also comprises specific services to feedback information to schools⁵.

System-level indicators

For the purpose of system-level monitoring, student performance data are complemented by a wide range of demographic, administrative and contextual data. The Ministry of Education, Culture and Science, through its unit dedicated to information management, DUO (*Dienst Uitvoering Onderwijs*, Implementation of Education Service), collects statistical snap-shot data from public and private schools. The data collections are conducted in co-operation with the schools. Schools are requested to periodically send their data in a standardised format to DUO.

The dataset includes information on students (e.g. type of enrolment, age, gender, completion, transition, transfers across school types, level of parental education [for primary education only], end-of-primary test results and school recommendation [for secondary education only]), teachers (e.g. functions, qualifications, age, gender, areas taught, teaching hours, absenteeism), non-teaching staff (qualifications, age, gender, category), schools (e.g. location, financial data, special needs advice) and school boards (e.g. number of schools, financial data). DUO also centralises data on student performance, namely results on national examinations and the Cito end-of-primary-education test. The Ministry has developed international indicators on the education system, as part of the joint UNESCO-OECD-EUROSTAT annual data collection on enrolment, graduation, finance and personnel.

The Ministry of Education, Culture and Science developed a sophisticated tool to monitor school drop-out figures for various target groups. The *School Drop-out Explorer* is an interactive tool that makes available quantitative and qualitative information on school drop-outs at the national, regional, local and school level. The Explorer allows to compare drop-out figures between schools and regions and to observe developments over time, which makes it possible to identify regional and other trends. The tool also includes other relevant information such as examples of promising projects to reduce school drop-out.

The National Bureau of Statistics (CBS) collects information on school finances. In addition, its databases allow for the education data developed by DUO to be linked to information about parents (e.g. labour market, population registry). Its social statistics database (with information on health, labour market, pensions and crime) is used by the Ministry, the Inspectorate of Education and independent research institutes. There is also collaboration between DUO and CBS involving exchange of data and agreement on definitions.

Reporting on the education system

The Ministry of Education, Culture and Science brings together the basic education statistics described above and publishes reports with system level indicators on education. The annual “Key Figures” report publishes key data and statistics on education, culture and sciences over a period of five years. It is mostly descriptive and covers areas such as student enrolment, transition and completion rates, transfers across school types, expenditures, financial situation of schools, school staff, and school size. Each edition also addresses special themes such as early school leavers, ethnic minorities, special needs advisory teams, selection of subject clusters in secondary education, non-subsidised education, top performers, and differences in performance between boys and girls.

Alongside “Key Figures”, the Ministry publishes “Trends in Pictures” (*Trends in Beeld*), which presents trends in the form of figures, tables or maps. Trends are featured by theme such as participation in education, links between education and the labour market, profile of the teaching profession, coverage, educational performance, lifelong learning, teacher education and instruction time. An online version of “Trends in Pictures” is also available. The National Bureau of Statistics (CBS) publishes a statistical compendium on education, the “Yearbook – Education in Figures”. Its 2012 edition included chapters on key figures by type of education, flows within the education system, studying beyond primary school, school leavers, transition to the labour market and expenditure on education. This is in addition to occasional publications on topics such as trends in educational expenditure and spending on vocational training. The Ministry of Education, Culture and Science also publishes the *Drop-out Atlas*, an annual publication bringing together yearly statistics on school drop-out rates at the national, regional and school level. It also provides background information and explores potential strategies to tackle school drop-out at the different levels of the system.

Regularly updated statistics and indicators on education are available on line at the *Open Education Data* web portal (www.data.duo.nl). It has been developed as a “one-stop-shop” for information on education statistics. The web portal is open to all audiences and brings together a wide range of information, including demographic and contextual information, and statistical information on students, qualifications, schools, personnel and school funding. Similarly, the National Bureau of Statistics offers a large section on education in its public data portal (<http://statline.cbs.nl>). It provides the general public with a wealth of information in areas such as participation, completion, flows across school types, the educational workforce, special needs education, expenditure on education, funding and lifelong learning.

Since 1801, the Inspectorate of Education has published an annual report on the state of education in the Netherlands. The *State of Education* report contains overall findings on the current state of affairs in education as identified during specific inspection activities (so-called “Education Report inspections”) and is complemented with data from other sources, such as examination results, school career data (based on the cohort studies), information from PPO and educational research studies. The report describes trends in primary and secondary education, including facts identified through the inspection framework (e.g. on the quality of teaching or quality assurance procedures in schools). It identifies challenges in education policy and provides recommendations. The report further analyses the performance of the primary and secondary education sectors. A number of specific themes are discussed such as the quality of school governance, the financial management of schools, teacher quality and the quality of education for students with special needs.

In the Netherlands surveys of students and parents are organised at the school level or by specific central organisations. For example, the National Students' Organisation (*Landelijk Aktie Komitee Scholieren, LAKS*) organises a national student survey regarding the quality of education and career support for students. Parents' satisfaction with schools is surveyed as part of the National Schools Survey conducted by the Center for Online Research (*Centrum voor Online Onderzoek*), and in the early 2000s, the Social and Cultural Planning Bureau (*Sociaal en Cultureel Planbureau, SCP*), funded by the Ministry of Education, Culture and Science, conducted a survey and research on parents' attitudes towards education (SCP, 2004). Surveys of teachers and school leaders are also occasionally organised to inform specific policy initiatives (e.g. research promoted by the Ministry to inform the Action Plan "Teacher 2020", which included a survey of teachers and school leaders).

Thematic evaluations and programme evaluation

In addition to its annual *State of Education* report, the Inspectorate publishes thematic reports, reflecting its own research, on issues such as teaching time, dyslexia and dyscalculia, and school governance. The Inspectorate of Education has also been involved in co-ordinating and partially undertaking two major policy evaluations of primary and lower secondary education (Scheerens et al., 2012).

In the Netherlands, programme evaluations have been undertaken in various forms and shapes during the last four decades. These evaluations range from fairly big and well documented projects (e.g. Parliamentary Committee Educational Innovation, which investigated innovations in secondary education) to a large set of smaller studies (Scheerens et al., 2012). Programme evaluations have been typically promoted by evaluation committees of the Ministry of Education, Culture and Science, which contracts out research studies to a variety of research groups. These are published as, for instance, "Monitors" by the Ministry.

Examples of issues addressed by recent studies include: early school leavers, class size, quality of the school infrastructure, school bullying, the implementation of the unique student identifier, the implementation of "needs-based" education, the introduction of the reference levels, and the effectiveness of remuneration measures, which are part of the Action Plan for Teachers. The studies also include evaluation of pilot projects, such as an experiment with flexible teaching time, and an organisation of student learning time within five school days of equal length. The Ministry also undertakes prospective studies, such as forecasts of teacher needs and estimations of the number of students there will be in one or two decades.

Participation in international student surveys

The Netherlands attributes a great deal of importance to international benchmarks of student performance, and has participated in most major international studies that provide trend data on outcomes at different stages of education in the Netherlands. It has participated in the triennial OECD Programme for International Student Assessment (PISA) since its inception in 2000 (although, in 2000 the required sampling standards for international comparison were not achieved in the Netherlands.) This programme tests 15-year-old students' knowledge and skills in reading, mathematics and science at the end of lower secondary education.

The Netherlands has also administered tests to students in Grades 4 and 8 as part of the International Association for Educational Achievement's (IEA) Trends in

Mathematics and Science Skills (TIMSS) studies in 1995, 1999, 2003, 2007 and 2011. The IEA's Progress in Reading Literacy Skills (PIRLS) study also provides an international benchmark for Grade 4 students' reading literacy over time, with the participation of the Netherlands in 2001, 2006 and 2011. As such, the Netherlands has a wealth of information on students' core skills in reading, mathematics and science at major points in school education to compare the system internationally. It provides benchmark information on the education system's performance and also allows monitoring of progress over time, for example via the trend data available from PISA.

Furthermore, the Netherlands participated in the 2013 IEA International Computer and Information Literacy Study (ICILS). The Netherlands also supports international comparisons on non-cognitive outcomes, including their participation in the 2009 IEA International Civic and Citizenship Education Study (ICCS). In addition, the Netherlands participated in 2011 in the European Commission's Survey on Language Competences, which tested students in their final year of lower secondary education in the two most widely taught foreign European languages (English and German).

The Netherlands regularly participates in the OECD's Teaching and Learning International Survey (TALIS) (2008 and 2013 rounds), which provides information on teachers' perceptions of various aspects of the school environment, including their profession. In the first round of TALIS (2008), however, the required sampling standards were not achieved in the Netherlands and its data was not included in the international comparisons.

Strengths

System evaluation receives considerable attention in the Netherlands

The monitoring of the education system is a well-developed component of the Dutch approach to evaluation and assessment. There has been considerable attention among policy makers and other education stakeholders on the development of instruments to assess the quality and progress of educational outcomes. This involves the collaboration of a wide range of players, such as the Ministry, the Inspectorate, representatives of school boards and research institutes. Education policy gives a strategic role to system evaluation as an essential part of policy planning and development.

This key focus on system evaluation is reflected in the establishment of comprehensive information systems, sample-based national assessments, longitudinal surveys and programme evaluations. It is also visible in the transparent reporting of indicators and education outcome measures in online databases and annual publications on the state of the education system. There has been increased attention to publishing and disseminating information widely and in different formats among educational agencies, schools, families and the general public. There is also wide acceptance of the principle that education policies and programmes should be continuously monitored and evaluated in order to inform future policy development and educational planning.

The approach to system evaluation is comprehensive

System evaluation in the Netherlands is achieved in several ways. First, information on student learning outcomes is collected in a variety of ways. Not only are student results from standardised assessments available at a given point in time (permitting analysis of differences across student groups or across schools), but analyses over time are also feasible through longitudinal surveys.

Second, demographic, administrative and contextual data are collected, which are useful to explain the outcomes of the education system. This is part of the education indicators framework managed by DUO. It is also noteworthy that student and parental views about well-being, satisfaction, engagement and motivation are collected at the school level. If these important stakeholder views were to be collected in a more systematic and standardised manner across schools, these data could be of greater policy and research relevance, through the analysis of the association between student performance and qualitative aspects of school life.

Third, system evaluation information goes well beyond quantitative measures of student performance data and education indicators. These are essential but only represent a subset of student learning objectives and, in particular, do not provide the richness of contextualised qualitative analysis. The qualitative analysis provided by the Inspectorate on the basis of individual school evaluations, as well as the secondary analyses of student performance and contextual data undertaken by the Ministry or research groups, provide evaluative information that broadens the base of evidence to inform policy development. Policy development is aided by the empirical research studies sponsored by the Education Council, which are typically conducted by researchers with long-standing expertise in the field.

Finally, there is also considerable openness to external views. For example, the Netherlands participates in international reviews of educational policy, such as the OECD projects on *Equity and Quality in Education*, *Skills Beyond School*, *Governing Complex Education Systems* and the project that this report belongs to: *Evaluation and Assessment Frameworks for Improving School Outcomes*. It has hosted OECD Reviews of adult education, career guidance, early childhood education and care, migrant education and tertiary education. This is in addition to its participation in a range of international surveys to benchmark the performance of the education system.

