

Opbrengstgericht werken

over de waarde van meetgestuurd onderwijs

Guuske Ledoux
Henk Blok
Marianne Boogaard
m.m.v. Meta Krüger

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Ledoux, G., Blok, H., Boogaard, M. Met medewerking van Krüger, M.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs.
SCO-Kohnstamm Instituut van de Faculteit der Maatschappij- en Gedragswetenschappen, Universiteit van Amsterdam (SCO-rapport nr. 812, projectnummer 40296).

ISBN 978-90-6813-870-2

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvuldigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen, of op enige manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher.

Uitgave en verspreiding:
SCO-Kohnstamm Instituut
Nieuwe Prinsengracht 130, Postbus 94208, 1090 GE Amsterdam
Tel.: 020-525 1201
www.sco-kohnstamminstituut.uva.nl
© Copyright SCO-Kohnstamm Instituut, 2009

Inhoudsopgave

Voorwoord	I
Managementsamenvatting	V
1 Aanleiding, probleemstelling en onderzoeksopzet	1
1.1 Aanleiding	1
1.2 Concept en context	3
1.3 Onderzoeksvragen	8
1.4 Opzet van het onderzoek	9
2 Van ‘data-driven teaching’ naar meetgestuurd onderwijs; een verkenning van de literatuur	15
2.1 Probleemstelling	15
2.2 Begripsanalyse	17
2.3 De evaluatieve cyclus	21
2.4 Effecten van meetgestuurd onderwijs	26
2.5 Het gebruik van vorderingengegevens in het basisonderwijs	31
2.6 Tussentijdse conclusies	35
3 Opvattingen van experts over meetgestuurd onderwijs	41
3.1 Doel van het deelonderzoek	41
3.2 Opzet van het deelonderzoek	42
3.3 Opvattingen van experts	45
3.3.1 Mate en wijze van voorkomen van meetgestuurd onderwijs	45
3.3.2 Voorbeelden van meetgestuurd onderwijs	48
3.3.3 Effecten	50
3.3.4 Draagvlak	52
3.3.5 Conditie	53
3.3.6 Neveneffecten en risico's	56
3.3.7 Aanbevelingen van de experts	57
3.4 Resultaten inventarisatieronde pabo's	59
3.5 Tussentijdse conclusies	64

4	Opvattingen van scholen over meetgestuurd onderwijs	69
4.1	Doel van het deelonderzoek	69
4.2	Opzet van het deelonderzoek	70
4.3	Resultaten quick scan	74
4.4	Resultaten opvolgingsonderzoek	81
	4.4.1 De analyse op groepsniveau	82
	4.4.2 De analyse van de individuele leerprestaties	87
	4.4.3 Uitvoeringsproblemen, tevredenheid en nadelen	90
	4.4.4 Voorwaarden voor verdere verbetering	95
4.5	De scholen voor speciaal basisonderwijs (sbo)	98
	4.5.1 Bezwaren tegen groepsanalyses	98
	4.5.2 Ontwikkelingen bij de individuele analyses	99
	4.5.3 Voorwaarden voor meetgestuurd onderwijs in het SBAO	100
4.6	Tussentijdse conclusies	101
5	Leraren over referentieniveaus	107
5.1	Achtergrond en onderzoeksvragen	107
5.2	Opzet van het deelonderzoek	108
5.3	Resultaten	109
5.4	Tussentijdse conclusies	115
6	Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen	119
6.1	Achtergrond en onderzoeksvragen	119
6.2	Opzet van het deelonderzoek: casestudies pilots taalbeleid	122
6.3	Resultaten: vijf voorbeelden van meetgestuurde onderwijspraktijk	124
	6.3.1 School 1	124
	6.3.2 School 2	132
	6.3.3 School 3	138
	6.3.4 School 4	142
	6.3.5 School 5	147
6.4	Knelpunten en succesfactoren	151
	6.4.1 Knelpunten en risico's	152
	6.4.2 Winstpunten en succesfactoren	154
6.5	Tussentijdse conclusies	156

7	Conclusie en discussie	161
7.1	Doel van het onderzoek	161
7.2	De begripsanalyse	161
7.3	Stand van de kennis in de literatuur	163
7.4	Beschikbare gegevens over leervorderingen	165
7.5	Gebruik van de gegevens over leervorderingen	166
7.6	De fasen in de evaluatieve cyclus	170
7.7	Draagvlak voor meetgestuurd onderwijs	172
7.8	Risicofactoren, succesfactoren, onbedoelde effecten	175
7.9	Conditie	179
7.10	Discussie en aanbevelingen	181
	Literatuur	189
	Bijlage	195
	Recent uitgegeven SCO-rapporten	197

Motto

'Teachers and administrators need data at more frequent intervals in order to improve their work with students'.

Data-driven teaching: Tools and trends (2006)

The teachers interviewed for our study suggest overwhelmingly that the concept of data-based decision making and continuous improvement is ideal but, under current conditions, is also unrealistic.

Ingram, Louis & Schroeder, 2004, p. 1283

Voorwoord

Verbetering van het taal- en rekenonderwijs is een centraal doel van de tegenwoordige beleidsagenda voor het primair onderwijs, zoals vastgelegd in de nota *Scholen voor morgen*. Een van de manieren om aan dit doel bij te dragen is een intensiever gebruik van toetsresultaten. Het afnemen van toetsen is in het basisonderwijs op zichzelf genomen geen nieuwe activiteit. Leraren doen dat al sinds jaar en dag, vooral om de vorderingen van individuele leerlingen te volgen en daarover aan ouders verslag te kunnen doen. Maar toetsen kunnen ook gebruikt worden om de effectiviteit van het voorafgaande onderwijs te onderzoeken en zo nodig te verbeteren. De term die het Ministerie van Onderwijs, Cultuur en Wetenschap voor deze manier van werken propageert, is *opbrengstgericht werken*. Ook andere instanties (zoals de PO-Raad, de Onderwijsraad, de Inspectie) zijn deze term gaan gebruiken. Opbrengstgericht werken betekent dat leraren zich in het bijzonder op leeropbrengsten zouden moeten focussen. Toetsresultaten spelen daarin een rol, maar ook de dagelijkse praktijk van observeren van leerlingen, nakijken van werk en dergelijke.

Opbrengstgericht werken is nog een betrekkelijk nieuwe term. Toen wij in november 2007 van het Ministerie het verzoek kregen een offerte uit te brengen voor het hier gerapporteerde onderzoek, was *data driven teaching* nog het sleutelwoord. Leraren zouden beter en meer gebruik moeten maken van objectieve gegevens over de leervorderingen. In de loop van het onderzoek hebben wij de Engelse benaming losgelaten en vervangen door een Nederlandse vertaling, *meetgestuurd onderwijs*. Om duidelijk te maken dat de twee termen qua betekenis grotendeels overlappen hebben we ze beide terug laten komen in de titel van dit rapport.

Wij hebben aan het rapport twee motto's meegegeven. Het eerste motto verwoordt de overtuiging dat leraren en beleidsmakers vaker gebruik moeten maken van toetsuitslagen als ze het onderwijs willen verbeteren. Dit motto stemt overeen met het uitgangspunt van het Ministerie als het over opbrengstgericht werken gaat.

Het tweede motto plaatst een waarschuwende noot. Leraren verwelkomen weliswaar in grote getale het idee van meetgestuurd onderwijs, maar de realisatie daarvan is in de praktijk nog niet zo simpel. Tezamen vormen de motto's een passende aanduiding van het spanningsveld waarbinnen het onderzoek tot stand is gekomen. Enerzijds is er de heldere opstelling van het Ministerie, anderzijds zijn er de voorzichtige reacties bij de *experts in the field*, die niet zozeer twijfelen aan het belang van beter onderwijs, maar wel aan de begaanbaarheid van de uitgestippelde route. De tijd zal leren welke van beide motto's toekomstige ontwikkelingen het beste weergeeft.

Wij danken het Ministerie van Onderwijs en Wetenschap voor het vertrouwen dat het ons heeft gegeven bij de opdrachtverlening tot het onderhavige onderzoek. Wij danken ook de vele leraren, schooldirecteuren en andere velddeskundigen die aan het onderzoek hebben meegewerkt. Velen van hen hebben ons belangeloos en met verve geïnformeerd over hun ervaringen met en overtuigingen over meetgestuurd onderwijs. Zonder hun inzet hadden wij het onderzoek niet tot een goed einde kunnen brengen.

Managementsamenvatting

Inleiding en vraagstelling

Verbetering van de taal- en rekenprestaties van leerlingen is een kernambitie in de beleidsagenda voor het primair onderwijs van het Ministerie van Onderwijs, Cultuur en Wetenschap. Een van de manieren om deze ambitie te realiseren is de invoering van *data driven teaching*. Data driven teaching houdt in dat leraren (beter) gebruik gaan maken van gegevens over de leervorderingen van leerlingen. Het gaat onder andere om het gebruik van scores op objectieve toetsen, waaronder de toetsen uit het leerlingvolgsysteem van het Cito. Weliswaar gebruiken vrijwel alle basisscholen zulke toetsen met een zekere regelmaat. Maar het vermoeden bestaat dat leraren de uitkomsten op toetsen onvoldoende benutten voor onderwijsevaluatie en -verbetering. Uit onderzoek is bekend dat leraren toetscores vooral benutten om de vorderingen van individuele leerlingen te volgen, maar in mindere mate om de effecten van het voorafgaande onderwijs op de groep als geheel te evalueren.

Over de toepassing van data driven teaching op Nederlandse basisscholen is nog weinig bekend. Met het oog op de implementatie van het beleid heeft het Ministerie ons een zevental onderzoeksvragen voorgelegd. Deze luiden als volgt.

1. Wat is volgens bestaand onderzoek de beste werkwijze voor data driven teaching? Uit welke handelingen of stappen bestaat data driven teaching?
2. Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen in het primair onderwijs tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?
3. In hoeverre en hoe worden deze gegevens gebruikt voor (1) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (2) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?
4. Welke handelingen of stappen van data driven teaching worden op scholen uitgevoerd, welke niet, en waarom?

Opbrengstgericht werken

5. Bestaat er draagvlak bij leerkrachten en schoolleiders voor data driven teaching?
6. Wat zijn risico- en succesfactoren bij data driven teaching? Zijn er eventueel onbedoelde effecten?
7. Wat hebben schoolleiders en leerkrachten nodig om data driven teaching te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

Opzet van het onderzoek

In de onderzoeksopzet vallen vijf deelonderzoeken te onderscheiden, gerapporteerd in afzonderlijke hoofdstukken van dit rapport. Begonnen is met een begripsanalyse en een korte verkenning van de onderzoeksliteratuur rond data driven teaching (hoofdstuk 2). In het kader van de begripsanalyse hebben we ons onder andere bezonnen op de vraag of data driven teaching voor het Nederlandse taalgebied een gelukkige terminologische keuze zou zijn. Wij hebben uiteindelijk de voorkeur gegeven aan de term *meetgestuurd onderwijs*. Meetgestuurd onderwijs hebben we omschreven als onderwijs waarbij betrokkenen (scholen, leraren, leerlingen) zich in hun taakuitvoering laten sturen door uitkomsten van metingen. Meetgestuurd onderwijs ligt qua betekenis dicht bij de betekenis die het Ministerie en anderen geven aan de term opbrengstgericht werken, maar geeft explicieter aan dat het gaat om het daarbij benutten van (meet)gegevens.

In het tweede deelonderzoek zijn vraaggesprekken gevoerd met onderwijsexperts, in het bijzonder met medewerkers van het Cito, onderwijsinspecteurs, schoolbegeleiders en lerarenopleiders. Een belangrijk gespreksonderwerp is geweest hoe scholen met toetsgegevens omgaan en hoe de invoering van meetgestuurd onderwijs zou kunnen worden ondersteund (hoofdstuk 3).