There is strong capacity to undertake system evaluation

The capacity for system evaluation is significant. Thousands of student standardised assessments and national examinations are processed every year that require a large logistical capacity and high levels of technical expertise. In the Netherlands, institutions such as Cito provide technical expertise, agencies such as CVE (*College Voor Examens*) provide guidance, and there is good implementation capacity within the Ministry. Work on national student assessments involves a large number of individuals, including top academics, distinguished teachers and experienced school leaders. The development of instruments and the processing of the data generally involves the top experts in the country. Areas such as educational measurement, psychometrics, test development, validation of test items or scaling methods are well developed in the Netherlands.

In addition, education system evaluation in the Netherlands benefits from a large and respected education research community. This is instrumental in the ability to respond to the Ministry's calls for programme and policy evaluations, as these are conducted in close co-operation with research groups based in institutions of higher education or in companies. The development of education indicators is undertaken within a specialised unit within the Ministry (DUO), which is provided with the needed technical expertise.

Other bodies playing a role in system evaluation boast strong capacity. The Inspectorate of Education has long-standing expertise in monitoring education at a system level through its thematic inspections and Education Report inspections. The latter are conducted specifically to collect information on a range of system indicators that feed

into the annual State of Education report (Dutch Inspectorate of Education, 2010). The Education Council has ten distinguished academics and education specialists as members, and the National Court of Audit is a large institution with a considerable proportion of its staff working on performance audits.

There is a great deal of co-operation between the Ministry, the Inspectorate, the National Bureau of Statistics (CBS), representatives of school boards and individual schools in the dissemination among school agents of evaluation information. This has been facilitated with the recent development of the online information system *Windows for Accountability* (see Chapter 5) and the maintenance of education data portals by DUO and CBS. The monitoring of school drop-out rates through the *School Drop-out Explorer* offers another example of advanced multilevel monitoring activities. It shows the strengths of the Dutch evaluation approach in connecting national, regional, local and school-level data and analysing these in a multilevel perspective.

System evaluation is undertaken with a good degree of independence

A further strength of system evaluation in the Netherlands is the substantial degree of independence that those with responsibilities have to make judgments on whether or not education system objectives are met. This is in part the result of the autonomy of the Inspectorate of Education, whose distance from political decision-making allows it to conduct rigorous and reliable inspections, analyse the relevant data, and be impartial in its conclusions about the education system (in particular, as outlined in its annual publication about the state of education in the Netherlands). Hence, it can provide a fresh, apolitical and constructive point of view informing the national debate.

The Inspectorate of Education is a key player in system evaluation in the Netherlands with the main responsibility for analysing the state of education in the country. Its annual report provides valuable information on a system-wide basis about specific aspects of schools' work that can inform policy makers. Information is reported over several years, so trends over time can be examined. It summarises the state of all quality indicators in the different types of schools and sheds light on identified priority areas within the school system, which can form a basis for the development or refinement of policy to address these priorities. The Inspectorate of Education also publishes specific reports on the results of thematic inspections conducted in a sample of schools throughout the system (e.g. teaching approaches in schools, partnerships between schools in the context of "needs-based" education, teaching time, addressing dyslexia and dyscalculia). Overall, the Inspectorate provides a unique evaluative contribution as it uses classroom- and school-level evidence based on its own framework, standards and indicators.

In addition, the variety of those involved with system evaluation that are external to educational governance greatly contribute to independent judgments. For example, sample-based student assessments (PPON, JPON) and the Cohort studies are carried out and analysed independently of the Ministry (e.g., by Cito). The Ministry also commissions a variety of policy and programme evaluations from independent research groups. In addition the Education Council provides advice based on the existing empirical evidence and offers its perspective on the state of education in the Netherlands in periodic reports (with its third edition published in 2013).

There is a good ability to monitor student learning outcomes at the national level

The Netherlands collects a wide range of data on student learning outcomes in relation to core learning objectives and reference levels. Information on student learning outcomes is collected from sample-based assessments (PPON, JPON) in primary education, longitudinal cohort studies, end-of-primary-education Cito tests, national school-leaving examinations and international student assessments. This shows recognition of the need for regular monitoring of key student outcomes in the education system to inform policy making.

The secondary school-leaving examinations are the only type of student assessment that targets the totality of the student cohort. The end-of-primary-education Cito tests, while not a mandatory student assessment covering an entire student cohort, is used by about 85% of schools to advise students on which track to take in secondary education. While the major purpose of both these assessments is to provide a summative measure of individual student performance, the annual results provide information on average performance for the participating schools and, as point in time measures, they can offer insight into average quality in measured areas, as well as the equity of outcomes.

The results from sample-based student assessments (PPON and JPON) give a comprehensive national picture of student performance in primary education in a broad range of subjects. One of the benefits of PPON is its broad coverage that ensures most curriculum areas are assessed. PPON is conducted every year, but assesses a different set of disciplines each year. A number of trend tasks are kept constant over the assessment cycles in order to obtain longitudinal data, allowing the analysis of time trends for each curriculum area. Both PPON and JPON are increasingly aligned with the reference levels, and, consequently, they enable the measurement of progress towards national education goals, thus broadening the national debate beyond results in international surveys. Both PPON and JPON gather information on the characteristics of students and schools through context questionnaires. The collection of this context information facilitates the identification of the social and school factors that are potentially associated with achievement. Longitudinal cohort studies allow for student progress to be analysed in light of contextual and school-level data, providing a valuable instrument to explain student performance.

Finally, the mandatory introduction of student monitoring systems such as the LVS (*Leerling Volg Systeem*) in all schools, together with the establishment of a unique student identifier, enables longitudinal analysis of student performance and is a key strength of system monitoring in the Netherlands. The unique identifier allows authorised users to monitor and ensure student enrolment and attendance, ensure that students' educational records are accurately maintained and issue the documentation that students need to present to other schools or education providers. The identifier also facilitates the analysis of student progress in the school system and data-sharing between schools, and can be used for reporting purposes by education agencies.

An Education Indicators Framework is established

An Education Indicators Framework is in place to help decision makers analyse the state of the education system, monitor trends over time, and provide information to the general public. The framework includes five core components (students, teachers, non-teaching staff, schools, school boards) and covers, as main areas, student enrolment and progression, teacher qualifications and working conditions, and school resources. It also

includes data on current priorities, such as early school leavers, special needs education and transitions within the education system and to the labour market. It involves well-established procedures for data collection in close cooperation with schools. A major advantage in the collection of data from individual schools is that data are provided at the individual student level. This creates the opportunity to obtain information on the background of individual students, and to link this data to their achievement at school. In addition, efforts are made to continuously monitor the teacher labour market, the financial circumstances of schools and transitions to the labour market.

Education statistics are widely disseminated. The Indicators Framework is the basis for statistical reporting on the education system used in publications such as *Key Figures* and *Trends in Pictures*, and the education data portals organised by DUO and the National Bureau of Statistics (CBS). The management of the Framework by DUO, as well as the collaboration with CBS, also ensures the monitoring of statistical development needs at all levels of schooling, and prioritises areas for improvement.

The principle of transparency in monitoring and publishing results is well established

In the Netherlands, the principle of transparency in monitoring and publishing results is well established. The Netherlands collects a wide range of data on education system performance, including through national student assessments, qualitative studies and inspection reports. The Ministry of Education, Culture and Science, as well as the National Bureau of Statistics, publishes a comprehensive set of educational statistics and reports its analysis in annual publications such as “Key Figures” and “Trends in Pictures”. The Inspectorate also publishes reports from individual external school evaluations alongside thematic studies on particular aspects of the education system. In addition, the Inspectorate publishes its annual report on the state of education in the Netherlands on the basis of its inspections. Results of school-leaving national examinations are published at the school level and analysed in reports produced by DUO and CVE (*College Voor Examens*). Similarly, a large set of indicators on individual schools is available from “Windows for Accountability”.

In addition, the Ministry publishes a series of analytical reports that cover trends over the last few years, and reports on themes of current policy relevance. These analytical reports draw on results from the monitoring system, plus other reports or programme evaluations that have been conducted. Analyses of national and international student assessments are also published, as well as the analytical reports of the longitudinal cohort studies.

There is a commitment to developing a strong evidence base in education

The principle of informing policies with evidence from research is well established in the Netherlands. The Ministry of Education, Culture and Science commissions a variety of research studies and promotes the evaluation of the implementation of specific policies and programmes. The aim of such studies is to determine how well education programmes are being implemented, and to enable the Ministry to draw lessons from such experience and either refine particular measures or better design future initiatives. These studies lead to a significant collaboration of the education authorities with educational researchers and companies, such as: the National Bureau of Statistics (CBS), the CPB Netherlands Bureau for Economic Policy Analysis, the Research Centre for Education and the Labour Market (ROA), ITS/Radboud Universiteit Nijmegen, the

Groningen Institute for Educational Research, Regioplan Policy Research, and Panteia. Informing policy on the basis of evidence is also at the heart of the work of the Education Council and the Foundation for Scientific Research (NWO).

A clear indication of the importance of evidence-based policy is that the Ministry of Education, Culture and Science has established a “Knowledge Directorate” to support an evidence-based approach to policy development. It works to ensure a strategic approach to the Ministry’s research, analysis and evaluation, and information management activities in view of supporting the provision of evidence-based policy advice to the Minister. The Knowledge Directorate keeps abreast of trends and developments shaping education (e.g. demographic, technological, economic) and promotes analyses to ensure education meets the needs of the society (e.g. improving its alignment with labour markets). It also co-ordinates the development of evidence by the Ministry to inform education policy development. Examples are the 2013 reviews of international best practices in funding arrangements for education and strategies to combat educational disadvantage that were commissioned by the Ministry to various consortia of research groups.

As other Ministries in the Netherlands, the Ministry of Education, Culture and Science established a Knowledge Chamber in 2007, with the purpose of promoting evidence-based policy. Knowledge Chambers bring together top Ministry officials, chairpersons of Advisory Councils and Planning Bureaus, as well as civil society organisations, knowledge institutes and researchers who are typically invited on an ad-hoc basis. The intention is to provide a forum where policy makers and researchers can meet and discuss policy-relevant research and evaluation results. The Knowledge Chamber provides an infrastructure for policy-makers to consider key issues that cut across individual departments of the Ministry, interact with researchers and be aware of relevant findings that could inform future policy development. Although the work of the education Knowledge Chamber was suspended at the time of the OECD review visit, other modes for consultation were being discussed (Scheerens, 2013).

Another strength of system evaluation in the Netherlands is the beneficial tradition of programme evaluation. The basic principle that education programmes and policies should be evaluated is widely present at all levels of policy making in the Netherlands. The Ministry of Education gives considerable attention to policy and programme evaluation by promoting and sponsoring such independent studies among research and policy groups.

Results of system evaluation are used extensively

There is a strong commitment at the national level to feed the results of system evaluations back into the development of policy. Within the Ministry of Education, Culture and Science, the Knowledge Directorate is responsible for building the Ministry’s evaluative capacity. This involves building the infrastructure and the expertise to ensure that the Ministry plans effectively and uses evaluative information strategically for decision making.