In de andere drie deelonderzoeken vormden leraren, interne begeleiders en schooldirecties onze gesprekspartners. In hoofdstuk 4 rapporteren we over een telefonische bevraging van ruim 250 schooldirecties of interne begeleiders, deels uit regulier en deels uit speciaal basisonderwijs. De respondenten hebben vragen beantwoord over hun huidige praktijk ten aanzien van meetgestuurd onderwijs. In hoofdstuk 5 zoomen we in op één deelaspect van meetgestuurd onderwijs, namelijk het gebruik van vaste, landelijk bepaalde streefniveaus, ook wel aangeduid als referentieniveaus. Respondenten waren 22 leraren, die onder leiding van het Cito betrokken zijn geweest bij het vaststellen van referentieniveaus voor medio jaar-

groep 5. In hoofdstuk 6 werpen we het licht op vijf afzonderlijke scholen, die wat betreft de invoering van meetgestuurd onderwijs als voorlopers beschouwd kunnen worden. De scholen zijn deelnemers aan de landelijke pilot Taalbeleid.

Conclusies

We vatten de conclusies samen aan de hand van de oorspronkelijke onderzoeksvragen. Tot slot formuleren we enkele discussie- en aandachtspunten, op basis van/naar aanleiding van de onderzoeksresultaten.

1. Wat is volgens bestaand onderzoek de beste werkwijze voor meetgestuurd onderwijs? Uit welke handelingen of stappen bestaat meetgestuurd onderwijs?

Meetgestuurd onderwijs is een vorm van onderwijs waarbij betrokkenen – leraren, schoolmanagers én leerlingen – zich in hun taakuitvoering laten leiden door uitkomsten van (vorderingen)metingen. Meetgestuurd onderwijs verloopt via de zogenoemde evaluatieve cyclus. De cyclus bestaat uit vijf stappen: (1) het vastleggen van doelen en standaarden, (2) het verzamelen van informatie, (3) het registreren ervan, (4) het interpreteren en (5) het nemen van beslissingen. Meetgestuurd onderwijs bestaat in twee varianten, formatief en summatief, afhankelijk van het type beslissing dat men wil nemen. Er is sprake van *formatief gebruik* als de meetuitkomsten worden gebruikt om het onderwijs in de loop van het leerproces bij te sturen. Er is sprake van *summatief gebruik* als het oogmerk is om achteraf vast te stellen welke leervorderingen zijn gerealiseerd. Van de formatieve variant is in onderzoek vastgesteld dat het bij kan dragen aan verbeterde leerresultaten. Van belang daarbij is vooral dat leerlingen inhoudelijke, taakgerichte feedback krijgen. De effectiviteit van de summatieve variant is meer omstreden. In de discussie hierover ligt het accent op mogelijke negatieve bijeffecten, waaronder onzuiver handelen van leraren en scholen, verarming van het curriculum en onderstimulering van betere leerlingen, omdat de meeste aandacht uitgaat naar zwakpresterende leerlingen.

Voor elke stap uit de evaluatieve cyclus zijn specifieke aandachtspunten te formuleren:

- Voor het *vastleggen van doelen en standaarden* geldt dat dit bij voorkeur moet plaatsvinden in meetbare termen en voldoende specifiek. Hierbij kan differentiatie plaatsvinden, bijvoorbeeld in minimumdoelen en streefdoelen.

Opbrengstgericht werken

- Het *verzamen van informatie* kan op verschillende manieren plaatsvinden en met verschillende instrumenten. Een combinatie van instrumenten verdient vaak de voorkeur omdat daarmee beperkingen van afzonderlijke instrumenten verminderd kunnen worden.
- *Registreren* kan eveneens op verschillende manieren, maar invoer van de gegevens in de computer is het meest gewenst, omdat dat betere analysemogelijkheden biedt. Belangrijk is dat er goede toegang tot de gegevens is voor zowel leerkrachten als schoolleiding.
- Voor *interpreteren* kan een onderscheid gemaakt worden tussen leerstofgerichte interpretatie (welk deel van de leerstof wordt beheerst?) en normgerichte interpretatie (wat is het prestatieniveau vergeleken met een norm?). Interpretatie wordt eenvoudiger als er vooraf gestelde doelen zijn (zie stap 1).
- Voor het *nemen van beslissingen* geldt dat dit niet altijd een beslissing in de richting van verbetering of verandering hoeft te zijn. Een beslissing kan ook zijn dat de geconstateerde situatie goed is zoals het is.

Meetgestuurd onderwijs is een wijze van feedback geven aan leerkrachten en/of leerlingen. Er is empirisch bewijs dat feedback geven prestaties verhoogt en dus gaat het in beginsel om een effectieve werkwijze. Voorwaarden zijn wel dat het om valide metingen en om realistische doelen gaat en dat er bereidheid is om conclusies te verbinden aan de uitkomst van de metingen.

De onderzoeksliteratuur geeft geen rechtstreekse aanwijzingen voor een 'beste werkwijze'. Wel zijn er verschillende valkuilen of risico's uit af te leiden: zie de bovengenoemde mogelijke negatieve bijeffecten.

2. Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen in het primair onderwijs tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?

Basisscholen beschikken voor de kernvakken Nederlandse taal en rekenen/wiskunde over een grote hoeveelheid vorderingengegevens. De gegevens zijn van uiteenlopende aard. De verspreidingsgraad van het Cito-leerlingvolgsysteem bedraagt rond de 90 procent. Daaruit mag men afleiden dat de overgrote meerderheid van leraren gebruik maakt van de daarin opgenomen toetsen voor technisch en begrijpend lezen, spelling en rekenen/wiskunde. Verondersteld mag bovendien

worden dat de meeste scholen zich houden aan het door het Cito aanbevolen afnameschema, dat grosso modo neerkomt op een halfjaarlijkse afname. Overigens is ook de verspreidingsgraad van de Eindtoets Basisonderwijs van het Cito hoog. Deze toets is in 2008 door 6300 van de bijna 7000 basisscholen afgenomen. Naast het gebruik van Cito-toetsen verzamelen leraren nog op andere manieren gegevens over leervorderingen. Voor technisch lezen gaat het om leesbeurten, voor spelling en rekenen/wiskunde om methodegebonden toetsen en ook observeren tijdens lessen wordt veelvuldig toegepast.

Voor de methodeonafhankelijke toetsen is meestal ook sprake van een vorm van centrale invoer; in de regel heeft de interne begeleider hierover de regie. Meer dan 60 procent van de scholen heeft het bij het Cito-leerlingvolgsysteem behorende computerprogramma in huis, dat bedoeld is als hulpmiddel bij het analyseren van toetsgegevens. Er zijn ook andere hulpmiddelen voor de analyse van gegevens, zoals software pakketten voor schooladministraties. Deze zijn eveneens op veel scholen aanwezig. Dat scholen beschikken over een geschikt hulpmiddel wil echter nog niet altijd zeggen dat ze het ook benutten voor analyses.

3. In hoeverre en hoe worden deze gegevens gebruikt voor (1) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (2) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?

Toetsgegevens worden primair gebruikt om de ontwikkeling van individuele leerlingen te volgen (75 procent van de basisscholen doet dit en 64 procent van de scholen voor speciaal basisonderwijs). In geval van zorgleerlingen spelen toetsgegevens ook een rol bij de planning en evaluatie van het onderwijsaanbod via handelingsplannen. Opvallend daarbij is dat toetsscores nauwelijks gebruikt worden om leerlingen feedback te geven, terwijl uit de literatuur bekend is dat het bieden van feedback een effectieve werkwijze is om prestaties te verbeteren.

Toetsgegevens worden in iets mindere mate gebruikt om de effectiviteit van het onderwijs op groepsniveau vast te stellen. Bovendien maken scholen daarbij weinig gebruik van vooraf bepaalde doelstellingen. Basisscholen met zwakke (taal)prestaties analyseren iets vaker op groepsniveau (74%) dan scholen die volgens de inspectie voldoende resultaten behalen (69%). In het speciaal basisonderwijs vindt analyse op groepsniveau heel weinig plaats (20%). De reden die

Opbrengstgericht werken

men daarvoor geeft is dat er te veel verschillen zijn tussen de leerlingen en groepsresultaten daarom weinig zeggen.

De aanleiding om analyses op groeps- en schoolniveau te gaan doen is nogal eens de wens van inspectie of bestuur.

De scholen zelf zijn positiever over hoe ze de toetsgegevens benutten dan externe experts, zoals de inspectie en medewerkers van het Cito. De oorzaak hiervan is een verschil in maatstaven. De Cito-medewerkers merken in hun contacten met scholen (bijvoorbeeld tijdens de cursussen die ze geven) dat er nog veel deskundigheid ontbreekt en dat schoolleiders en interne begeleiders lang niet altijd voldoende doorzien wat ze uit hun gegevens kunnen halen. De inspectie verwacht van scholen dat ze de leeropbrengsten analyseren en daarbij per vak naar ontwikkelingen in de tijd kijken. Volgens de inspectie gebeurt dit nog lang niet altijd goed genoeg. De scholen weten echter niet altijd wat er mogelijk is, respectievelijk verwacht wordt en vinden het in veel gevallen zelf voldoende als zorgleerlingen worden gesignaleerd en als globaal wordt bijgehouden of de gemiddelde leerprestaties ongeveer op het verwachte niveau liggen.

4. Welke handelingen of stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom?

Volgens de onderwijsexperts voeren veel scholen de evaluatieve cyclus onvolledig uit: weinig aandacht voor het vastleggen van doelen en standaarden, wel veel aandacht voor het verzamelen en registreren van informatie, weer veel minder aandacht voor het analyseren en het nemen van beslissingen. Uit de telefonische ronde langs ruim 250 schoolleiders en interne begeleiders komt een gunstiger beeld naar voren, namelijk het volgende:

Stap 1, doelen stellen

In het reguliere basisonderwijs zeggen de meeste scholen (ca 80 procent) vastgelegde leerstandaarden te hebben, behalve in de groepen 1 en 2. Dit geldt voor zowel lezen als rekenen. Van de scholen voor speciaal basisonderwijs heeft maar ongeveer een kwart vastgestelde leerstandaarden, dit geldt voor zowel lezen als rekenen en voor alle groepen.

Stap 2 en 3, verzamelen en registreren

Vrijwel alle scholen beschikken over een leerlingvolgsysteem met opslag van gegevens over leervorderingen op verschillende vakken. Nagenoeg altijd voor lezen en rekenen, zeer vaak voor overige onderdelen van het taalonderwijs (zoals spel-

ling, woordenschat) en nog relatief weinig voor andere leerdomeinen. Dit geldt voor het regulier en het speciaal basisonderwijs.

Stap 4, interpreteren

Een groot deel van de scholen geeft aan de verzamelde gegevens ook te analyseren en te interpreteren, namelijk ruim 80 procent van de basisscholen en 74 procent van de scholen voor speciaal basisonderwijs.

Stap 5, beslissingen nemen

Beslissingen over te ondernemen stappen worden voornamelijk genomen voor individuele leerlingen, het gaat hier om het opstellen en bijstellen van handelingsplannen voor zorgleerlingen. Indien nodig worden ook beslissingen genomen over het onderwijsaanbod voor de hele groep, dan gaat het bijvoorbeeld om de keuze van een nieuwe methode of het inzetten van extra lesmateriaal.

Een discrepantie tussen norm en eigen resultaten is niet altijd aanleiding tot actie. Soms vindt men de norm voor de eigen school te hoog gegrepen of heeft men het vertrouwen/de ervaring dat een geconstateerde achterstand nog wel bijtrekt.

Het verschil tussen de oordelen van de experts en dat van de scholen zelf is ook hier weer een verschil in maatstaven: de experts zijn strenger in hun oordeel over wat een goede manier is van doelen stellen, gegevens interpreteren, etc.

Moeilijkheden waar scholen tegenaan lopen, zijn: tijdgebrek (vooral volgens scholen zelf), onvoldoende expertise ten aanzien van het analyseren van gegevens (volgens onderwijsexperts) en beperkte mogelijkheden om de geïndiceerde beslissingen – bijvoorbeeld dat leraren meer moeten differentiëren – ook daadwerkelijk in praktijk te brengen (volgens beide groepen).

Leraren worden niet altijd bij de interpretatiefase betrokken. Resultaten worden dan alleen met hen besproken als er ‘onder de maat’ wordt gepresteerd. Op de voorbeeldscholen zijn de leraren wel altijd in beeld. Wanneer het om individuele leerlingen gaat, is er meestal sprake van overleg tussen de groepsleerkracht en de interne begeleider over toetsresultaten, ontwikkelingsverloop en het handelingsplan. De groepsresultaten op schoolniveau worden over het algemeen besproken in een teamoverleg. Een enkele keer vormen ze ook een aandachtspunt in het functioneringsoverleg.

Opbrengstgericht werken

5. Bestaat er draagvlak bij leraren en schoolleiders voor meetgestuurd onderwijs?

Volgens de onderwijsexperts is er niet, of nog niet, een sterk draagvlak voor meetgestuurd onderwijs. Als scholen al aandacht besteden aan zelfevaluatie, maken de leervorderingen van de leerlingen daarvan volgens hen slechts een bescheiden onderdeel uit. Vooral scholen met weinig kinderen uit achterstandsgroepen hebben volgens de experts geringe belangstelling voor meetgestuurd onderwijs. Als de resultaten toch wel goed zijn, ziet men geen aanleiding om veel aandacht te besteden aan analyse van resultaten, is de ervaring. Op zwakke scholen of scholen met veel kansarme leerlingen is volgens de experts het draagvlak groter: daar is de noodzaak om prestaties te bewaken en waar nodig te verhogen groter.

Uit de telefonische ronde onder schooldirecteuren en interne begeleiders komt echter naar voren dat er voor meetgestuurd onderwijs al wel een duidelijk draagvlak is. Ze zijn het bijvoorbeeld in meerderheid eens met de volgende uitspraken:

- verbetering van het onderwijs in taal en rekenen kan niet zonder regelmatig meten van je onderwijsopbrengsten
- op onze school is er voldoende openheid in het team om met elkaar de leeropbrengsten in ieders groep te bespreken
- jezelf vergelijken met andere scholen met hetzelfde type leerlingenpubliek is een goede manier om te evalueren of de leerlingen presteren wat verwacht mag worden.

In het kader van de verkenning van draagvlak is in interviews met een twintigtal schoolleiders/interne begeleiders ook gevraagd naar hun mening over de invoering van landelijke referentieniveaus. Deze kunnen immers een hulpmiddel zijn bij het bewaken van de onderwijskwaliteit en een prikkel vormen om te proberen leeropbrengsten te verhogen. Over referentieniveaus blijkt wisselend te worden gedacht. Een klein aantal (minder dan 20%) vindt de plannen voor landelijke referentieniveaus een goede ontwikkeling. Eveneens een klein aantal (minder dan 20%) is er beslist tegen. Een deel (ca 35%) is het wel met invoering eens, maar onder de voorwaarde dat er bij het opstellen van referentieniveaus rekening wordt gehouden met verschillen tussen scholen in leerlingenpubliek. De rest vindt dat er al referentieniveaus zijn, en doelt daarmee op de normeringen in het Cito Leerlingvolgsysteem.

Ook de leraren die betrokken zijn geweest bij de Cito-pilot 'Referentieniveaus medio jaargroep 5' plaatsen kanttekeningen bij het vastleggen van landelijke refe-

rentieniveaus. Men ziet voordelen (een toenemende duidelijkheid), maar ook nadelen (grote druk op scholen met achterstandsleerlingen). Ongeveer de helft van de leraren onderschrijft de verwachting dat referentieniveaus zullen bijdragen aan betere resultaten. De andere helft heeft deze verwachting (nog) niet, omdat men meent dat de marges voor onderwijsverbetering smal zijn.

6. Wat zijn risico- en succesfactoren bij meetgestuurd onderwijs? Zijn er eventueel onbedoelde effecten?

De scholen uit de telefonische ronde noemen de volgende successen van meetgestuurd onderwijs: je kunt je onderwijs gerichter bijstellen; planmatig werken verbetert; er komt meer aandacht voor differentiatie en voor afstemming tussen de groepen; het maakt van de school een lerende organisatie; er ontstaat een beter inzicht in zwakke plekken.

De leraren uit de casestudies melden bovendien dat zij meer greep hebben op hun onderwijs en bewustere keuzes maken bij de selectie van leerstof. De constatering dat de leerprestaties inderdaad vooruit gaan werkt daarbij voor zowel de leerkrachten als de leerlingen zelf extra motiverend. Dat werkt ook weer positief uit voor het draagvlak voor meetgestuurd onderwijs ('je ziet waarvoor je het doet').

Risico's of onbedoelde effecten zijn er ook. De experts en de onderzoeksliteratuur brengen de volgende (mogelijke) risico's naar voren: verwaarlozing van de curriculumelementen die niet worden getoetst; vermindering van de leertijd (als gevolg van veelvuldig toetsen); een verschuiving van de balans tussen formatieve en summatieve toetsing in het nadeel van de eerste; een ongezonde prestatiedruk op leraren en leerlingen; strategisch gedrag – of erger, onzuiver handelen – van scholen; bevorderen van competitiegedrag; te veel nadruk op verantwoordingseisen; te sterke focus op zwakke leerlingen; alleen maar toetsen 'omdat het moet'.

De scholen zelf noemen veel minder risico's. Een risico dat ze wel onderschrijven is dat er een kans is op 'stigmatiseren' van leerlingen of leraren. Ook de angst voor een afrekencultuur is bij hen soms aanwezig.

7. Wat hebben schoolleiders en leraren nodig om meetgestuurd onderwijs te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

We onderscheiden aanbevelingen op verschillende gebieden.

Opbrengstgericht werken

- (1) De schoolcultuur moet zodanig zijn dat er systematische en ruime aandacht is voor leeropbrengsten én voor de condities die daarop van invloed zijn. Het schoolteam moet bovendien in staat zijn tot een open gedachtewisseling op basis van onderling vertrouwen.
- (2) Scholen hebben instrumenten nodig om meetgestuurd onderwijs te kunnen realiseren. Landelijke standaarden kunnen daarbij helpen, maar de toepassing daarvan roept nog wel de nodige vragen op. Ook zou er (voorbeeld)documentatie kunnen worden ontwikkeld om leraren en scholen te helpen bij de uitvoering van de complete evaluatieve cyclus.
- (3) Er zou verdere deskundigheidsbevordering moeten plaatsvinden van leraren, schoolleiders en interne begeleiders, maar ook van schoolbegeleiders en docenten van lerarenopleidingen. Daarbij gaat het niet alleen om het toevoegen van technische kennis over data-analyse. Wat ook geleerd moet worden, zijn strategieën om in de scholen ‘leerconversaties’ op gang te brengen.
- (4) Het schoolbestuur en de oudergeleding moeten gerichte belangstelling tonen voor wat leerlingen hebben geleerd. De inspectie vervult al een stimulerende rol en zou dit wellicht nog meer kunnen doen, vooral waar het gaat om het prikkelen van scholen die niet zwak zijn maar wel nog beter kunnen.

Discussie- en aandachtspunten

Meetgestuurd onderwijs is in potentie een krachtig middel om leraren en schoolleiders bewuster en doelgerichter met verbetering van leerprestaties bezig te laten zijn. Maar een kanttekening bij deze verwachting is op zijn plaats: van meten en analyseren alleen wordt het onderwijs niet beter. Uiteindelijk moet de focus liggen op verbeteringen van het onderwijs in de klas, zoals meer leertijd, betere instructie, meer en betere feedback aan leerlingen, hoge eisen stellen, aandacht voor leerstrategieën.

Volgens de geïnterviewde schoolleiders vinden niet alle leerkrachten op hun scholen het al een normaal deel van hun taak om gegevens over leervorderingen te verzamelen, vast te leggen, te analyseren en te bespreken in de school. Zij zien meetgestuurd onderwijs eerder als verhoging van de werkdruk dan als onderdeel

Managementsamenvatting

van hun professionaliteit. Voor hen is een omslag in denken op dit punt (wat hoort bij een professionele taakuitoefening?) nog nodig.

Het draagvlak voor meetgestuurd onderwijs lijkt in het regulier basisonderwijs redelijk aanwezig (in het speciaal basisonderwijs is het minder). Maar dat betekent niet dat het 'zomaar' overal een praktijk zal zijn. De complete toepassing is een innovatie. Bij de verdere ontwikkeling is het van belang inzichten uit de innovatieliteratuur te benutten (zie verder hoofdstuk 7).

Bij de eventuele invoering van referentieniveaus moet bedacht worden dat bij onvoldoende differentiatie alleen scholen met veel achterstandsleerlingen voor de opgave zullen komen te staan om de prestaties te verhogen, omdat zij de gestelde standaarden nog niet halen. Scholen met een kansrijker leerlingenpubliek zullen niet geprikkeld worden tot prestatieverbetering, omdat zij die standaarden al wel halen. Als prestatieverbetering in de breedte het doel is, is het verstandig om ook voor de scholen met een kansrijk leerlingenpubliek de lat hoog te leggen en voor hen passende referentieniveaus te formuleren.

1 Aanleiding, probleemstelling en onderzoekopzet

1.1 Aanleiding

In november 2007 heeft de staatssecretaris van onderwijs de beleidsagenda primair onderwijs voor de komende jaren uitgebracht, in de nota *Scholen voor Morgen* (Ministerie van OCW, 2007). Een centraal element in het beleid is het streven naar verbetering van de taal- en rekenvaardigheden van leerlingen in het primair onderwijs. Het achterliggende motief hiervoor is tweeledig. Enerzijds wil men de op zichzelf goede gemiddelde prestaties van leerlingen in het basisonderwijs nog verder verhogen, vanwege het cruciale belang daarvan voor de kenniseconomie. In het primair onderwijs wordt immers de basis gelegd waarop alle verdere opleiding moet worden gebouwd. Anderzijds bestaan er nog punten van zorg over het huidige kennisniveau. Bijvoorbeeld over het aantal leerlingen dat beschouwd moet worden als ‘zwakke lezer’ bij het verlaten van het basisonderwijs (Inspectie van het Onderwijs, 2006), het aantal leerlingen bij wie sprake is van onderpresteren (Onderwijsraad, 2007) en de nog altijd forse achterstanden bij leerlingen uit de doelgroepen van het onderwijsachterstandenbeleid (Mulder, Roeleveld & Van der Veen, 2005).