Similarly, data from national assessments are used at the system level to analyse learning areas in need of particular attention. For example, the new policy focus on “excellence” results from analysis of student outcomes in both national and international assessments. Results from school-leaving examinations and end-of-primary-education Cito tests are also scrutinised to monitor developments in particular subject areas, including student progress in literacy and numeracy rates. Results from policy and

programme evaluations inform adjustments to policy and the identification of new areas of policy attention.

There has been considerable progress in the use of system-level data to inform school management. Through *Windows for Accountability*, schools are provided with individual detailed data about their activities that allows comparisons with other schools. This system is also developing specific indicators to support schools in their internal analysis and further planning (*The Management Window*). In addition, schools receive feedback on the results of their students in national school-leaving examinations and end-of-primary-education Cito tests. The same occurs for schools that participate in PPO, JPO and the longitudinal cohort studies.

Challenges

System evaluation would benefit from clearer national education goals

The OECD review team formed the impression that education system evaluation in the Netherlands would benefit from a clearer set of education goals. As mentioned above, the government does not clearly express its overall priorities for educational policy in a single document, which could provide a framework for policy development. National targets for the education system do not seem to be aligned to broader social and economic goals. This raises challenges in establishing both a solid foundation on which to build the evaluation and assessment framework, and a set of benchmarks against which the results of the overall system evaluation can be compared. As discussed in Chapter 2, the education authorities do not provide a vision or future-oriented strategy regarding key learning objectives preparing Dutch students life and work in the mid-21st century.

There has been good development with the establishment of Action Plans for education, which identify priority areas for action within specific areas of education policy. These are increasingly expressed in terms of explicit attainment targets, which facilitate policy and programme evaluation through the availability of clearer reference standards. However, these Action Plans typically focus on discrete education issues, such as primary or secondary education, youth or teachers, and there is no coherent strategic document bringing them together. In addition, the achievement targets established in these Action Plans cover a limited set of learning outcomes, primarily Dutch language and mathematics, and they usually focus on short- and medium term goals (with the exception of *Teaching 2020*).

The monitoring of complex competencies is limited

Designing measures of education system performance that are broad enough to capture the whole range of student learning objectives is a challenge shared across OECD countries. While all countries have some mechanism in place to monitor student learning, monitoring systems typically only provide performance information with respect to a subset of student learning objectives (OECD, 2013). The evolution of student learning objectives brings new demands on the monitoring systems. In particular, heightened importance in student learning goals for cross-curricular competencies and student ability to demonstrate and apply their learning in real-life settings are rarely adequately reflected in national measures of learning outcomes.

This is also the case in the Netherlands, as the national monitoring surveys rely on paper-and-pencil tasks, which are based on syllabi of single subjects. By their nature these types of questions can only measure a subset of student learning objectives. In

primary education particularly, standardised assessments rely heavily on multiple choice assessment formats, and this creates challenges to monitoring students' performance in more cognitively demanding areas. As a result, there is limited information regarding aspects such as creativity, critical thinking and team work, or the processes used by students to solve problems or conduct experiments.

There are concerns about the multiple uses of primary and secondary school examination results

There are concerns about the use of the Cito end-of-primary test and the secondary school examinations for system level target setting. Both assessments were initially designed to assess and certify individual student performance. As mentioned in Chapters 3 and 5, their use for multiple other purposes raises challenges. The Cito end-of-primary test was developed to identify student proficiency and readiness to enter different tracks of secondary education. It is not mandatory and about 15% of schools do not use this test, hence it is not fully representative of the education system. In the future, an end-of-primary school test will be mandatory, but schools will be able to choose which test they use, thus limiting comparability (more on this in Chapter 5). In the case of the secondary school examinations, their use as a system-level indicator also raises challenges of comparability. One of the two components of the examination is developed at the school level. This means that the coverage of learning targets for education in the school component of the national examination may vary considerably across schools. Although scoring guidelines are centrally provided, there is no systematic training for teachers who mark the national examinations (see Chapter 3). As a result, school-level average results in secondary school examinations are not strictly comparable across schools or over time, thus limiting their value for system-level target setting.

Equity aspects could receive more attention in system evaluation

The education system does not set specific targets for reducing educational disadvantage for particular groups, such as those students from lower income families, with a disability or with an immigrant status. As a result, system evaluation does not include measures to assess whether or not equity objectives are being achieved. The OECD review team formed the impression that relatively little differential analysis is undertaken on student performance across specific groups such as migrant students or students from disadvantaged families.

An area that seems insufficiently covered by national education statistics is the first language/language spoken at home by students. The OECD's Review of Migrant Education in the Netherlands also raised concerns about a narrow focus of data on the group of "non-Western" immigrants. The National Bureau of Statistics (CBS) uses this classification for practical purposes as the four major groups identified within the "non-Western" category (Turkish, Moroccan, Surinamese and Antillean) represent two-thirds of the immigrant population (Shewbridge et al., 2010). However, with recent trends of reduced numbers of "non-Western" immigrants entering the Netherlands, it would make sense to review this categorisation and consider adding categories for the newer immigrant groups arriving, in order to monitor the educational outcomes of all first-generation immigrants.

Information on school professionals, school practices and perceptions is not regularly collected

School autonomy and self-management create good conditions for school leader and teacher professionalism and continue to be strongly valued by school agents. This governance structure recognises that schools know their contexts best and allows school professionals to adopt diverse practices, according to their best judgment, and thereby creating conditions for innovation and system evolution.

At the same time, in such a system, characterised by a high level of local autonomy, school-based innovations in areas such as leadership, human resource management and the organisation of teaching and learning are usually neither documented nor disseminated. This may result in schools spending a lot of time on reinventing practices, some of which may be sub-optimal. In the context of freedom of education, individual schools can be relatively isolated and may have limited opportunities for learning from effective practice from across the country. Among the stakeholders interviewed by the OECD review team, there appeared to be a demand for more systematic channels and networks that could help schools share effective practice (Chapter 2). There is potential for the newly created *Initiative for Education Research* (NRO) to play a greater role in disseminating research results on effective educational practices.

From a system perspective, one option to learn from school-based expertise and perceptions is to go further in developing user surveys. The OECD review team commends the Netherlands for its regular use of student and parent surveys as part of education system monitoring. Going further, consideration could be given to conducting cyclical surveys to collect qualitative information about school processes directly from school staff. While surveys of teachers and school leaders are occasionally organised to inform specific policy initiatives, the Netherlands does not administer a regular survey of school staff. Collecting information from stakeholder surveys can provide important information for system monitoring as well for disseminating information on the views and practices of school professionals.

As teacher registration becomes increasingly common among teachers, more information regarding the teacher workforce will also become available through this channel. There is potential to use the registration system as a way to share knowledge on teachers' professional development and practices and to analyse the aggregate information with a view to informing future policies in support of teacher professionalism.

Policy recommendations

The above analysis indicates that system evaluation in the Netherlands is already highly developed. The OECD review team commends the sophisticated use of knowledge by decision-makers in the Netherlands, as well as the intensity and depth of communication between the educational administration and researchers. The Dutch approach to education system evaluation is a good example of how evaluation can be used to enhance the problem-solving capacity of a “smart state”. Continuous involvement of educational researchers and stakeholders in policy and programme evaluation exercises has strong potential to further improve the value of educational research for policy development. This section suggests a number of potential policy directions for future development of education system evaluation:

- Maintain the comprehensive approach to system evaluation.

- Develop clearer national goals for education.
- Recognise the main purposes and limitations of different indicators.
- Consider monitoring a broader range of student competencies.
- Further develop the monitoring of key aspects related to equity in education.
- Incorporate information on school professionals, school practices and perceptions in system evaluation.

Maintain the comprehensive approach to system evaluation

The Netherlands' comprehensive approach to education system evaluation is commendable. Educational authorities, as well as other education stakeholders, make it clear that education system evaluation is to play an important role in providing a firm evidence base for education policy making. This is demonstrated by the regular collection of data (student outcomes, administrative and contextual data), the systematic use of evidence for policy development (secondary analysis of data, qualitative studies, policy and programme evaluation), the importance given to dissemination of education data and research results, the involvement of the research community in education policy analysis, and the strategic approach to research, analysis and evaluation taken by education authorities (e.g. through the co-ordination of the Knowledge Directorate in the Ministry of Education, Culture and Science).

The centrality of education system evaluation in the Dutch evaluation and assessment framework should be sustained. This implies ensuring technical capacity and continuous investment in achieving the following major purposes of education system evaluation: (i) to monitor student outcomes at a given point in time, including differences among different regions within the education system and given student groups (e.g. by gender, socio-economic background or immigrant status); (ii) to monitor changes in student outcomes over time; (iii) to monitor the impact of given policy initiatives or educational programmes; (iv) to collect and analyse demographic, administrative and that are useful to explain the outcomes of the education system; (v) to maintain systems through which the relevant information is provided to the different stakeholders in the education system; and (vi) to use the generated information for analysis, development and implementation of policies.

As elaborated below, greater attention could be devoted to the analysis of learning outcomes across student groups, measures of student learning outcomes could be broader, and further work is needed to disseminate best practices across schools. Also, more could be done to ensure that programme evaluations are consistently conducted for all education programmes. To facilitate the evaluation of programme effectiveness and impact, it is important that all new programmes have an evaluation component in their original design, i.e. the programme plan should include elements to facilitate its evaluation, such as targets and baseline indicators.

Develop clearer national goals for education

In the broadest sense, education system evaluation aims to make a judgment regarding the extent to which the education system has met national goals and objectives for education. A key recommendation coming out of the *OECD Reviews of Evaluation and Assessment in Education* is to situate school system evaluation in the broader context of public sector performance requirements, given the key role of education in achieving

social cohesion and economic growth (OECD, 2013). When Ministries and other bodies with specific responsibilities for system evaluation need to show accountability for their performance, this stimulates demand for procedures to monitor progress in the school system and, where necessary, to establish appropriate systems to collect evidence on progress. This is particularly the case in systems where high level targets are set by the government related to economic and social improvement.

In the Netherlands, clear national goals for the education system, aligned to broader social and economic goals, could be more explicitly articulated. Such goals could respond to social and economic needs and reflect perspectives and views from outside the education sector. They can range, for example, from specific goals to provide high-quality education to students, to wider goals of promoting national values, civic responsibilities and innovation, creativity and entrepreneurship. There should also be particular attention given to equity in the provision of education services, and the need to improve educational outcomes for particular student groups. As a first step, bringing together the goals articulated in different Action Plans into a single document could help provide a broader vision and strategy for the education system (see Chapter 2).

Education system evaluation could benefit from a clearer and more comprehensive set of education policy priorities and targets with the associated indicators, so that progress towards achieving them can be measured. Such targets could cover a larger set of education system objectives and could cover a longer time span. One option would be for the Ministry of Education, Culture and Science to conduct an exercise to map out key objectives for the education system, followed by a set of specific goals or targets to be realised. To ensure that education policy is not driven by the availability of data, there should be a regular review of education system objectives and available measures within a meaningful and nationally agreed education system framework. This could be the foundation of strategies to prioritise further measurement development and/or refinement. Such an exercise will contribute to reminding all stakeholders of the full spectrum of national priorities, while also communicating that not all of these are currently measured or measurable (OECD, 2013).