Het beleid voor het primair onderwijs legt daarom voor de komende jaren de focus bij verbetering van het onderwijs in de basisvaardigheden. Die verbetering wordt nagestreefd over de hele linie, maar vooral daar waar nog ‘reservekapitaal’ vermoed kan worden. Men verwacht dat die verbetering mogelijk is als er meer aandacht is voor opbrengstgericht werken, als de zorg voor zwakke leerlingen wordt opgevoerd en als het onderwijs beter wordt afgestemd op verschillen. Dit zijn allemaal punten waar volgens de onderwijsinspectie op veel scholen nog een kwaliteitssprong mogelijk en gewenst is. Als middelen die scholen kunnen gebruiken om die verbetering tot stand te brengen worden genoemd: leerstandaarden (referentieniveaus taal en rekenen), systematisch en evaluatief gebruik van het leerlingvolgsysteem en gebruik van goede methoden die bij voorkeur ‘evidence based’ zijn.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

In de nota *Scholen voor morgen* is ook aandacht voor de condities die hiervoor nodig zijn. Gewezen wordt op het belang van schoolleiders die in staat zijn om kwaliteitsbeleid te voeren en een ontwikkel- en verbetercultuur op de school te creëren en op de rol van de inspectie die moet stimuleren en intensief toezicht moet houden op scholen die zwak of risicovol zijn. Verder wil men stimuleren dat er wordt geleerd van elkaar, binnen de school maar ook schooloverstijgend, en van deskundigen buiten de school (zie ook Ministerie van OCW en PO Platform Kwaliteit en Innovatie, Innovatieagenda Primair Onderwijs, 2006).

Eén van de manieren om leerprestaties te verbeteren is volgens het Ministerie het beter gebruik maken van objectieve gegevens over het prestatieniveau en de leervorderingen van leerlingen. De meeste scholen beschikken over dergelijke gegevens in hun leerlingvolgsysteem, maar het vermoeden bestaat dat die nog onvoldoende worden benut voor evaluatie van het onderwijs. Wanneer scholen de beschikbare gegevens beter of vaker zouden benutten als bron voor feedback, zou dat leiden tot gerichte aanpassingen of interventies en tot betere formulering van doelen, zo is de verwachting. Deze werkwijze wordt in de Engelstalige literatuur *data driven teaching* genoemd en wordt gezien als een methodiek die effectief is voor het bereiken van betere leerresultaten (o.a. Fullan, 2007). In de Nederlandse onderwijswereld is deze term voorlopig vertaald als *opbrengstgericht werken*.

Over de toepassing van data driven teaching/opbrengstgericht werken op basisscholen in Nederland is nog weinig bekend. Meer inzicht daarin is nodig om te kunnen bepalen of er stimuleringsmaatregelen nodig zijn om het te bevorderen of te verbeteren. Het SCO-Kohnstamm Instituut heeft in opdracht van het Ministerie van OCW hier naar een onderzoek uitgevoerd. Het voorliggende rapport bevat daarvan de resultaten.

In dit eerste hoofdstuk beschrijven we de opzet van het onderzoek. We gaan in op concept en context, op de onderzoeksvragen en op de gekozen methode. Het onderzoek bestaat uit een aantal deelonderzoeken. Elk deelonderzoek wordt kort toegelicht. In de rest van het rapport worden de uitkomsten per deelonderzoek gepresenteerd, in afzonderlijke hoofdstukken (hoofdstuk 2 tot en met 6). Meer uitvoerige gegevens

over opzet en uitvoering van elk deelonderzoek zijn in deze hoofdstukken te vinden. In het laatste hoofdstuk (hoofdstuk 7) brengen we de resultaten bij elkaar en beantwoorden we alle onderzoeksvragen. Dit hoofdstuk wordt afgesloten met enkele aanbevelingen.

1.2 Concept en context

In de onderzoeksliteratuur wordt data driven teaching geplaatst binnen de context van schoolontwikkeling. Schoolontwikkeling is het proces van vernieuwen en verbeteren van het onderwijs in de school, bij voorkeur op permanente basis (het onderwijs staat immers nooit stil, verbetering/verandering is altijd nodig) en geleid door een duidelijke toekomstvisie. Schoolontwikkeling kan gericht plaatsvinden als het wordt verbonden met kwaliteitszorg. Kwaliteitszorg is het op systematische wijze zorgen voor behoud en verbetering van de kwaliteit van het onderwijs. Scholen zijn wettelijk verplicht om een kwaliteitszorgsysteem te hebben. De inspectie, die hierop toezicht houdt, hanteert ter beoordeling hiervan indicatoren die onder meer betrekking hebben op systematische evaluatie en planmatigheid van verbeteractiviteiten (Hoffmanns, 2007). Scholen die systematisch en planmatig naar hun eigen kwaliteit kijken, hebben daarmee kennis in handen waarmee ze gericht beleid kunnen voeren. Het kan daarbij gaan om kennis op allerlei terrein, kwaliteitszorg gaat immers niet alleen over leeropbrengsten maar ook over andere domeinen van het onderwijs. Het principe is echter steeds hetzelfde: het gaat om een cyclische manier van werken die steeds de stappen plannen, uitvoeren, evalueren en handelen (in de zin van verdere beslissingen nemen) omvat. Evalueren moet dan gebeuren op basis van gegevens, bij voorkeur zo betrouwbaar mogelijk. Wanneer dat gebeurt, is sprake van toepassen van *datafeedback*, en wanneer dat wordt verbonden met beslissingen die op basis daarvan worden genomen over de inrichting van het onderwijs, is er sprake van *data driven teaching*. Het hoeft hierbij niet perse te gaan over gegevens die betrekking hebben op leerresultaten van leerlingen. In het onderhavige onderzoek is dat echter wel het kernonderwerp. Het kader voor het onderzoek is immers het beleidsstreven naar verhoging van prestaties in de basisvaardigheden.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Er is nog niet veel (empirisch) onderzoek beschikbaar naar toepassen van datafeedback. In hoofdstuk 2 van dit rapport gaan we in op de uitkomsten van een literatuurstudie naar dit onderwerp en rapporteren we over enkele specifieke vragen die we met behulp van deze studie hebben gesteld. Uit een eerste, algemene verkenning van onderzoeksliteratuur die we eerder hebben uitgevoerd ten behoeve van de opzet van het onderzoek kwamen enkele aandachtspunten naar voren die we hebben gebruikt bij het uitwerken van de opzet. We vermelden die hier kort.

Datagebruik in scholen

Van Petegem en Van Hoof (2004) deden onderzoek naar het terugkoppelen van gegevens over leerprestaties naar scholen in het (Vlaamse) basisonderwijs. Zij concludeerden dat dit type data-feedback leidt tot benutting voor schoolontwikkeling als schoolleiders en leerkrachten:

- de gepresenteerde informatie als relevant percipiëren
- de gepresenteerde informatie correct kunnen interpreteren
- de manier van vergelijken (met andere scholen, met landelijk gemiddelde) onderschrijven
- de 'knowhow' hebben om de informatie te gebruiken
- in een niet-bedreigende context werk kunnen maken van schoolontwikkeling

Geijsel & Krüger (2005) hebben op 18 scholen voor basisonderwijs onderzocht hoe datafeedback moet worden georganiseerd. Het ging in dit geval over data over de innovatiecultuur in de school. Duidelijk werd uit hun evaluatie dat het zorgdragen voor relevantie van data een gezamenlijke zoektocht is van onderzoekers en de scholen zelf.

Ledoux et al (2005b, 2005c) constateerden op basis van een evaluatie van het Rotterdamse Onderwijskansenbeleid dat scholen het moeilijk vinden om zelfstandig verbeteractiviteiten te evalueren en informatie te verzamelen waarmee ze de voortgang van de verbeteringen kunnen monitoren. Hulp van onderzoekers bleek daarbij gewenst, op verschillende terreinen: het logisch ordenen van doelen en activiteiten, het kiezen van passende evaluatie-instrumenten, het stellen van goede onderzoeksvragen bij al be-

schikbare data (zoals scores op toetsen in het leerlingvolgsysteem), het interpreteren van gegevens en het vertalen van verworven inzichten naar beleidsconclusies.

Blok, Otter en Roeleveld (2001) deden enkele jaren geleden onderzoek naar het gebruik van leerlingvolgsystemen in het primair onderwijs. Ze betrokken daarbij ook de gegevens die in de periode 1994-2000 hierover waren verzameld in het PRIMA cohort onderzoek. Uit hun onderzoek blijkt dat het vastleggen van leervorderingen en het opslaan daarvan in een leerlingvolgsysteem zo langzamerhand een praktijk is op alle scholen in het primair onderwijs, althans voor de kernvakken lezen en rekenen in de groepen 3 t/m 8. Voor kleuters vindt opslag van gegevens in een centraal systeem lang niet altijd plaats, ook wordt in de kleutergroepen minder getoetst dan in de hogere groepen (zie ook Emmelot et al, 2006). Centrale opslag van gegevens komt verder wat minder vaak voor in het speciaal basisonderwijs.

Ten aanzien van het gebruik van deze gegevens op scholen kwamen uit dit onderzoek enkele belangrijke punten naar voren:

- de gegevens worden het meest gebruikt voor rapportages (aan ouders, bij overdracht van een klas) en minder voor diagnose of evaluatie
- een tweede belangrijke functie is het opsporen van leerlingen die extra zorg nodig hebben
- de gegevens worden het minst gebruikt voor evaluatie van prestaties op groepsniveau
- zowel leerkrachten als schoolleiders hechten weinig waarde aan vergelijkingen tussen groepen of vergelijkingen met andere scholen/het landelijk gemiddelde; dit wordt dan ook weinig gedaan. Dit laatste geldt voor het speciaal basisonderwijs nog sterker dan voor het regulier basisonderwijs
- de wijze waarop scholen de gegevens opslaan maakt het gebruik voor kwaliteitsbewaking moeilijk: er wordt vaak niet voldaan aan condities als het gebruik van toetsen met goede meettechnische eigenschappen, een gestandaardiseerde wijze van verzamelen en invoer in een centraal systeem met toegang tot statistische bewerking.

Deze onderzoeksresultaten geven aan dat het van belang is om in het onderzoek naar data driven teaching/opbrengstgericht werken aandacht te besteden aan:

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

- de vraag in hoeverre scholen in staat zijn op eigen kracht data te analyseren en te benutten voor schoolontwikkeling
- de attitude van schoolleiders en leerkrachten jegens data driven teaching, vooral wat betreft de reserves die men heeft tegenover vergelijkingen maken
- de technische en organisatorische aspecten van verzamelen en invoer van data. Het lijkt er immers op dat er wel veel gegevens beschikbaar zijn over leerprestaties, maar niet altijd van goede kwaliteit en niet altijd goed toegankelijk
- mogelijke verschillen tussen reguliere basisscholen en scholen voor speciaal (basis)onderwijs: het is aannemelijk dat in het speciaal (basis)onderwijs het draagvlak voor data driven teaching geringer zal zijn dan in het regulier basisonderwijs.

Systematische kwaliteitszorg

Tot nu toe zijn er in Nederland nog maar weinig scholen die op een systematische en cyclische wijze aan kwaliteitszorg doen en die dat ook consequent doen in functie van schoolontwikkeling (Inspectie van het Onderwijs, 2003, 2007; Hofman et al, 2004, 2005; Blok et al, 2006; Van Eck & Boogaard, 2007). Over de oorzaken daarvan is nog weinig bekend. We noemden al de bevindingen in het onderzoek naar het Onderwijskansenbeleid, die laten zien dat schoolleiders het moeilijk vinden om doelen, middelen en (verbeter)activiteiten goed te ordenen en om verbeterprocessen te monitoren (Ledoux et al, 2005a, 2005b). Men is nog weinig evaluatiegericht, maakt nog weinig gebruik van bestaande en beschikbare gegevens en beschikt ook niet over de vaardigheden om systematisch te evalueren. In ander onderzoek wordt dit bevestigd, ook voor scholen buiten Nederland (Earl & Fullan, 2003; Leithwood, Aitken & Jantzi, 2001). In dit kader wordt gesproken over de noodzaak tot 'capacity building': het ontwikkelen van het vermogen in scholen om te leren van data over de eigen school teneinde de data te benutten voor het interne veranderingsproces (Stoll, Bolam & Colarbone, 2002). Fullan (2007) spreekt over 'capacity building with a focus on results'. Volgens Earl & Katz (2006) vergt dit een mentale omslag (mind-shift) van schoollei-

ders en leraren. In plaats van zich te baseren op impliciete kennis (tacit knowledge) zullen zij moeten gaan werken met expliciete evidence-based kennis.