The establishment of concrete education targets can help provide short- medium- and long-term goals for the education system, and a reference against which the effectiveness of education policies can be assessed. It can also help school agents better focus on the main policy challenges for the school system, as well as offer an opportunity for reflection on strategies to address the challenges, including a redeployment of resources to progress towards new goals. Box 6.1 below provides some information on the measurement framework currently in place in Australia.

Box 6.1 The Measurement Framework for Australian Schooling

In Australia, national targets related to schooling have been agreed and align with performance measures outlined in the *Measurement Framework for Australian Schooling*. This framework provides the basis for government reporting on the performance of schooling in Australia. A national target is defined as “a measurable level of performance expected to be attained within a specified time”. Australian Ministers (at the federal and state levels) first defined national Key Performance Measures in early 2000 as “a set of measures limited in number and strategic in orientation, which provides nationally comparable data on aspects of performance critical to monitoring progress against the National Goals for Schooling in the 21st Century”. The core of the framework is a schedule setting out key performance measures and an

Box 6.1 The Measurement Framework for Australian Schooling *(continued)*

agreed assessment and reporting cycle for the period 2006-2014. The framework clearly presents the agreed measures and their source for each of the priority areas:

- Literacy
- Numeracy
- Science literacy
- Civics and citizenship
- Information and communication technologies (ICT) literacy
- Vocational education and training (VET) in schools
- Student participation
- Student attainment
- Student attendance

Source: Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.

Recognise the main purposes and limitations of different indicators

The previous section described challenges around the use of the Cito end-of-primary test and the secondary school examinations for setting system-level targets and monitoring education system performance. Despite the fact that the primary purpose of the Cito end-of-primary test and the secondary school examination is for individual student assessment, they constitute a valuable element within a broader strategy for education system evaluation. For example, these assessments can provide information on the types of tasks students are able to master in different subjects, and they offer information to compare regions across the Netherlands at a particular point in time. Specific assessments to monitor the progress of the education system, such as the PPON, JPON and COOL studies, complement these external assessments and allow the setting of targets and monitoring of trends over time. Effective system level evaluation makes use of a broad set of available indicators and not just one measure. It is important, however, to clearly communicate the main purposes and limitations of different measures (OECD, 2013).

Consider monitoring a broader range of student competencies

Consideration should be given to further developing the sample-based assessments (PPON) by introducing a greater variety of tasks to assess a broader range of student outcomes. For example, the national monitoring assessments could include performance-type tasks where students are assessed on elements such as reasoning processes, problem-solving and oral communication skills. The range of competencies covered by sample-based assessments could also be extended to cover cross-curricular skills such as civic and citizenship skills, ICT literacy and learning-to-learn skills, and a range of personal

and interpersonal skills and attitudes. For example, in Australia the triennial sample-based assessments include an assessment of civics and citizenship skills and, in Finland, a survey is used to monitor students' "learning to learn" skills (OECD, 2013).

Across OECD countries it is not common practice to include oral questions and answers in national assessments, although this is done in Austria, Iceland, New Zealand and Sweden, as well as occasionally in the Flemish Community of Belgium. Austria and Sweden also include oral presentations in their national assessments. Both New Zealand and Sweden use the students' teachers in the scoring of assessment results, but in the other countries, students' oral tasks are centrally marked (OECD, 2013). In New Zealand, a broad range of performance-based tasks are included in the sample-based National Education Monitoring Project (NEMP), as explained in Box 6.2.⁶

Performance tasks are often seen as being more strongly aligned with learning goals that emphasise the development of higher-order thinking skills and the capacity to perform complex tasks. Introducing such measures would allow benchmarking different regions or specific student groups on an externally validated measure. It could also contribute to monitoring the extent to which students are stimulated to achieve excellence. However, performance-based tasks require a great deal of investment to ensure the comparability of scoring, for example through the development and provision of scoring guides and rubrics as well as through training of scorers. While the training of scorers implies higher costs, it can be a good source of professional development (OECD, 2013; see also Chapter 3).

Box 6.2 New Zealand's National Education Monitoring Project (NEMP)

Many of the NEMP assessment tasks are performance-based, requiring students to transfer learning to authentic close-to-real life situations. There are different assessment situations including one-to-one interviews, work stations and teamwork. As the assessment does not carry high stakes for students it is particularly important that tasks are meaningful and enjoyable to them. The assessment provides rich information on the processes used by students to solve problems or conduct experiments. Most assessment tasks are carried out orally so as to analyse what students can do without the interference of reading and writing skills. Some of the tasks are videotaped to allow for an in-depth analysis of student responses and interaction with teachers. NEMP also assesses students' cross-curricular skills, and attitudes towards the learning areas being assessed. Students' enjoyment of particular assessment tasks is also surveyed.

Source: Nusche, D., et al. (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116917-en>.

Further develop the monitoring of key aspects related to equity in education

There is room to give more prominence to the monitoring of inequities in learning outcomes between specific student groups. For example, education system targets could pay attention to the achievement of different student groups to monitor the equity of outcomes of, for example, boys and girls, students not speaking Dutch at home, students with a less advantaged socio-economic background, or students with a disability. The OECD's *Review of Migrant Education in the Netherlands* also recommended adding categories in regular data collection for the newer immigrant groups arriving in the Netherlands, with a view to introducing a regular monitoring of their educational outcomes and progress (Shewbridge et al., 2010). It would be of interest to provide

national assessment data broken down by specific student groups in order to monitor trends and analyse whether certain groups face particular challenges with some tasks. Overall, the value of annual monitoring reports could be further enhanced by regularly reporting information on student learning outcomes for groups where there is evidence of system underperformance. This would allow tracking the education system's progress in responding to the needs of diverse groups.

Incorporate information on school professionals, school practices and perceptions in system evaluation

An important aspect of knowledge management at the system level is to set up systematic processes to identify best practices within the school system and ensure that they are spread and shared across schools. In line with the government's emphasis on teacher professionalism and excellence, there should be a reflection about how to best monitor the quality of professional environments in schools. As teacher registration becomes more common in the Netherlands, there is a challenge for the Netherlands to tap the potential of this system to support knowledge management and system-wide learning. The Education Cooperative, together with key stakeholders and the national authorities, should engage in a reflection on how aggregate information from the teacher register can best be used and analysed to feed into future policies to support teacher professionalism.

In addition, to further expand the areas covered by education system evaluation, consideration could be given to establishing a mechanism to collect information from school professionals on a regular basis. A regular survey of teachers and/or school leaders could help collect more qualitative feedback on the education system, including on the teaching and learning environment, overall satisfaction, human and financial resource management, professional development and perhaps on specific innovations in the school system. This could also help mobilise and disseminate school-based expertise and promising approaches (see Chapter 2). While the Netherlands' participation in the OECD's *Teaching and Learning International Survey* (TALIS) will provide valuable information, a national survey would allow collecting more in-depth information on national priority topics.

Among the countries participating in the OECD Reviews of Evaluation and Assessment in Education, teacher surveys are administered in Australia, Austria, Chile, Estonia, Iceland, Israel, New Zealand, Norway, Portugal, Sweden and Northern Ireland in the United Kingdom, as well as in the Flemish Community of Belgium, Canada and Hungary as part of the national sample assessment (OECD, 2013). New Zealand, for example, uses national surveys on a three-year cycle that are completed by a random sample of primary and secondary schools (principals, teachers, trustees and parents) on school finance, strategic management, professional learning and collaboration with communities, plus current priority topics such as the introduction of national standards in 2010 (Nusche et al., 2012). Norway has established a multi-year framework for administering sample-based user surveys to ensure a cyclical coverage of key topics, while limiting demands on users to complete surveys (Norwegian Directorate for Education and Training, 2011).

Notes

- ¹ The CPB is one of three independent applied policy research institutes of the Dutch government.
- ² Nationaal Plan Onderwijs/Leerwetenschappen, Vergaderjaar 2010-2011, Kamerstuk 31288, No. 151.
- ³ Previously (until 2013) the NWO hosted two programme councils for education research (the programme council for fundamental education research [PROO] and the council for policy-related research into primary education [BOPO]), whereas practice-oriented research was subsidised by the government through the Law on Subsidising Educational Support Activities.
- ⁴ This includes general secondary education (HAVO), pre-vocational secondary education (VMBO) and pre-university secondary education (VWO).
- ⁵ At the time of publication of this report, the potential to use cohort studies in combination with register data on students' progress through education (using the student number) was being studied by a special committee of the Netherlands Initiative for Education Research (NRO).
- ⁶ In 2012, NEMP was discontinued and replaced by The National Monitoring Study of Student Achievement (NMSSA). NMSSA builds on the strengths of NEMP (see <http://nmssa.otago.ac.nz>).

References

- Cito (2008), *Onderwijs op peil? Een samenvattend overzicht van 20 jaar PPO*, Cito, Arnhem.
- Dutch Ministry of Education, Culture and Science (2011a), *Actieplan po Basis voor presteren*, www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/05/23/actieplan-po-basis-voor-presteren.html.
- Dutch Ministry of Education, Culture and Science (2011b), *Actieplan vo Beter presteren*, www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/05/23/actieplan-vo-beter-presteren.html.
- Dutch Inspectorate of Education (2010), *Risk-Based Inspection as of 2009, Primary and Secondary Education*, Dutch Inspectorate of Education, Utrecht, www.onderwijsinspectie.nl/actueel/publicaties/Riskbased+Inspection+as+of+2009.html.
- Norwegian Directorate for Education and Training (2011), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for Norway*, www.oecd.org/edu/evaluationpolicy.
- Nusche, D., et al. (2012), *OECD Reviews of Evaluation and Assessment in Education: New Zealand 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116917-en>.
- OECD (2013), *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264190658-en>.
- Santiago, P., et al. (2011), *OECD Reviews of Evaluation and Assessment in Education: Australia 2011*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264116672-en>.
- Scheerens, J., et al. (2012), *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands*, University of Twente, www.oecd.org/edu/evaluationpolicy.
- Scheerens, J. (2013), Addendum to *OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Country Background Report for the Netherlands*, University of Twente, Netherlands.
- Shewbridge, C., et al. (2010), *OECD Reviews of Migrant Education: Netherlands 2010*, OECD Reviews of Migrant Education, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9789264086210-en>.
- SCP (2004), *Ouders over opvoeding en onderwijs*, Sociaal en Cultureel Planbureau, The Hague.

Conclusions and recommendations

School system context

The Netherlands combines strong school autonomy with central steering, inspections and support

The Netherlands has one of the OECD's most devolved education systems, with schools enjoying a high degree of autonomy. School autonomy is grounded in the principle of "freedom of education", which gives the right to any natural or legal person to set up a school and to organise teaching. This constitutional arrangement puts public and private schools on an equal footing, with all schools receiving public funding, provided that they meet the requirements for schools in their sector. While "freedom of education" implies that schools are free to determine the content and methods of teaching, the central government sets learning objectives and quality standards that apply to both public and private schools. The Inspectorate of Education monitors school quality and compliance with central rules and regulations. There is also a large infrastructure of school service providers, most of them private companies, offering advice, instruments and support for school organisation, quality management and development.

The Dutch education system achieves good results internationally and is aiming to go further

The Dutch education system achieves very good results by international standards. Attainment rates of the Dutch population are similar to the OECD average and show positive trends. Dutch students have consistently shown above average performance in the OECD's Programme for International Student Assessment (PISA), with fewer low performers and more high performers among 15-year-olds than the OECD average. However, although the Dutch education system has a high standing on international assessments, there is a general appreciation it should continue to further improve the quality and equity of education. The performance of Dutch students in PISA since 2003 has decreased in mathematics and remained unchanged in reading and science. There are also marked performance differences across Dutch schools. National data indicate that the distribution of students across different education programmes is closely related to parental income and origin. There are also concerns about high levels of grade repetition, a large performance gap between immigrant and non-immigrant students and an increasing number of students being diagnosed as having special educational needs.

National priorities relate to promoting excellence among schools, enhancing teacher professionalism and supporting “results-oriented work”

The Dutch government has launched a general policy emphasising excellence in education. This includes a focus on providing better support to gifted and talented students and stimulating schools to aim for higher levels of achievement. In this context, the Ministry of Education, Culture and Science has implemented a prize for excellent schools, and the Inspectorate of Education is developing a differentiated inspection approach with the intention of helping schools that already provide basic quality education to further improve towards excellence. The Inspectorate has also conducted pilot studies to measure the “value added” by schools. As a cornerstone to promoting excellence in education, the Dutch government has made the enhancement of teacher professionalism a priority for education policy. Recent developments in this area include the definition of competency standards for teachers, a requirement for school boards to monitor their teachers’ competencies, the establishment of a teacher registration system and a voluntary teacher peer review project. A key element of this agenda is to strengthen teachers’ capacity for “results-oriented work”, using various assessment approaches to enhance student learning outcomes. Recent laws on student assessment make it mandatory for primary schools to implement regular student monitoring systems and a standardised end-of-primary test.

Strengths and challenges

Evaluation and assessment approaches in the Netherlands are highly developed, but some elements need further attention

The Netherlands has a strong tradition of well-developed evaluation and assessment systems. Central mechanisms for student assessment, school evaluation and education system evaluation have been in place for several decades, along with requirements for schools to assure their own quality. This long-standing focus on developing reliable evaluation and assessment systems has allowed the central actors to experiment with different approaches and develop and deepen their expertise over time. The country’s longstanding tradition and expertise in educational evaluation and assessment provides the Netherlands with a strong basis to further develop and integrate the evaluation and assessment framework. While most components of evaluation and assessment are well developed, this report finds there is room to further strengthen approaches to classroom-based student assessment, teacher appraisal and school self-evaluation, and to embed these with the broader evaluation and assessment system.

A balanced evaluation and assessment approach, which relies on the involvement of multiple stakeholders

The Netherlands is characterised by a complex system of governance with multiple actors shaping evaluation and assessment practices. In addition to the three traditional actors (the Ministry, the Inspectorate and the schools) a broad range of stakeholder organisations, private educational service providers and educational research institutions

also have an important influence on evaluation and assessment in the country. The involvement and continuous negotiation among these different groups has supported the development of a comprehensive and balanced approach to evaluation and assessment in the Netherlands. It combines school-based and central elements, quantitative and qualitative approaches, improvement and accountability functions, and vertical and horizontal responsibilities of schools. However, assuring coherence when several actors are pursuing different goals and following different approaches requires intensive cooperation, as well as the continuous development of shared goals and principles to guide evaluation and assessment practice.

Schools have broad autonomy in evaluation and assessment, which creates both opportunities and challenges for the system

Schools have broad autonomy in the area of evaluation and assessment. School autonomy and self-management create good conditions for school leader and teacher professionalism and continue to be strongly valued by school agents. This governance structure recognises that schools know their contexts best and allows school professionals to adopt diverse practices, thereby creating conditions for innovation and system evolution. At the same time, in such a system, characterised by a high level of school autonomy, school-based innovations related to evaluation and assessment are usually neither documented nor disseminated. This may result in schools spending a lot of time on reinventing practices, some of which may be sub-optimal. In the context of freedom of education, individual schools can be relatively isolated and may have limited opportunities for learning from effective practice implemented elsewhere. Without a strategy to effectively support, leverage and disseminate school-based innovations, many promising ideas will remain localised or even fade away for lack of external support.

Evaluation and assessment are policy priorities in the Netherlands, but they are not embedded in a broader vision or strategy for Dutch education in the 21st century

The national authorities regularly launch new priorities or initiatives related to evaluation and assessment in education, but they typically do not provide a narrative of how such policy priorities fit together into a coherent plan. Several stakeholder groups, including the professional organisations of teachers and school leaders, voiced concerns about the absence of a common vision for schooling in the Netherlands. By extension, at the time of the OECD review visit, there did not seem to be an articulated overarching vision or strategy for how evaluation and assessment should fit within broader educational improvement strategies, and which role they should play in achieving system goals. If evaluation and assessment are to be tools for improving learning, rather than the drivers of education in the Netherlands, it is critical that efforts are made to achieve a national consensus on the education goals for future generations. Among others, the Education Council advocates the need to re-focus the education system on broader learning goals, such as creative thinking, problem-solving and collaboration. Defining such learning goals for the 21st century would allow key stakeholders to engage in reflection and dialogue on how evaluation and assessment should evolve in order to support a future-oriented education system.

The focus on teacher professionalism is commendable and has potential to become a central element in the Dutch evaluation and assessment framework

The commitment of the government to boost teacher professionalism is a key contextual factor for evaluation and assessment, and there is room to further promote it as a central element in the Dutch evaluation and assessment framework. Concerning student assessment, although there is a good balance between school-based and standardised student assessments, there is a challenge in the Netherlands to build the competencies of all teachers to fully exploit the potential of assessment to transform and improve classroom practices. Regarding teacher appraisal, there is room to develop more systematic appraisal and feedback processes to ensure that all teachers receive adequate feedback on their performance along with relevant support for improvement and recognition of their professional achievements. The emphasis on teacher professionalism also creates new challenges for the design of school evaluation, as it requires the Inspectorate to monitor more closely the internal capacity of schools to foster teacher professionalism, manage human resources and create professional learning communities.

Standardised assessments provide reliable measures of student learning but there are risks that they are becoming too dominant

The Netherlands stands out internationally with regards to the development of high-quality standardised assessments at key stages of education. Major advantages of external standardised assessments include their high reliability and low cost of administration. They also help to clarify learning expectations for all schools and motivate teachers and students to work towards high standards and steer their teaching and learning strategies towards common goals. At the same time, there is a risk that the high visibility of standardised assessment might lead to distortions in the education process. In the primary sector, the end-of-primary test developed by the Central Institute for Test Development (Cito) uses only multiple choice items, with all the accompanying constraints on what skills can be adequately measured. Because of the role of the test results in determining placement in the secondary sector, and their use in the evaluation of schools, there is pressure on both students and teachers to concentrate on the specific content and format of the tests, which can lead to a narrowing of the delivered curriculum. Similar issues pertain to the assessments administered at the end of secondary education. Although schools have substantial autonomy in developing the school-based part of the examination, teachers interviewed by the OECD review team indicated that they tended to align their own assessments to the formats used in the central examination.

The newly introduced reference levels provide greater clarity on expectations for student learning, but they are only available for a sub-set of learning goals

The Dutch education system is organised around national “core objectives”, which describe in very general terms the knowledge and skills that students should achieve at the end of primary and secondary education. Additional reference levels for Dutch language and mathematics were recently published to provide more clarity about

instructional objectives at the primary and secondary level. Greater clarity should help teachers create syllabi at each grade level that better represent national learning goals, as well as develop assessments with improved coverage of those learning goals. Ideally, this will result in more equality in students' opportunity to learn and better alignment of instruction across different year levels and sectors of education. However, the reference levels cover only a sub-set of learning goals and some stakeholders expressed concerns that they remained quite vague. In the absence of clear and specific central expectations, learning goals in many areas are set with reference to existing assessments, in particular the Cito end-of-primary test and the national examinations at the end of secondary education.

The focus on “results-oriented work” is a good policy but requires further investment in capacity building

The government is promoting “results-oriented work” work in schools, which involves helping educators to more fully exploit student monitoring systems and, by analysing the information generated, to design appropriate teaching and learning strategies. Student monitoring systems, such as Cito's student monitoring system (LVS), are important tools that support results-oriented work at the school level. They can provide signposts for teachers and students by indicating the learning goals that are expected nationally and they offer timely data that may inform teaching strategies. At the same time, there is evidence that many teachers have difficulties in interpreting and effectively using the information generated by such assessments to improve teaching and learning and to provide effective feedback to students. It is likely that these challenges are shared by school leadership teams and that the information provided by student monitoring systems could be better utilised for decision-making at the classroom and school level. Teachers also face challenges in developing their own classroom-based formative assessment approaches to identify and respond to their students' specific learning needs, and to ensure that students are engaged in their own learning and assessment.

Summative student assessment relies on a good balance between central and school-based elements, although there are some concerns

Overall, there is a reasonable balance at key decision points between the use of school-based results and central examination scores. In the transition from primary to secondary education, for example, the school's recommendation is as important as the standardised test results in determining the school and the track most suitable for the student. At the end of secondary school, a subject assessment consists of both a centrally prepared examination and a school-developed examination. The fact that schools have some flexibility in deciding the content of their examinations is a mark of educators' professional autonomy, and can lead to improvements in the coverage of the syllabus. This flexibility, together with the integral role that teachers play in the construction and scoring of the central examinations, is likely to account for the general credibility that the secondary examination system has among teachers. While the involvement of teachers in this aspect of the final examinations is commendable, it requires substantial training of teachers and systematic moderation processes to build teachers' assessment competencies

and ensure reliability of marking. There is also some concern about teachers designing the school-based part of the examination to resemble the central part, which may reduce the scope of material covered in examinations, with potential adverse consequences on teaching and learning.

Professional standards for teachers exist, but they are not widely used in schools

There has been considerable reflection in the Netherlands around what is considered “quality teaching”. The Education Professions Act includes a description of teacher competencies, which functions as a professional standard for teachers. In addition, the Inspectorate’s classroom observation framework provides guidance on aspects of good teaching and the Education Council provides advice on how to develop teachers’ “personal professionalism”. Professional standards are an important element in any teacher appraisal system, as they provide a common reference to effectively review teacher competencies and offer the potential to frame and align the organisation of key elements of the teaching profession, such as initial education, teacher appraisal and professional development. However, while the competency requirements are a mandatory element of initial teacher education, they do not appear to be carried forward into the appraisal and professional development planning of teachers in schools. The lack of common reference standards for teacher appraisal is likely to limit the capacity of schools to effectively review their teachers’ competencies and risks weakening the alignment between different elements of teachers’ professional practice and development.