Van Eck & Boogaard (2007) hebben laten zien dat scholen die een compleet en cyclisch systeem van kwaliteitszorg hebben vaker dan andere scholen zelfgeïnitieerde vernieuwingen hebben, die bovendien kunnen bogen op een goed draagvlak bij leerkrachten. Verder melden zij dat scholen die verder zijn met kwaliteitszorg enthousiast zijn over de opbrengsten ervan: het geeft meer grip, structuur, betere onderlinge communicatie en betrokkenheid. Ook leidt het vaker tot afspraken die structurerend zijn voor het onderwijs, zoals afspraken over leerlijnen en (tussen)doelen. We nemen deze aandachtspunten mee in het onderhavige onderzoek.

Condities

Verschillende onderzoekers wijzen op het belang van een ‘onderzoekende cultuur’ in scholen als conditie voor onderwijsverbetering (Earl & Katz 2006; Verbiest 2003; Ponte, 2002). Earl & Katz (2006), die een omvattende studie naar datafeedback en leiderschap hebben uitgevoerd, wijzen er op dat het werken met data reflectie stimuleert over de wijze waarop onderwijsvernieuwing kan worden aangepakt, maar dat een voorwaarde hiervoor is dat schoolleiders en leerkrachten bereid zijn tot kritisch zelfonderzoek. Geijsel, Slegers, Stoel en Krüger (2009) hebben laten zien dat er een aantal condities zijn voor schoolontwikkeling, namelijk: leerkrachten die vertrouwen hebben in eigen kunnen en die de schooldoelen internaliseren, samenwerking tussen leerkrachten, participatie van leerkrachten in de besluitvorming rond de schoolontwikkeling (nodig voor draagvlak, eigenaarschap) en schoolleiders die in staat zijn leerkrachten te steunen en te stimuleren en die een eenduidige visie uitdragen. Verder is voor toepassen van datafeedback specifieke deskundigheid van belang: schoolleiders en leerkrachten moeten ‘data onderlegd’ zijn, dat wil zeggen data kunnen analyseren, verwerken, interpreteren en benutten (Earl & Katz, 2006).

Ook schoolextern zijn er belangrijke condities, zo blijkt uit verschillende publicaties: de rol van externe ondersteuning en van eventuele ‘critical friends’, de rol van het schoolbestuur of het bovenschools management (initiërend, steunend, belemmerend?) en de rol van de inspectie. Elk van deze actoren kan invloed uitoefenen op het werken

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

met datafeedback in scholen (Van Eck en Boogaard, 2007; Ledoux et al, 2005a, 2005b). Ook deze aandachtspunten nemen we mee in het onderhavige onderzoek.

1.3 Onderzoeksvragen

De vragen die het Ministerie met het onderzoek beantwoord wilde zien, zijn de volgende.

1. Wat is volgens bestaand onderzoek de beste werkwijze voor data driven teaching? Uit welke handelingen/stappen bestaat data driven teaching?
2. Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen in het primair onderwijs tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?
3. In hoeverre en hoe worden deze gegevens gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?
4. Welke handelingen/stappen van data driven teaching worden op scholen uitgevoerd, welke niet, en waarom?
5. Bestaat er draagvlak bij leerkrachten en schoolleiders voor data driven teaching?
6. Wat zijn risico- en succesfactoren bij data driven teaching? Zijn er eventueel onbedoelde effecten?
7. Wat hebben schoolleiders en leerkrachten nodig om data driven teaching te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

1.4 Opzet van het onderzoek

Voor het onderzoek is een opzet gekozen met de volgende deelonderzoeken:

1. Literatuurstudie
2. Interviews met deskundigen
3. Telefonische verkenning bij een representatieve steekproef scholen
4. Studie naar leerkrachten/scholen die betrokken zijn bij het Cito-pilotproject Hantering van leerstandaarden
5. Casestudies bij scholen die deelnemen aan de pilots Taalbeleid

Literatuurstudie

Het doel van de literatuurstudie is tweeledig geweest. In de eerste plaats is een begripsanalyse uitgevoerd van de term data driven teaching. Het is niet ongebruikelijk dat er rond de betekenis van een nieuwe term verwarring ontstaat. Het oogmerk van een begripsanalyse is het opsporen van interpretatieverschillen en het ontwikkelen van een omschrijving waarin zoveel mogelijk gebruikers van de term zich kunnen vinden. Daarbij hebben we ons laten leiden door de wens om voor het Nederlands taalgebied een toegankelijker term te kiezen dan data driven teaching. We hebben uiteindelijk de voorkeur gegeven aan de term *meetgestuurd onderwijs*. Meetgestuurd onderwijs hebben we omschreven als onderwijs waarbij betrokkenen (scholen, leraren, leerlingen) zich in hun taakuitvoering laten sturen door de uitkomsten van metingen. Meetgestuurd onderwijs ligt qua betekenis dicht bij het begrip ‘opbrengstgericht werken’, waarvoor het Ministerie gekozen heeft. In onze ogen verdient ‘meetgestuurd onderwijs’ echter de voorkeur, omdat dit begrip meer recht doet aan de betekenis van de term data driven teaching. Beide termen hebben namelijk gemeen dat ze de nadruk leggen op het gebruik van data of de uitkomsten van metingen. Opbrengstgericht werken heeft die connotatie niet, deze term mist een verwijzing naar systematische gegevensterugkoppeling of feedback. In het vervolg van het rapport zullen we vooral de term meetgestuurd onderwijs hanteren.

In de tweede plaats is via de literatuurstudie gezocht naar een (eerste) antwoord op de onderzoeksvragen 1, 2 en 3. Een bruikbaar model voor het beschrijven van de stappen (vraag 1) is gevonden in de reeds ingeburgerde *evaluatieve cyclus*. Verder is verkend

Opbrengricht werken; over de waarde van meetgestuurd onderwijs

wat in de onderzoeksliteratuur bekend is over effecten van meetgestuurd onderwijs; daarbij zijn enkele relevante onderscheidingen aangebracht (naar soorten betrokkenen en naar soorten evaluatie). Dit deel geeft tevens inzicht in mogelijke belemmeringen voor meetgestuurd onderwijs en eventuele onbedoelde effecten (vraag 6). Ten slotte is weergegeven wat in de onderzoeksliteratuur al bekend is over het gebruik van data over leervorderingen in scholen (vragen 2 en 3).

Zoals reeds vermeld zijn de resultaten van de literatuurstudie te vinden in hoofdstuk 2.

Interviews met deskundigen

Omdat we verwachtten dat de hoeveelheid empirisch-wetenschappelijk onderzoek naar meetgestuurd vrij gering zou zijn (zie Sleegers & Ledoux, 2006), hebben we als tweede stap in het onderzoek een aantal gesprekken met deskundigen opgenomen. Het doel hiervan was het verzamelen van praktijkkennis over toepassingen van meetgestuurd onderwijs in het primair onderwijs. Die praktijkkennis hebben we gezocht bij:

- a) *de inspectie van het onderwijs*: inspecteurs beoordelen alle scholen, onder meer op standaarden met betrekking tot opbrengsten en kwaliteitszorg, en hebben daardoor een goed zicht op zowel beschikbaarheid als (typen) gebruik van gegevens over leervorderingen in scholen
- b) *schoolbegeleiders* die ervaring hebben met toepassen van meetgestuurd onderwijs en scholen daarbij ondersteunen (bijvoorbeeld begeleiders die ervaring hebben met processen van schoolverbetering), en deskundigen op het gebied van kwaliteitszorg, schoolontwikkeling en het gebruik van ken- en stuurgetallen.
- c) *Cito-medewerkers* die ervaring hebben met het gebruik van het Cito-Leerlingvolgsysteem (een systeem met gestandaardiseerde toetsen voor het basisonderwijs) door scholen, en Cito-medewerkers die betrokken zijn bij het project hantering van leerstandaarden (zie verder bij deelonderzoek 5).

Met de inspectie en de Cito-medewerkers zijn twee afzonderlijke groeps gesprekken gevoerd. De schoolbegeleiders en overige deskundigen zijn individueel geïnterviewd,

Aanleiding, probleemstelling en onderzoeksopzet

telefonisch. Vooraf is een gespreksleidraad opgesteld en toegezonden aan de gesprekspartners. Alle onderzoeksvragen zijn daarin meegenomen.

Verder zijn als onderdeel van dit deelonderzoek ook enkele coördinatoren taal en rekenen van Pabo's geïnterviewd (individueel, telefonisch). Het doel hiervan was te achterhalen of/in welke mate Pabo's in hun opleidingsprogramma aandacht besteden aan (elementen van) meetgestuurd onderwijs, zoals het leren afnemen van toetsen, analyseren van toetsresultaten, afstemmen van het onderwijsaanbod op deze resultaten en inzicht in processen van kwaliteitszorg.

Alle resultaten van dit deelonderzoek staan in hoofdstuk 3.

Telefonische interviews bij scholen

Om een beeld te krijgen van de huidige praktijk in scholen voor primair onderwijs wat betreft meetgestuurd onderwijs (werkwijze, condities) is een telefonische vragenlijst afgenomen bij een representatieve steekproef scholen. Deze vragenlijst bestond uit twee delen. Het eerste deel was kort (slechts enkele vragen) en bedoeld om een overzicht te krijgen van de mate waarin scholen voor primair onderwijs bezig zijn met (onderdelen van) meetgestuurd onderwijs. Voor deze vorm (telefonische afname, korte vragenlijst met een afnameduur van circa 10 minuten) is gekozen om een hoge respons te kunnen halen zonder veelvuldig rappelleren. Responspercentages bij schriftelijke surveys liggen tegenwoordig vrij laag (rond 30%) en we verwachtten dat bij het nog vrij onbekende onderwerp 'datafeedback' een schriftelijke vragenlijst vaak terzijde zou worden gelegd. Vanwege de wens om een representatief beeld te krijgen was het streven om een respons te halen van 70% of hoger.

We hebben 350 scholen benaderd: 200 reguliere basisscholen, 100 scholen die 'taalzwak' zijn volgens de inspectie en 50 scholen voor speciaal basisonderwijs. De taalzwakke scholen zijn opgenomen als aparte subgroep omdat het denkbaar is dat scholen met relatief zwakke scores voor de basisvaardigheden meer of eerder dan andere scholen aandacht zullen besteden aan meetgestuurd onderwijs, vanwege de noodzaak om de prestaties te verbeteren. Voor taalzwakke scholen is gekozen omdat op het moment van het onderzoek voor deze groep een bestand beschikbaar was bij de inspectie (voor rekenzwakke scholen was dat nog niet het geval).

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Alle scholen zijn aselekt gekozen uit hun subpopulatie. De overall respons is 74% en het beoogde responspercentage is dus gehaald.

Het tweede deel van de vragenlijst was uitgebreider en bedoeld om vervolgvragen te stellen aan scholen die in de eerste telefonische ronde gemeld hadden relatief actief te zijn op het gebied van meetgestuurd onderwijs. Deze vragenlijst is afgenomen bij 20 scholen, proportioneel verdeeld over de drie substeekproeven, eveneens telefonisch. In alle gevallen waren directeuren of intern begeleiders de respondent.

Ook in dit deelonderzoek zijn alle onderzoeksvragen meegenomen, met uitzondering van vraag 1. Het accent lag op de vragen 3, 4 en 5. De resultaten staan in hoofdstuk 4.