There are school-based approaches to teacher appraisal, but no guarantee that all teachers receive regular feedback on their performance

There are a range of channels through which teachers may receive school-based feedback, including through performance interviews with the school leader, internal coaching systems and informal peer learning within teacher teams and departments. However, while some teachers receive extensive feedback and support from their peers and school leaders, there is no mechanism in the Netherlands to ensure minimum standards for teacher appraisal and no guarantee that every teacher receives proper induction, appraisal and professional feedback. Most school boards delegate responsibility for the appraisal of teachers to the school leaders. As there is no mandatory school leadership training for principals, the preparation and competencies of principals vary across schools. Even though examples of school leaders exemplifying strong pedagogical leadership and human resource management certainly exist, there is a challenge for the Dutch system in building up the role and capacity of their full cohort of school leaders. School boards are responsible for ensuring that their schools have functioning personnel and appraisal policies in place, but there are large variations in the degree to which school board members have the background, capacity and commitment to do so in a systematic manner.

Teacher professional development is encouraged, but not sufficiently linked to teacher appraisal

The importance of professional development is clearly recognised in Netherlands and the introduction of the registration system further emphasises the expectation that all teachers engage in ongoing professional learning. Informal mentoring arrangements within schools also appear to be common practice. However, there are indications that teachers' choices of formal professional development are only rarely linked to a thorough evaluation of their strengths and areas for development. The Dutch Inspectorate of Education found that schools vary widely regarding the support they provide teachers to facilitate their professional development and that training is often too discretionary and lacking in focus on the actual teaching and learning process. There is also scope to better link teacher professional development to school development and improvement.

The introduction of a teacher register is positive and has potential for further development

The Education Cooperative (*Onderwijscoöperatie*), a teacher professional organisation created in 2011, maintains a voluntary registration system for teachers. To be registered, teachers need to meet criteria regarding the amount and content of professional development they have undertaken. However, teachers' actual practice and use of the new skills in the school and classroom are not evaluated as part of the registration process. Although the registration system has potential to become a key element in supporting teacher professionalism, it is still in the early stages of development and its role in the teacher career has not yet been clearly defined. Registration and registration renewal do not grant teachers access to a higher career step. The award of senior teacher positions is not typically linked to an appraisal of teachers' performance in relation to the competency requirements and there does not seem to be a formalised career path for teachers. The further development of the teacher registration system may provide valuable opportunities to further develop the teacher career in the Netherlands and it could help establish a link between teacher appraisal and teacher career development.

School self-evaluation relies on increasingly well-developed instruments and capacity, but it would benefit from further support

School self-evaluation is an important component of school evaluation in the Netherlands, and there has been considerable investment in building tools and capacity for this at the school level. At the same time, school self-evaluation and school inspection are conducted as parallel rather than cooperative activities. Hence, there is a risk that the methods, language and used criteria are not well aligned, which may diminish the developmental function of school evaluation. In addition, recent changes in the educational legislation and the Inspectorate's Supervision Framework indicate that the Inspectorate's attention in inspection visits has shifted away from school leadership and more towards school governing boards. However, most aspects of internal quality management and self-evaluation can be undertaken effectively only by the leadership teams at the school level, and it is therefore important that they benefit from external

review and feedback on their practices. Many school boards rely on the information provided to them by school principals, but there is little evidence that school boards conduct regular appraisal and review of their principals' work.

Inspections have been successful at helping underperforming schools improve, but new inspection approaches appear necessary to stimulate the progress of all schools

The Dutch Inspectorate of Education is a highly structured and professionalised institution whose operations follow clear goals and are guided by well elaborated and transparent rules. This is an important strength of the Dutch approach. The Inspectorate operates a risk-based approach, whereby schools at risk of underperformance are evaluated more frequently and more thoroughly than others. Its capacity to detect risks and launch effective interventions is well illustrated by the success of the risk-based approach to stimulate improvements in weak and unsatisfactory schools. Several studies conducted in the Netherlands confirm the positive impact of inspections on schools identified as providing weak or unsatisfactory quality. However, there are some doubts about the nature and strength of the impact of risk-based inspection on other schools. To further stimulate improvement in all schools including those that are already achieving basic quality, the Inspectorate will develop a “differentiated” approach to school inspection, adapting external inspection to the situation of individual schools. This is likely to create a more complex evaluation environment for inspectors, and will require ongoing development of their professional competences, quality criteria and inspection approaches. Inspectors will have to maintain a balance between the current data-oriented approach and new trends towards evaluating aspects of quality that are not easy to measure.

The availability of reliable information on learning outcomes is a strength but concerns exist that assessment results are increasingly put to multiple uses

A major strength of the Dutch school evaluation system is the highly developed system of collecting, managing and analysing student-level data, including in time series format, and the availability of such data in an aggregated form at school level. A related significant development has been the creation of user-friendly online information systems which connect different existing databases, summarise information about each school and grant opportunities for institutions to share information with each other and with a wide range of stakeholders. However, concerns exist that student assessment results are increasingly being put to multiple uses. In particular, the Cito end-of-primary test and the secondary school examinations, which were developed to identify the level of proficiency attained by individual students, are also used as key indicators of school quality. The high stakes for schools associated with these scores exert pressure on schools to improve test performance through a focus on the specific item formats employed by the test, and other strategies that may lack educational value.

The introduction of value-added modelling brings both opportunities and new challenges for school evaluation

While school-level information on student test results has been publicly available in the Netherlands for a long time, there is a general recognition that evaluating schools on the basis of their pupils' performance can be problematic because it does not take into account their status upon entry. Therefore, a range of pilot studies have been conducted to review the utility of value-added models (VAM) in judging school quality. VAM take into account the differences in intake among schools, so as to better "isolate" the contributions of schools to their students' progress. The development of value-added modelling is intended to address concerns regarding the potential misinterpretation of school league tables and to provide a more useful and credible identification of schools with high relative efficacy. Although the use of VAM addresses some of the concerns with conventional league tables, it does raise other issues. First, it is unlikely that the covariates employed in the regression model fully capture the differences among schools in enrolled students. These covariates are also subject to measurement error. Second, Dutch research found that differences in value-added among schools were relatively small and the (estimated) uncertainty associated with the estimates of school value-added were quite large. Thus, value-added analysis is a somewhat crude tool for accountability. Third, the complexity of the statistical models involved in VAM makes such approaches difficult to understand for teachers, parents and the general public. Unfortunately, the superficial, and often sensationalist, treatment of test results in the media often adds to the confusion.

The Inspectorate operates as an advanced "knowledge organisation", whose success also depends on effective collaboration with multiple actors

The Inspectorate of Education has many features of an advanced knowledge organisation. It has a dedicated knowledge directorate responsible for data collection, analysis, research and internal knowledge management and it maintains close cooperation with the Dutch educational research community. Thanks to its research-intensive operation, the Inspectorate seems capable of foreseeing a number of future challenges and of identifying appropriate responses, as illustrated by ongoing innovation efforts to: (i) make inspection more differentiated; (ii) integrating pedagogical-didactical and financial inspections; (iii) strengthening value-added approaches. However, in the Dutch multi-stakeholder environment, many school evaluation activities happen outside the formal activity sphere of the Inspectorate, for example through self-evaluation, supervision by school boards, school development programmes containing evaluation components and horizontal accountability mechanisms. Not all actors in this multi-actor landscape are equally prepared, which calls for heightened attention to clarifying responsibilities and building the capacity of those involved.

Education system evaluation receives considerable attention in the Netherlands but would benefit from clearer national education goals

The monitoring of the education system is a well-developed component of the Dutch approach to evaluation and assessment, which involves the collaboration of a wide range of players. The capacity for system evaluation is significant and the actors involved operate with a substantial degree of independence. The importance accorded to system evaluation in the Netherlands is reflected in the establishment of comprehensive information systems, sample-based national assessments, longitudinal surveys and programme evaluations; in the transparent reporting of system-level information in online databases and annual publications; and in the regular monitoring and evaluation of education policies and programmes to inform future educational planning. However, the government does not typically express its priorities for educational policy and the associated indicators in a single document which could provide a framework for policy development. This raises challenges in establishing both a solid foundation on which to build the evaluation and assessment framework, and a set of benchmarks against which the results of the overall system evaluation can be compared.

There is a good ability to monitor learning outcomes at the national level, but measures of student learning could be broader

The Netherlands collects a wide range of data on student learning outcomes in relation to core learning objectives and reference levels. Information on student learning outcomes is collected from sample-based assessments (PPON, JPON) in primary education, longitudinal cohort studies, end-of-primary-education tests, national school-leaving examinations and international student assessments. However, as other OECD countries, the Netherlands faces challenges in designing measures of education system performance that are broad enough to capture the whole range of student learning objectives. The national monitoring surveys rely on paper-and-pencil tasks, which are based on syllabi of single subjects. In primary education particularly, standardised assessments rely heavily on multiple choice assessment formats, and this creates challenges to monitoring students' performance in more cognitively demanding areas. As a result, there is limited information regarding aspects such as creativity, critical thinking, problem-solving, project development and team work.

There is room to further learn from the practices and perceptions of school professionals

While freedom of education enables schools to undertake their own experiments in relation to evaluation and assessment, these are typically not documented and collaboration among schools tends to be local. Dutch schools also have considerable freedom in deciding whether to participate in reform initiatives or research studies supported by the Ministry. Consequently, it is often problematic to obtain the representative samples of schools and students that are essential to high quality research. This can delay both innovation and a shift to more evidence-based policy making. There is also relatively limited information at the national level regarding the practices and

perceptions of school professionals. While surveys of teachers and school leaders are occasionally organised to inform specific policy initiatives, the Netherlands does not administer a regular survey of school staff. As teacher registration becomes increasingly common among teachers, more information regarding the teacher workforce will also become available through this channel. There is potential to use the registration system as a way to share knowledge on teachers' professional development and practices and to analyse the aggregate information with a view to informing future policies in support of teacher professionalism.

Policy recommendations

Embed the evaluation and assessment framework with broader learning goals for the 21st century

A crucial aspect for the successful implementation of evaluation and assessment is their alignment with student learning objectives. Thus, it seems essential to begin as soon as possible a broad-based consultative process to build consensus on a set of long-term learning goals for Dutch students that will prepare them well for the mid-21st century. In the Dutch governance context, this will need to involve multiple stakeholders, and negotiations are likely to be difficult given the principle of schools' freedom of education. However, a national conversation on how traditional learning goals should be augmented to meet the challenges of the 21st century will help identify the changes that must be made to the evaluation and assessment framework to support innovative, future-oriented and reflective teaching and learning. To make the evaluation and assessment system coherent, it is important that the learning goals to be achieved are placed at the centre of the framework and that all other evaluation and assessment activities align to work towards these goals. For example, competency descriptions for school professionals and quality indicators for school evaluation should reflect the learning goals that the school system is aiming to achieve.