Studie naar leerkrachten/scholen die betrokken zijn bij het Cito-pilotproject Hantering van leerstandaarden

In de beleidsagenda is een belangrijke plaats ingeruimd voor de ontwikkeling van leerstandaarden. Verwacht worden dat deze structurend kunnen werken voor acties gericht op prestatieverbetering. Er is een expertgroep Doorlopende leerlijnen taal en rekenvaardigheden in het leven geroepen die een advies heeft uitgebracht over referentieniveaus voor taal en rekenen (Expertgroep Doorlopende Leerlijnen Taal en Rekenen, 2008). Hiermee wordt voortgebouwd op het advies 'Zeker Weten' van de Onderwijsraad (1999). Verder heeft het Cito de opdracht gekregen om proefreferentieniveaus taal en rekenen te ontwikkelen voor groep 5 en 8, dat wil zeggen halverwege en aan het eind van het primair onderwijs. Gestart is met groep 5, in een pilotproject Hantering van leerstandaarden. Als eerste stap zijn in dit pilotproject uit de toetsen van het Cito LVS opgaven geselecteerd die als referentieniveau kunnen dienen voor leerlingen in groep 5. Aan een aantal leerkrachten uit basisscholen in de regio's Arnhem en Den Haag is vervolgens gevraagd om per opgave te beoordelen of leerlingen in groep 5 dit moeten beheersen, volgens twee niveaustandaarden: een minimumniveau (= ca. 95% van de leerlingen zou dit moeten kunnen) en een voldoende niveau (= ca. 75% van de leerlingen zou dit moeten kunnen). Vervolgens kregen leerkrachten te zien hoeveel leerlingen (in een landelijke normgroep) die opgaven werkelijk goed maken, op basis daarvan konden ze hun oordeel eventueel nog bijstellen. Dit heeft uiteindelijk geresulteerd in referentieniveaus voor groep 5 voor de domeinen

technisch lezen, begrijpend lezen en rekenen. Het Cito publiceert deze in een brochure.

Het was de wens van de opdrachtgever om voor het onderzoek naar meetgestuurd onderwijs ook scholen te benaderen die meededen aan dit pilotproject. In de vorm van casestudies zou bestudeerd moeten worden wat de ervaringen van de scholen waren met het werken met (proef)referentieniveaus. Door wijzigingen in de opzet van het pilotproject van het Cito werden echter geen hele scholen in het project betrokken, maar alleen individuele leerkrachten. Na overleg is daarom besloten om voor de vragen rond referentieniveaus geen casestudies uit te voeren, maar in plaats daarvan een schriftelijke vragenlijst af te nemen bij de leerkrachten die deelnamen aan het Cito project, na afloop van hun werkzaamheden in het project (juni 2008). Daarin is gevraagd naar hun ervaringen in het project en naar hun opvattingen over de zinvolheid van referentieniveaus en het gebruik daarvan in de eigen praktijk.

Van dit deelonderzoek wordt verslag gedaan in hoofdstuk 5.

Casestudies pilots taalbeleid

Een andere wens van de opdrachtgever was om bij het onderzoek ook scholen te betrekken die meedoen aan de pilots taalbeleid. De pilots taalbeleid worden uitgevoerd in een aantal gemeenten en regio's door scholen die een hoog percentage leerlingen uit achterstandsgroepen hebben (gewichtleerlingen) en voor wie het dus van belang is om extra aandacht te besteden aan het taalonderwijs. Het gaat in totaal om circa 250 scholen. De looptijd van de pilots is van 2007 tot 2010. Per gemeente/regio zijn afspraken gemaakt over de activiteiten die de scholen in deze periode uitvoeren. Deze kunnen verschillen, uitgangspunt is wat voor de betreffende school nodig is. De scholen die meedoen aan de pilots worden gevolgd in een monitor die wordt uitgevoerd door de Hogeschool Utrecht. De monitor is recent van start gegaan en omvat onder meer verzameling van toetsgegevens op de betreffende scholen (technisch lezen, begrijpend lezen en woordenschat) en een vragenlijst voor interne begeleiders en schoolleiders. Daarin wordt onder meer gevraagd hoe zij omgaan met toetsgegevens. Verder zijn van alle scholen de plannen voor de projectperiode beschikbaar; deze worden voor de monitor ook geanalyseerd. De monitor zal daarmee in de breedte een beeld geven van (onder andere) de wijze waarop de scholen leervorderingen meten en

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

gebruik maken van prestatiegegevens om hun onderwijs te sturen. Voor het onderzoek naar meetgestuurd onderwijs hebben we er voor gekozen om deze informatie aan te vullen met een meer diepgaand beeld van enkele scholen die een werkwijze hebben die correspondeert met kenmerken van meetgestuurd onderwijs. Hiervoor zijn vijf scholen geselecteerd. De selectie is gedaan door de landelijke coördinator van de pilots taalbeleid, op basis van bij haar beschikbare informatie over de werkwijze van de scholen. De scholen zijn vervolgens door de onderzoekers benaderd. Bij de selectie is gezorgd voor een spreiding naar percentage gewichtenleerlingen en naar type gewichtenleerlingen (allochtoon/autochtoon). Ook is één casestudie uitgevoerd in een school voor speciaal (basis)onderwijs.

In de casestudies is het accent gelegd op het voeren van gesprekken met leerkrachten. Leerkrachten zijn de sleutelpersonen vanuit het perspectief van de doelstellingen van de beleidsagenda: zij moeten immers zorgen voor de gewenste prestatieverbetering. In de casestudies was het mogelijk leerkrachten een centrale plaats te geven. Per case (school) zijn drie gesprekken gevoerd:

- twee groepsgesprekken met leerkrachten, verdeeld over kleuterbouw-
onderbouw en middenbouw-bovenbouw
- een groepsgesprek met de schoolleider, interne begeleider(s) en taalcoördinator

Verder zijn documenten opgevraagd en bestudeerd over de werkwijze van de school. Onderwerpen in de gesprekken waren de toepassing van de verschillende stappen bij data driven teaching, draagvlak, opbrengsten, succes- en risicofactoren en vereiste kwaliteiten en houdingen bij leerkrachten, schoolleiders en intern begeleiders. Ook is gevraagd naar de rol van het bestuur/bovenschools management. Het accent lag op beantwoording van de onderzoeksvragen 4, 5, 6 en 7.

De resultaten van dit deelonderzoek zijn te vinden in hoofdstuk 6.

2 Van 'data-driven teaching' naar meetgestuurd onderwijs; een verkenning van de literatuur

2.1 Probleemstelling

Op de beleidsagenda voor het primair onderwijs – Scholen voor morgen (Ministerie van OCW, 2007) – neemt het verbeteren van taal- en rekenprestaties van leerlingen een prominente plaats in. Goede taal- en rekenvaardigheden zijn, zo wordt in de nota gesteld, onmisbaar voor het succes in andere vakken. En ook, wil Nederland mee blijven doen in de mondiale kenniseconomie, dan moet het taal- en rekenniveau omhoog.

In de genoemde nota worden vijf sporen of wegen genoemd waarlangs de komende jaren gewerkt zal worden. In het tweede spoor ('zicht op succes') wordt uitgewerkt dat leraren en scholen beter gebruik gaan maken van informatie over de onderwijskwaliteit. Het gaat hier om een systematische benutting van toetsgegevens. Dit wordt in het offerteverzoek aangeduid als *data-driven teaching*. Rond dit verschijnsel stellen we in dit hoofdstuk drie onderwerpen aan de orde.

Het eerste onderwerp is de betekenis van *data-driven teaching*. Het is een nieuwe term en over nieuwe termen zijn nuttige vragen te stellen (De Groot & Medendorp, 1986). Een belangrijke vraag is bijvoorbeeld of *data-driven teaching* alleen maar een nieuwe naam is voor een al langer bestaande werkwijze, of dat er sprake is van een nieuw begrip, verwijzend naar een werkwijze die nog niet eerder benoemd is. Anders gezegd: is alleen de term nieuw, of ook de werkwijze waarnaar de term verwijst? Een andere vraag is of gebruikers van de nieuwe term er allemaal hetzelfde mee bedoelen. Het is, zeker voor een nieuwe term, niet ongewoon dat er onduidelijkheid bestaat over de betekenis ervan. Die onduidelijkheid kan leiden tot een verschillende interpretatie. Oppervlakkig gezien lijken gebruikers het met elkaar eens te zijn, maar bij nadere beschouwing bedoelen ze verschillende dingen.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Een ander onderwerp is welke effecten men mag verwachten van *data-driven teaching*. Valt van deze werkwijze inderdaad een prestatieverhoging te verwachten, en zo ja, onder welke voorwaarden? En niet onbelangrijk, moet men bedacht zijn op nadelige neveneffecten? Zulke vragen kunnen beantwoord worden op basis van een inventarisatie van ervaringen die elders al zijn opgedaan. Zo'n inventarisatie kan helpen om valkuilen te vermijden en bestaande verwachtingen van een realistische basis te voorzien.

Een derde onderwerp is hoe *data-driven teaching* zich verhoudt tot de gangbare werkwijzen in het basisonderwijs. Over welke vorderingengegevens beschikken scholen en hoe gebruiken ze deze? Inzicht in gangbare werkwijzen stelt ons in staat een beoordeling te geven in welke opzichten *data-driven teaching* inderdaad een andere of vernieuwende werkwijze is. Op basis daarvan kan een indruk gegeven worden van de aard en omvang van de van scholen gevraagde veranderingen.

In dit hoofdstuk gaan we op de aangeduide onderwerpen nader in. We doen dat aan de hand van de volgende onderzoeksvragen.

- a. Wat valt er te verstaan onder *data-driven teaching*?
- b. Welke effecten worden er in de onderzoeksliteratuur van *data-driven teaching* beschreven?
- c. Over welke vorderingengegevens beschikken scholen voor basisonderwijs?
- d. Hoe maken scholen voor basisonderwijs gebruik van de beschikbare vorderingengegevens?
- e. Welke belemmeringen zijn er bij de invoering van *data-driven teaching*?

Deze onderzoeksvragen zijn beantwoord op basis van een literatuuronderzoek. We hebben een *quick scan* uitgevoerd van de internationale onderzoeksliteratuur aan de hand van trefwoorden als *data-driven teaching*, *leerlingvolgsysteem*, *assessment for learning* en verwante termen. Gebruik is gemaakt van onder andere ERIC, ISI Web of Knowledge, Picarta en zoekmachines op internet. Aanvullend trokken we verwijzingen na in de literatuurlijsten van relevante publicaties (de zogenoemde sneeuwbalmethode). Gezien de korte duur van de literatuurstudie hebben we ons beperkt tot overzichtsartikelen en enkele andere kernpublicaties.

2.2 Begripsanalyse

Het offertezoek van het Ministerie van OCW, dat we als startpunt nemen voor onze analyse, biedt een eerste context voor wat *data-driven teaching* inhoudt. In die context komen drie elementen naar voren: het gaat om het gebruik a. van objectieve gegevens over leervorderingen, b. door leraren, c. met het oog op verbeterde leerprestaties. De veronderstelling is dat leerlingen meer of sneller leren als hun leraren duidelijke doelen stellen en regelmatig via objectieve toetsen bepalen of de doelen zijn bereikt. Indien dat niet het geval blijkt, kunnen leraren vervolgens achterhalen waarom de doelen niet zijn bereikt en hun vervolghandelingen afstemmen op het alsnog halen van de doelen.

In thesauri op het terrein van de onderwijskunde komt *data-driven teaching* als term niet voor. We kunnen daaruit afleiden dat het geen aanvaarde vakterm is met een vastgelegde betekenis. De term wordt in de literatuur bovendien nauwelijks gebruikt. We vonden slechts een brochure van het Amerikaanse *Rennie Center for Education Research & Policy*. De titel luidt: 'Data-driven teaching: Tools and trends' (Rennie Center, 2006). In de brochure wordt een pleidooi gehouden voor formatieve toetsen ('formative assessments'), die – in tegenstelling tot summatieve toetsen – leraren helpen de vorderingen van hun leerlingen te monitoren met het oog op het instructieproces.