Draw up an overarching strategy to further develop and integrate the evaluation and assessment framework

While the Dutch evaluation and assessment system benefits from strong central expertise, sophisticated measurement instruments and a range of targeted policy initiatives, it could benefit from a clearer overarching strategy linking its different elements and providing the rationale for its further development and implementation across the country. The process itself would help to identify gaps, missing links and potential imbalances to be addressed in future policy development. As elaborated further below, this report finds that there is room to further develop: research-based approaches to classroom assessment and the use of assessment results for improvement; systematic teacher appraisal linked to professional and school development; and effective school self-evaluation that is well articulated with external inspection. Taken together, these elements would contribute to strengthening the school-based components of evaluation and assessment, which would help maintain a balanced approach to evaluation and assessment. Such processes actively involving students, teachers and school leaders have strong potential to produce results that will be useful to shape future teaching and learning and have an impact on actual classroom practice.

Continue to build on teacher professionalism

As the most important school-level factor in student achievement, teachers are key to improving education outcomes. The OECD review team commends the current focus of the Dutch government to build on and further develop teacher professionalism in the Netherlands. Defining and rethinking the framework for evaluation and assessment also provides an opportunity to place teacher professionalism firmly at the heart of the evaluation and assessment agenda. Further work to enhance teacher professionalism can take place at different levels of the education system. Channels that are likely to reinforce the professionalism of teachers and to build links to classroom practice include: an emphasis on teacher appraisal for the continuous improvement of teaching practices; ensuring teaching standards are aligned with student learning objectives; involving teachers in collaborative school self-evaluation processes; preparing inspectors to understand the complexity of school-based human resource management and professional learning; ensuring that teachers feel the ownership of student assessment and accept it as an integral part of teaching and learning; enhancing teacher capacity to use assessment results for improvement; and building teachers' ability to assess students in relation to national goals and reference levels. As teacher professionalism is not yet appropriately reflected in the evaluation criteria of the Inspectorate, special emphasis could be given to these aspect in the next round of revisions.

Engage stakeholders and build networks for system learning

Effective implementation seeks to strike the right balance between top-down and bottom-up initiatives, which is generally believed to foster consensus. Given the traditional autonomy accorded to schools, any top-down imposition of innovative evaluation and assessment approaches is likely to be problematic in the Dutch context. It seems more feasible to develop evaluation and assessment policies through the co-operation of different stakeholders towards a common goal. This calls for practitioners, such as school leaders and teachers, to be engaged in the design, management and analysis of evaluation and assessment policies. New projects to develop innovation in evaluation and assessment should involve partnerships between evaluation and assessment organisations and groups of schools where pilots would take place. The Inspectorate should also play a key role in recognising and disseminating promising innovations developed at the school level. In addition, developing more deliberate improvement networks among practitioners can be a powerful organisational tool. The central authorities can contribute to creating an ambition-friendly and innovation-friendly environment, by providing funding and support for schools and networks of schools to accelerate their work and showcase their efforts to a broader audience.

Consider developing learning progressions to complement curriculum goals

Research-based learning progressions describing the way in which students typically move through learning in each subject area can provide a picture from beginning learning to expertise and help provide teachers, parents and other stakeholders with concrete images of what to expect in student learning, with direct links to the final learning objectives and reference levels. Such learning progressions can provide a clear conceptual basis for a coherent assessment framework, along with assessment tools that are aligned

to different stages in the progressions. They could be promoted as voluntary resources that teachers use as signposts in their assessment. They can help raise aspirations and communicate a focus on excellence and continuous improvement. In line with the government's focus on teacher professionalism, such guidance could help teachers design their instructional plans and classroom assessment strategies in alignment with national objectives and progressions. Teachers should also be encouraged to share and co-construct intermediate learning goals and assessment criteria with students so that they understand different levels of work quality. Such common work on goals and criteria can promote both student learning and reflective teaching practice with shared national goals in mind.

Develop an assessment strategy in line with 21st century learning goals

Although some 21st century skills are already incorporated in national learning goals in the Netherlands, it is likely that a broader set of these skills and competencies will become part of the goals that are set at the national level. Current paper and pencil tests with their limited item formats will not be able to appropriately assess these skills, neither for formative nor for summative purposes. Thus, there will be a need to develop the expertise and technical capacity to design, develop, deliver and evaluate more complex assessments. Given the novelty of 21st century skills for most teachers, formative assessment should be the primary focus, as it can contribute directly to improved learning. Ideally, a coherent set of formative assessments (across grades, within a sector), along with the corresponding scoring rubrics and exemplars of student work, will help to provide illustrations of both the learning goals and the expectations for student performance. In parallel, work needs to be done both to develop the assessment infrastructure and the expertise that will facilitate the introduction of the new forms of assessment demanded by new curricular goals.

Strengthen teacher competencies for effective development and use of assessment

In addition to further developing the assessment infrastructure, it is equally important to continue to build assessment competencies among both teachers and school leaders. To this end, assessment capacity, including the ability to use results for improvement, should be reflected in teacher standards and be addressed in a coherent way across teacher preparation programmes and publicly funded professional development programmes. Eventually, assessment-related competencies should become part of the teacher registration system and teacher appraisal approaches. This report identifies priorities for professional learning in the following areas: classroom-based formative assessment; assessment of complex student competencies; interpretation and use of assessment results for improvement; and reliable summative assessment and marking of examinations. This human capital development agenda will also require professional development of teacher educators and of providers of in-service teacher training. Fortunately, teacher education programmes can draw on the rapidly expanding resources available internationally. Inducements and support from the Ministry will be essential in this regard.

Review and refine teaching standards

A framework of teaching standards is an important reference point for teacher appraisal. While competency requirements for teachers exist in the Netherlands and are

widely used in initial teacher education, their use for regular appraisal and professional development in schools appears limited. To ensure coherence between initial teacher education, registration, appraisal and professional development, it would be helpful to promote the wider use of the competency requirements as a working document in schools underlying all of these processes. The current co-existence of several types of references for the evaluation of teaching call for their consolidation into a single set of standards, to develop a shared understanding of what counts as accomplished teaching in the Netherlands. The consolidated standards should also build on practice-based expertise and could be informed by a thematic review on the use of teacher appraisal standards and criteria by schools. Another helpful adjustment could be to develop clearer descriptions of competency requirements for different roles and career steps of teachers, with appraisal criteria specific to distinct career levels.

Strengthen school-based appraisal for teacher professional development

School-based formative teacher appraisal takes place in many schools across the Netherlands, typically with senior teachers, team leaders or department heads conducting classroom observations, and principals holding performance conversations with their teachers. However, further steps are necessary to ensure that all teachers across the country benefit from meaningful appraisal and feedback, pursue relevant professional development, and are able to implement improvements in the classroom. To make developmental appraisal processes more effective and consistent across the country, the OECD review team recommends that it should be: (i) school-based but underpinned by common reference standards; (ii) firmly rooted in classroom practice; (iii) carried out by qualified internal evaluators; (iv) externally validated by school governing boards and the Inspectorate. In addition, the Dutch education system could benefit from the introduction of more systematic school-based induction and feedback systems for new teachers.

Further develop the teacher career structure and link career development to a revised registration system

There is room to further develop the teacher career structure in the Netherlands in order to recognise and reward teaching excellence and allow teachers to diversify their careers. The revised teacher career structure could comprise a number of key stages, with access to each stage being associated to a more formal appraisal process, which could potentially be organised through the teacher registration system. Registration processes that are linked to career development could help provide incentives for teachers to perform at their best, bring recognition to effective teachers, support professional learning, and help recognise and spread good practice more widely. Registration and registration renewal processes at certain key stages in the teacher career could also provide useful information for accountability, hiring and tenure decisions, professional development and promotion opportunities, or, in particular circumstances, responses to underperformance. Such appraisal for registration/career advancement would be more summative in nature than the regular appraisal for professional development, and it would need to be ensured that processes are fair and the same standards are applied across schools. Given the stakes attached to appraisal for registration, decisions should draw on several types of evidence and multiple evaluators, and encompass the full scope of the work of the teacher.

Enhance the evaluation capacities of school boards and school leadership teams

As responsibility for self-evaluation, internal quality management and human resource management is shared between the school boards and school leadership teams, special attention should be paid to whether both of these partners assume their related responsibility appropriately. The management capacity of school boards could be improved through supporting the professional development of their permanent staff. There is a need to develop a differentiated approach to deal with school boards of different size and capacities, taking into account that some boards are highly professionalised organisations while others are more loosely coupled formations of volunteers. Research on the operation of school boards that reveals their capacities, possible shortcomings and potential should also be supported. At the same time, the capacities and actions of the school leadership teams within each school should remain a strategic aspect of external evaluation. Efforts aimed at improving the capacities of school leaders should be continued through targeted development interventions, peer evaluation, professional networks and school partnerships. Such support should also be targeted at teachers to promote the involvement of teacher teams in analysing student performance data and developing school-level pedagogical strategies.

Continue to adapt school evaluation to emerging needs

Given the multi-actor nature of school evaluation in the Netherlands, the Inspectorate should continuously map the environment in which it operates at both the national and local level, and take into account the potential impact of the other partners when designing its evaluation approaches. The current move of the Inspectorate towards the introduction of differentiated inspections provides a good opportunity to reconsider the relationship between internal and external school evaluation, and self-evaluation documents could gain a greater role in school inspections. To evaluate and further stimulate schools that have developed innovative approaches, external evaluators need in-depth knowledge of innovative learning environments and an understanding of how to evaluate practices that are not yet widely proven as resulting in good quality. The policy focus on teacher professionalism means that external evaluators will need to understand the complexities of internal human resource management and development practices in schools. Further innovations or pilot experiments may be required, for example, modifying the classroom observation framework or involving teachers as part of inspection visits. The current initiatives to integrate financial and educational inspection and to enhance the evaluation of school level human resource management practices should be realised in this context, taking stronger teacher professionalism into account. This could also lead to the creation of broader and more diverse inspection teams.

Further explore the formative use of value-added information for school evaluation

Given the general concerns regarding the use of value-added models for accountability, as well as the challenges revealed by studies conducted in the Dutch context, the Inspectorate should move cautiously in formally incorporating VAM estimates into the school indicator system. More empirical work is required before a reliable decision can be made on the choice of the statistical model. Further, the

complexity of the statistical models employed in value-added analysis results in a lack of transparency that is likely to meet with strong opposition. Consequently, the current focus should be on helping participating schools to use this information as a component of a more comprehensive self-evaluation effort, while enlisting a broader and much larger sample of schools to participate in future pilot studies. The current focus of the Dutch education system on promoting results-oriented work and encouraging the analysis of student growth and learning gains at the school level has potential to contribute to building schools' capacity in analysing trends and devising adequate improvement strategies.

Develop and communicate clear national goals for the education system

Education system evaluation could benefit from a clearer and more comprehensive set of education policy priorities and targets with the associated indicators, covering a larger set of education system objectives and a longer time span than is currently the case. Such goals could respond to social and economic needs and reflect perspectives and views from outside the education sector. There should also be particular attention given to equity in the provision of education services, and the need to improve educational outcomes for particular student groups. One option would be for the Ministry of Education, Culture and Science to conduct an exercise to map out key objectives for the education system, followed by a set of specific goals or targets to be realised. Bringing together the goals articulated in different Action Plans into a single document could help provide a broader vision and strategy for the education system. To ensure that education policy is not driven by the availability of data, there should be a regular review of education system objectives and available measures within a meaningful and nationally agreed education system framework. This could be the foundation of strategies to prioritise further measurement development and/or refinement. Such an exercise will contribute to reminding all stakeholders of the full spectrum of national priorities, while also communicating that not all of these are currently measured or measurable.