Ook al is *data-driven teaching* als term vrijwel niet in gebruik, de ermee aangeduide werkwijze is wijd en zijd bekend, zij het onder andere benamingen. We geven een kort overzicht van vaak gebruikte termen. *Diagnostic teaching* is een erkende vakterm, als volgt omschreven in de thesaurus van ERIC: 'Process of diagnosing student abilities, needs, and objectives and prescribing requisite learning activities'. *Diagnostic teaching* is aan de orde in situaties waarin het onderwijs wordt afgestemd op een individuele leerling, meestal een leerling die extra zorg of aandacht behoeft. Gebeurt de afstemming minder nauwkeurig, niet op individuele leerlingen maar op groepen van leerlingen, dan is de Nederlandse term *adaptief onderwijs* te gebruiken. Voor deze term stelden wij zelf enkele jaren geleden de volgende omschrijving voor: adaptief onderwijs is het doelbewust afstemmen van de onderwijsleersituatie op verschillen tussen leerlingen in dezelfde leergroep (Blok & Breetvelt, 2002, p. 18). In deze om-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

schrijving is niet expliciet vastgelegd hoe de verschillen tussen leerlingen zouden moeten worden bepaald. Maar de veronderstelling ligt hier voor de hand dat leraren zich daarbij bedienen van een geëigend instrumentarium om de leervorderingen te bepalen.

Data-based decision making is een term die we vonden in Ingram, Louis en Schroeder (2004) en die afgewisseld wordt met *teacher decision making*. Het artikel gaat over het gebruik dat leraren van vorderingstoetsen maken om hun onderwijs te kunnen verbeteren. De termen worden in het artikel niet expliciet omschreven, maar de context maakt veel duidelijk. Het gaat erom dat leraren gebruik maken van 'standardized tests and other sources ... to make decisions about school and classroom practice' (p. 1258). Beide termen liggen qua betekenis heel dicht bij *data-driven teaching* zoals beschreven in het eerder genoemde offerteverzoek. In beide gevallen gaat het om leraren die vorderingstoetsen gebruiken om hun onderwijs te verbeteren. Een verschil is er wel wat betreft het type toetsen. In het offerteverzoek is sprake van objectieve toetsen, Ingram e.a. hebben het over 'standardized tests and other sources'.

Met dit verschil raken we ons inziens aan een kernpunt. Leraren behouden graag zeggenschap over de manier waarop ze de vorderingen van hun leerlingen bepalen. Zo is van Nederlandse basisschoollerares bekend dat ze bij het bepalen van de vorderingen veelvuldig gebruik maken van schriftelijk werk dat leerlingen tijdens de lessen produceren (Blok, Otter & Roeleveld, 2001). Ook nemen ze regelmatig methodegebonden toetsen af. Zulke methoden lenen zich goed voor een directe terugkoppeling: zijn de doelen van de les (of van de lessenserie) bereikt, wie is nog niet ver genoeg, welke onderdelen verdienen nog extra aandacht? Maar beoordeling van schriftelijk werk en methodegebonden toetsen zijn geen bruikbare basis voor een vergelijking van grotere groepen leerlingen, bijvoorbeeld leerlingen van verschillende scholen (een scholenvergelijking) of leerlingen van verschillende leerjaren (een cohortvergelijking). En het zijn juist deze vergelijkingen die voor managers en beleidsambtenaren boeiend zijn. Het kernpunt is dus: wat is eigenlijk het type beslissingen dat genomen moet worden? Gaat het om leraren die hun eigen onderwijs willen optimaliseren, of om managers die mogelijkheden willen onderzoeken het onderwijsrendement in hun eigen 'eenheid' te verhogen, of zijn beide opties relevant?

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

Deze vraag heeft veel te maken met het bekende onderscheid tussen formatieve en summatieve evaluatie of toetsing. Formatieve evaluatie is evaluatie die wordt ondernomen met het oog op feedback en bijsturing. Summatieve evaluatie vindt achteraf plaats en is gericht op het vaststellen van de effectiviteit. Formatieve evaluatie is met name voor leraren van belang, summatieve evaluatie raakt de verantwoordelijkheid van leraren en managers samen. Voor goed onderwijs zijn vanzelfsprekend beide evaluatievormen van belang. Daarom lijkt het ons onverstandig een term te kiezen die ofwel naar de formatieve functie, ofwel naar de summatieve functie verwijst. In de rest van het rapport zullen we beide functies aan bod laten komen, zoals overigens ook de bedoeling is van de opdrachtgever van het onderzoek, het Ministerie van OCW.

Wij achten *data-driven teaching* voor het vervolg van deze tekst geen geschikte term. Het is allereerst een Engelse term, een term die nauwelijks gebruikers kent. De term is bovendien – via 'teaching' – exclusief verbonden aan leraartaken, en niet expliciet ook aan taken en verantwoordelijkheden van managers. Wij geven de voorkeur aan een andere term, namelijk '*meetgestuurd onderwijs*':

Meetgestuurd onderwijs is onderwijs waarbij betrokkenen zich in hun taakuitvoering laten sturen door uitkomsten van metingen.

We voorzien deze omschrijving van enkele annotaties.

De eerste kanttekening is een toelichting op de term *betrokkenen*. We maken hier een ideaaltypisch onderscheid tussen drie partijen, die elk een specifieke positie hebben: de leerling voor wie het onderwijs bedoeld is, de leraar die het onderwijs geeft en de manager die het onderwijs organiseert. Voor deze drie partijen heeft de term *meetgestuurd onderwijs* een specifieke invulling, afhankelijk van de taken die men vervult. De leerling leert en kan zich in zijn leren laten sturen of bijsturen door de uitkomst van een vorderingenmeting. Hij kan bijvoorbeeld besluiten bepaalde oefeningen nog eens over te doen of om extra uitleg te vragen. De leraar onderwijst en kan zich bij de uitoefening van deze taak eveneens laten sturen door de uitkomst van metingen. Hij kan bijvoorbeeld besluiten bepaalde leerlingen een extra opdracht te geven. De mana-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

ger organiseert het onderwijs en kan op grond van metingen bijvoorbeeld tot het oordeel komen dat de school zich op een andere rekenmethode moet oriënteren.

De tweede kanttekening betreft de term *metingen*. De eerste gedachte gaat natuurlijk uit naar metingen van een specifiek type, de vorderingenmetingen. Wij vermijden hier expliciet de term toetsen, een term die vaak gekoppeld wordt aan objectieve toetsen. We wensen ruimte te laten aan andere vormen van vorderingenmetingen, zoals beoordelingen van schriftelijk werk, mondelinge overhoringen, gedragsobservaties, portfolio's en methodegebonden toetsen. Het zijn vaak deze – zachtere – metingen die van nut zijn voor leerlingen en leraren. Maar naast vorderingenmetingen kunnen de gedachten ook uitgaan naar andere metingen die voor de taakuitvoering relevant kunnen zijn: metingen van de doorstroom naar het voortgezet onderwijs, van het welbevinden van leerlingen, van de tevredenheid van ouders, van de satisfactie van leraren, ja zelfs van het CO₂-niveau in klaslokalen. Van de laatste is namelijk onlangs vastgesteld dat een te hoog gehalte nadelig is voor leerprestaties (Meijer, Hasselaar & Snepvangers, 2007).

Onze derde kanttekening is misschien niet nodig, maar zekerheidshalve maken we hem toch. Meetgestuurd onderwijs impliceert niet dat het onderwijs alléén door uitkomsten van metingen wordt gestuurd. Onderwijs wordt gestuurd door doelstellingen die men wil bereiken, en natuurlijk ook door de faciliteiten waarover men beschikt. Sturing door meetuitkomsten kan alleen bijsturing zijn. *Door-meetuitkomsten-bijgestuurd onderwijs* zou wellicht een meer precieze term zijn, maar overigens weinig praktisch in het gebruik.

Tot slot, *meetgestuurd* is zoals veel adjectieven gradeerbaar. We kunnen dus spreken over onderwijs dat in meer of mindere mate meetgestuurd is. Het is geen kwestie van alles of niets, van wel of niet meetgestuurd onderwijs. De werkelijkheid is natuurlijk veel genuanceerder. Scholen verschillen in de mate waarin ze meetgestuurd onderwijs verzorgen, zoals we in hoofdstuk 4 zullen laten zien.

2.3 De evaluatieve cyclus

In deze en volgende paragrafen beperken we ons tot een specifiek type van metingen, de vorderingenmetingen. De reden daarvoor is vastgelegd in het offerteverzoek en het daarbij ontwikkelde onderzoeksplan. De belangstelling van het Ministerie van OCW gaat met name uit naar een systematische benutting van gegevens over leerprestaties van leerlingen.

Het gebruik van vorderingengegevens is onlosmakelijk verbonden met de evaluatieve cyclus. Deze cyclus bestaat uit vijf stappen: het vastleggen van doelen en standaarden, het verzamelen van informatie, het registreren ervan, het interpreteren en het nemen van beslissingen. We vinden het belangrijk aan de cyclus aparte aandacht te besteden, omdat deze in de praktijk vaak gebrekkig functioneert. Een veel voorkomend probleem is dat de cyclus niet in zijn geheel wordt uitgevoerd, maar alleen de tweede en de derde stap. We beschreven de evaluatieve cyclus al eens eerder (Blok, Otter & Roeleveld, 2002), overigens in een iets gewijzigde vorm, maar ditmaal gaan we explicieter in op de taken en verantwoordelijkheden van de drie partijen die we in het voorafgaande hebben onderscheiden: de leraren en de managers, maar óók de leerlingen. In de praktijk van meetgestuurd onderwijs gaat de meeste aandacht uit naar leraren en managers. Maar ook leerlingen hebben bij metingen van hun vorderingen eigen taken en verantwoordelijkheden, zoals we zullen laten zien.

Het vastleggen van doelen en standaarden

De evaluatieve cyclus start idealiter met het vastleggen van de leerdoelen en de daarbij te hanteren standaarden. Wat zijn de doelen die men wil evalueren en welke standaard of standaarden horen daarbij om te beoordelen in welke mate de doelen zijn bereikt? Deze vragen zijn soms niet zo gemakkelijk te beantwoorden, omdat doelen in het onderwijs nogal eens worden geformuleerd in termen van onderwijsaanbod, te onderscheiden van doelen in termen van beheersing door leerlingen. Omdat het hier om vorderingenmetingen gaat, verdient de laatste vorm de voorkeur. Een standaard geeft het streefniveau aan, het niveau dat men bij leerlingen wil bereiken. Men kan één standaard hanteren, die dan noodzakelijkerwijs samenvalt met de grens tussen voldoende en onvoldoende. Men kan ook meer standaarden vaststellen, rekening

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

houdend met niveauverschillen tussen leerlingen. Zo hanteert het project Periodieke Peiling van het Onderwijsniveau (PPON) twee standaarden, de standaard 'minimum' (te bereiken door 90 tot 95 procent van de leerlingen) en de standaard 'voldoende' (te bereiken door 70 tot 75 procent van de leerlingen).

Het vastleggen van doelen en standaarden is in het basisonderwijs typisch een gemeenschappelijke verantwoordelijkheid van leraren en managers, maar niet van leerlingen.