Consider monitoring a broader range of student competencies

With respect to education system evaluation, consideration should be given to further developing the national assessments for system monitoring (PPON) by introducing a greater variety of tasks to assess a broader range of student outcomes. For example, the national monitoring assessments could include performance-type tasks where students are assessed on elements such as reasoning processes, problem-solving and oral communication skills. The range of competencies covered by sample-based assessments could also be extended to cover cross-curricular skills such as civic and citizenship skills, ICT literacy and learning-to-learn skills, and a range of personal and interpersonal skills and attitudes. Performance tasks are often seen as being more strongly aligned with learning goals that emphasise the development of higher-order thinking skills and the capacity to perform complex tasks, although they require a great deal of investment to ensure the comparability of scoring, for example through the development and provision of scoring guides and rubrics as well as through training of scorers. While the training of scorers implies higher costs, it can be a good source of professional development.

Incorporate information on school professionals, school practices and perceptions in system evaluation

An important aspect of knowledge management at the system level is to set up systematic processes to identify best practices within the school system and ensure that they are spread and shared across schools. In line with the government's emphasis on teacher professionalism and excellence, there should be a reflection about how to best monitor the quality of professional environments in schools. As teacher registration becomes more common in the Netherlands, this opens opportunities to tap the potential of this system to support knowledge management and system-wide learning. The Education Cooperative, together with key stakeholders and the national authorities, should engage in a reflection on the degree to which aggregate information from the teacher register can best be used and analysed to feed into future policies to support teacher professionalism. In addition, to further expand the areas covered by education system evaluation, consideration could be given to establishing a mechanism to collect information from school professionals on a regular basis through a regular survey of teachers and/or school leaders.

Annex A: The OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes

The OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes is designed to respond to the strong interest in evaluation and assessment issues evident at national and international levels. It provides a description of design, implementation and use of assessment and evaluation procedures in countries; analyses strengths and weaknesses of different approaches; and provides recommendations for improvement. The Review looks at the various components of assessment and evaluation frameworks that countries use with the objective of improving student outcomes. These include student assessment, teacher appraisal, school evaluation and system evaluation. The Review focuses on primary and secondary education.¹

The overall purpose is to explore how systems of evaluation and assessment can be used to improve the quality, equity and efficiency of school education.² The overarching policy question is “How can assessment and evaluation policies work together more effectively to improve student outcomes in primary and secondary schools?” The Review further concentrates on five key issues for analysis: *(i)* designing a systemic framework for evaluation and assessment; *(ii)* ensuring the effectiveness of evaluation and assessment procedures; *(iii)* developing competencies for evaluation and for using feedback; *(iv)* making the best use of evaluation results; *(v)* implementing evaluation and assessment policies.

Twenty-five countries are actively engaged in the Review. These cover a wide range of economic and social contexts, and among them they illustrate quite different approaches to evaluation and assessment in school systems. This will allow a comparative perspective on key policy issues. These countries prepare a detailed background report, following a standard set of guidelines. Countries can also opt for a detailed Review, undertaken by a team consisting of members of the OECD Secretariat and external experts. Fourteen OECD countries have opted for a Country Review. The final comparative report from the OECD Review, bringing together lessons from all countries, will be completed in 2012.

The project is overseen by the Group of National Experts on Evaluation and Assessment, which was established as a subsidiary body of the OECD Education Policy Committee in order to guide the methods, timing and principles of the Review. More details are available from the website dedicated to the Review: www.oecd.org/edu/evaluationpolicy.

Notes:

- ¹ The scope of the Review does not include early childhood education and care, apprenticeships within vocational education and training, and adult education.
- ² The project’s purposes and scope are detailed in the OECD (2009) document entitled “OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes: Design and Implementation Plan for the Review”, which is available from the project website www.oecd.org/edu/evaluationpolicy.

Annex B: Visit itinerary

Monday 3 June	
08.15 – 09.30	Director-General for Primary and Secondary Education
09.30 – 10.30	Ministry of Education, Culture and Science
10.30 – 12.00	Inspectorate of Education
12.00 – 13.00	Student organisation (LAKS)
14.00 – 15.00	College for Examinations (CVE)
15.00 – 16.00	Schoolinfo: meeting on the <i>Windows for Accountability</i> system
16.00 – 17.00	Primary Education Directorate of the Ministry of Education, Culture and Science: meeting on the <i>Schools have the initiative</i> programme
17.00 – 18.30	Education Council
Tuesday 4 June	
09.00 – 10.00	Primary and Secondary Education Councils (Po-raad and VO-raad)
10.00 – 11.00	General Teacher Union (AOB) and Christian Teacher Union (CNVO)
11.00 – 12.00	Education Cooperative (<i>Onderwijscoöperatie</i>)
12.00 – 13.00	School Leaders Association (AVS)
14.00 – 17.00	School Visit One: Aloysius College Secondary School (HAVO, VWO), The Hague
17.00 – 18.00	Central Institute for Test Development (Cito)
Wednesday 5 June	
10.00 – 13.00	School Visit Two: Josef Hayden Primary School, Groningen
14.00 – 17.00	School Visit Three: Zernike College Secondary School (VMBO, HAVO, VWO), Groningen
Thursday 6 June	
10.00 – 13.00	School Visit Four: Prinses Beatrix Primary School, Hulshorst
15.00 – 16.00	Parents organisation (Ouders en COO)
16.00 – 17.00	Teacher educators : InHolland University of Applied Sciences, VU University Amsterdam
17.00 – 18.00	National Knowledge Centre on Special Education (LECSO)

Friday 7 June

09.30 – 12.30 School Visit Five:
Libanon Lyceum Secondary School (MAVO, HAVO, VWO), Rotterdam

14.00 – 17.00 School Visit Six:
Zuidwester Primary School, The Hague

17.00 – 18.00 Netherlands Bureau for Economic Policy Analysis (CPB) [cancelled]

Monday 10 June

08.15 – 09.00 Minister and State Secretary for Education, Culture and Science

09.00 – 10.00 National Bureau of Statistics (CBS)

10.00 – 11.00 Netherlands Court of Audit (*Algemene Rekenkamer*)

11.00 – 12.00 School support services: Edventure; Netherlands Institute for Curriculum Development (SLO); 678 Onderwijsadvisering; Centraal Netherlands

13.30 – 14.30 Education Foundation (*Stichting van het Onderwijs*) [cancelled]

14.30 – 17.00 Education Researchers

Annex C: Composition of the Review team

Henry Braun, an American national, earned a B.Sc. (Hon.) in mathematics from McGill University and a Ph.D. in mathematical statistics from Stanford University. After serving as an assistant professor of statistics at Princeton University, he joined Educational Testing Service in 1979, where he held a series of increasingly responsible positions. He was vice-president for research management from 1990 to 1999 and held the title of distinguished presidential appointee from 1999-2006. In 2007, he retired from ETS and assumed the position of Boisi Professor of Education and Public Policy in the Lynch School of Education at Boston College. His current interests include school and teacher accountability, the role of testing in education policy, the analysis of large-scale survey data and standard setting. In recent years, he has published on a variety of topics including the Black-White achievement gap, comparative school effectiveness, applications of multi-level modeling, and policy issues in the use of value-added analyses. He has done considerable work in the area of value-added modeling and authored *Using Student Progress to Evaluate Teachers: A Primer on Value-Added Models* (2006). He was a major contributor to the OECD monograph *Measuring Improvements in Learning Outcomes: Best Practices to Assess the Value-added of Schools* (2008), and was chair of the NRC panel that recently issued the publication *Getting value out of value-added: Report of a workshop* (2010).

Gábor Halász, a Hungarian national, is doctor of the Hungarian Academy of Sciences and professor of education at the Faculty of Pedagogy and Psychology of the University Eötvös Loránd in Budapest where he is leading the Centre for Higher Educational Management. He teaches, among others, education policy, education and European integration and global trends in education. He is the former Director-General of the National Institute for Public Education in Budapest (now Institute for Educational Research and Development) where he is now scientific advisor. His research fields are education policy and administration, comparative and international education, educational research and innovation and the theory of education systems. As an education policy expert and policy adviser, he took an active part in Hungary's educational-change process in the 1990s. Dr. Halász is one of the founders and president of the Board of the Hungarian School for Education Management. He actively participates in the professional training of Hungarian school leaders. Dr. Halász has worked as an expert consultant for a number of international organizations, particularly OECD, the European Commission, the World Bank, and the Council of Europe.

Deborah Nusche, a German national, is a Policy Analyst in the OECD Directorate for Education and Skills, where she has been since 2007. As part of the OECD Reviews of Evaluation and Assessment in Education, she has led country-specific reviews in New Zealand, Norway and Sweden and has been part of the review teams in Belgium (Flemish Community), Chile, the Czech Republic, Mexico, Portugal, the Slovak Republic and the United Kingdom (Northern Ireland). She previously worked on thematic reviews on *Education and Diversity* and *Improving School Leadership*. She co-authored the OECD reports "Synergies for Better Learning: An International Perspective on Evaluation and

Assessment” (2013), “Closing the Gap for Immigrant Students” (2010) and “Improving School Leadership” (2008). Deborah has previous work experience with UNESCO and holds an M.A. in International Development from Sciences Po Paris. She co-ordinated the review of the Netherlands and acted as *Rapporteur* for the review team.

Paulo Santiago, a Portuguese national, is a Senior Analyst in the OECD Directorate for Education and Skills, where he has been since 2000. He is currently the co-ordinator of the OECD Reviews of Evaluation and Assessment in Education. He has previously assumed responsibility for two major cross-country reviews, each with the participation of over twenty countries: a review of teacher policy (between 2002 and 2005, leading to the OECD publication “Teachers Matter”) and the thematic review of tertiary education (between 2005 and 2008, leading to the OECD publication “Tertiary Education for the Knowledge Society”). He has also led reviews of teacher policy and tertiary education policy in several countries. He holds a PhD in Economics from Northwestern University, United States, where he also lectured.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

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

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

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

OECD Reviews of Evaluation and Assessment in Education

NETHERLANDS

How can student assessment, teacher appraisal, school evaluation and system evaluation bring about real gains in performance across a country's school system? The country reports in this series provide, from an international perspective, an independent analysis of major issues facing the evaluation and assessment framework, current policy initiatives, and possible future approaches. This series forms part of the OECD Review on Evaluation and Assessment Frameworks for Improving School Outcomes.

Contents

Chapter 1. School education in the Netherlands

Chapter 2. The evaluation and assessment framework

Chapter 3. Student assessment

Chapter 4. Teacher appraisal

Chapter 5. School evaluation

Chapter 6. Education system evaluation

www.oecd.org/edu/evaluationpolicy

Consult this publication on line at <http://dx.doi.org/10.1787/9789264211940-en>.

This work is published on the OECD iLibrary, which gathers all OECD books, periodicals and statistical databases. Visit www.oecd-ilibrary.org for more information.