Het verzamelen van informatie

Het verzamelen van informatie over het vorderingsniveau van een leerling is in feite een reductieproces. Het gaat erom het niveau van een leerling te vertalen in een toetscore (bijvoorbeeld het aantal goed of fout), een schoolcijfer of een kwalitatief oordeel. Het verzamelen kan op heel verschillende manieren gebeuren. De leraar kan bijvoorbeeld gebruik maken van gedragsobservaties (zoals het geven van leesbeurten), productobservaties (zoals het laten schrijven van een opstel) of toetsen. Welke techniek een leraar kiest, hangt sterk samen met het domein en de leerdoelen die aan de orde zijn. Voor rekenen/wiskunde bijvoorbeeld zijn heel andere technieken geëigend dan voor schrijven, bewegingsonderwijs of kunstzinnige oriëntatie.

Het is belangrijk te beseffen dat de technieken om informatie te verzamelen alle hun meettechnische tekortkomingen hebben. Zo is het oordeel van een leraar over een leesbeurt of over een opstel gevoelig voor een persoonlijke vertekening die niets met het niveau van de leerling te maken heeft. Toetsen hebben andere beperkingen. Vaak meten ze slechts een selectief deel van een domein. Een spellingtoets bijvoorbeeld meet alleen of een leerling van gegeven woorden de correcte schrijfwijze kent. Maar waar het bij het spellingonderwijs om draait, gaat net iets verder. Het gaat erom dat leerlingen leren om eigen teksten zonder fouten te schrijven. Daarbij hoort dat leerlingen bereid en in staat zijn een eigen tekst op fouten te controleren.

De kwaliteit van de evaluatieve cyclus hangt rechtstreeks af van de kwaliteit van de verzamelde informatie. Vage of onbetrouwbare informatie kan aanleiding zijn tot verkeerde beslissingen. Het is daarom van belang dat de verzamelde informatie van goede kwaliteit is. Mogelijkheden om de kwaliteit van vorderingsgegevens te verbe-

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

teren zijn onder meer het regelmatig verzamelen ervan – toevallige meetfouten middelen elkaar immers uit – en het verzamelen van informatie langs meerdere wegen, bijvoorbeeld door naast het lerarenoordeel ook toetsen af te nemen.

Bij het verzamelen van informatie zijn in hoofdzaak twee partijen betrokken: de leerlingen en de leraren. De leraar is verantwoordelijk voor alle organisatorische aspecten van de gegevensverzameling, waaronder de keuze van het meetinstrument en van het moment van de meting. De leerlingen hebben de taak aan de meting mee te werken en de verantwoordelijkheid zich maximaal in te spannen, opdat de uitkomst van de meting geen onderschatting vormt van hun vorderingen.

Het registreren van vorderingen

Het eindresultaat van de vorige stap moet vastgelegd worden, op papier of met een computer. Vastleggen met een computer levert extra werk op, maar biedt ook extra mogelijkheden. Welke mogelijkheden dat zijn, hangt af van de gebruikte software. Een nuttig voordeel is dat een eenmaal ingevoerde score op uiteenlopende manieren weer uitgevoerd of gerapporteerd kan worden. Zo maakt de computer in een handomdraai zowel groepsoverzichten als individuele overzichten. Traditioneel moesten leraren kiezen tussen een administratie op leerlingniveau (per leerling een kaart of een bladzij) of een administratie op groepsniveau (per toets een kaart of een bladzij). Wie op beide niveaus het overzicht wilde bewaren, moest twee administraties voeren en elk gegeven tweemaal invoeren.

Een ander voordeel dat computerprogramma's kunnen bieden, is dat ze de interpretatie van de vastgestelde vorderingen ondersteunen. Moderne programma's kunnen ruwe toetsscores (aantal goed of fout) omzetten in normscores. Deze normscores brengen tot uitdrukking hoe ver een leerling gevorderd is ten opzichte van leeftijdgenoten.

In de tijd dat vorderingen alleen nog op papier werden vastgelegd, was de registratie ervan de uitsluitende taak van leraren. Met de komst van de computer is dat veranderd. Steeds vaker is registratie mede een managementtaak geworden. Op veel scholen behoort de opbouw en het beheer van het vorderingenregistratiesysteem tot het takenpakket van de interne begeleider. Hoe de taken ook verdeeld zijn, het is belangrijk dat er afspraken zijn over de toegankelijkheid van de vorderingengegevens in het

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

systeem. Leerlingen moeten worden geïnformeerd over hun vorderingen en managers moeten eveneens toegang krijgen in verband met hun beheers- en beleidstaken.

Het interpreteren van vorderingen

Geregistreerde leervorderingen krijgen in de evaluatiecyclus pas hun waarde als ze naast een standaard worden gelegd. Die standaard verschaft de basis voor een oordeel in termen van meer of minder voldoende. Op theoretische gronden wordt vaak een onderscheid gemaakt tussen twee soorten standaarden. We gebruiken een leerstofgerichte standaard als we willen beoordelen in hoeverre de leerling de leerstof onder de knie heeft. Van een normgerichte standaard is sprake als we het vorderingenniveau relateren aan het niveau van leeftijdgenoten.

Dit theoretische onderscheid is in de praktijk niet altijd herkenbaar. Het oordeel van een leraar over een leesbeurt of over een opstel is een gegeven waarin impliciet een standaard is verwerkt. Het schoolcijfer 5 staat nu eenmaal gelijk aan onvoldoende. Volgens welke standaard een schoolcijfer tot stand is gekomen, verschilt per leraar en is vaak niet precies te reconstrueren. Sommige leraren hechten in hun oordeel een groot gewicht aan de inzet van de leerling. Een leerling die erg zijn best heeft gedaan, kan dan soms bijna alleen al op grond hiervan een voldoende krijgen. Andere leraren ruimen in hun standaard vooral plaats in voor het niveau waarop de leerstof wordt beheerst. Bij zo'n leraar moet een leerling daadwerkelijk vorderingen boeken om een voldoende te krijgen.

In feite speelt het onderscheid tussen een leerstofgerichte en een normgerichte interpretatie vooral een rol bij toetsscores. Methodegebonden toetsen hanteren vaak een leerstofgebonden standaard, terwijl methodeonafhankelijke toetsen een normgerichte standaard aanbieden. Voorbeelden van scores die een normgerichte interpretatie impliceren, zijn de niveauscores A tot en met E zoals gebruikt in het Cito-leerlingvolgsysteem en de dle-systematiek (het didactische leeftijdsequivalent) die zijn oorsprong vindt bij het onderwijs in aanvankelijk lezen/spellen, maar die later ruimere verspreiding heeft gekregen.

Bij het interpreteren van de vorderingen zijn alle drie de partijen betrokken. Leraren kiezen een standaard en beoordelen aan de hand daarvan de vorderingen van afzon-

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

derlijke leerlingen en van de klas als geheel. Ook managers kiezen een standaard – het kan een andere standaard zijn dan die van de leraar – en vormen zich een oordeel over de vorderingen van met name groepen van leerlingen. Leerlingen vormen zich over de eigen vorderingen een oordeel. Dit oordeel vormt de basis voor het beeld dat ze opbouwen van de eigen competentie. De betrokkenheid van leerlingen bij deze stap van de evaluatieve cyclus is misschien wat minder zichtbaar, maar theoretisch gezien niet onbelangrijk.

Het nemen van beslissingen

De evaluatiecyclus is pas compleet als er ook beslissingen worden verbonden aan de geïnterpreteerde vorderingengegevens. Lang niet altijd gaat het om zichtbare beslissingen. Immers, als de voorgaande stap uitwijst dat het geregistreerde vorderingenniveau naar tevredenheid is – anders gezegd, voldoet aan de eerder vastgelegde standaard(en) – dan is er geen reden om bij te sturen. Maar in feite ligt aan het voortgaan op de ingeslagen weg natuurlijk ook een beslissing ten grondslag.

Beslissingen kunnen op meerdere niveaus genomen worden, het niveau van de leerling, de klas, de school of zelfs op bovenschools niveau. Leraren nemen met name beslissingen op de eerste twee niveaus, vanuit hun taak het leerproces te ondersteunen. Ze kunnen bijvoorbeeld besluiten een of meer leerlingen, of ook de hele groep, extra oefening of leertijd te bieden. Leerlingen beperken zich in hun beslissingen tot het eigen niveau. Ze kunnen onder andere beslissen hun inzet te veranderen of te vragen om een aangepast programma, zoals een herhalingsprogramma of een versneld programma. Managers nemen in de praktijk hun beslissingen vooral op het groeps-, het school- en eventueel het bovenschoolse niveau. Een voorbeeld is de beslissing een studiemiddag te organiseren om de vorderingen in een breder verband te bespreken.

2.4 Effecten van meetgestuurd onderwijs

Net als in de vorige paragraaf perken we de bespreking ook hier weer in tot een specifieke vorm van meetgestuurd onderwijs, namelijk sturing op basis van vorderingen van leerlingen.

De vraag naar de effecten van meetgestuurd onderwijs is niet gemakkelijk te beantwoorden, al kunnen we vooraf al wel bepaalde verwachtingen formuleren. Zo zijn er goede redenen om van meetgestuurd onderwijs een positief effect te verwachten. Immers, uitkomsten van vorderingenmetingen lijken voor alle drie de partijen – leerlingen, leraren en managers – buitengewoon nuttig als feedback. En van feedback is bekend dat het een effectieve bijdrage kan leveren aan het resultaat, hoe dan ook gedefinieerd.

Maar een positief effect veronderstelt dat aan zekere voorwaarden is voldaan. Zo moeten de meetuitkomsten valide zijn voor de nagestreefde leerdoelen. Met deze eis stuiten we al meteen op een serieus probleem. Voor lang niet alle kerndoelen van het basisonderwijs zijn valide meetprocedures beschikbaar, zelfs niet als we ons beperken tot de domeinen Nederlandse taal en rekenen/wiskunde. Een andere voorwaarde voor een positief effect van meetgestuurd onderwijs is dat de gehanteerde maatstaven, de criteria om tot een oordeel te kunnen komen, realistisch zijn. En ook hier ligt een probleem. Is het bijvoorbeeld wel realistisch te eisen dat alle leerlingen aan het eind van groep 3 tenminste AVI-niveau 3 hebben gerealiseerd? In deze eis is niet verdisconteerd dat sommige leerlingen leerproblemen hebben en dat andere leerlingen na één jaar lesonderwijs veel verder zouden kunnen zijn. Voor belangrijke leerdomeinen zijn, blijkens het peilingsonderzoek PPO, zelfs experts te optimistisch over het niveau dat voor basisschoolleerlingen haalbaar is (Van der Schoot, 2008). Nog een andere voorwaarde is dat leerlingen, leraren en managers zich daadwerkelijk gedragen zoals de in paragraaf 2.2. beschreven evaluatieve cyclus voorschrijft. Ze moeten dus ook de laatste stap nog zetten, dat wil zeggen na de stap van interpretatie nog een beslissing nemen. Vaak valt het niet mee om na een tegenvallende uitkomst de vinger op de zere plek te leggen en een verbetertraject te starten.

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

Vanzelfsprekend is er empirisch onderzoek verricht naar de effectiviteit van meetgestuurd onderwijs. De uitkomsten van zulk onderzoek bieden een belangrijke aanvulling op onze analyse 'ex ante'. Er is een omvangrijk en heel divers corpus aan onderzoeksliteratuur beschikbaar. De diversiteit hangt samen met de verschillende perspectieven: zijn het leerlingen, leraren of managers die gebruik maken van voorde-ringengegevens? We hebben ervoor gekozen aan afzonderlijke onderzoeken weinig of geen aandacht te geven. We beperken ons hier tot de hoofdlijnen, zo veel mogelijk op basis van overzichtsartikelen. In onze bespreking sluiten we aan bij het bekende onderscheid tussen formatieve en summatieve toetsing. De eerste is gericht op bijsturing van het onderwijsleerproces, de tweede op evaluatie achteraf (zie ook paragraaf 2.2). Onze conclusies zijn in feite weinig hoopvol. Van formatieve toetsing staat weliswaar vast dat het bij kan dragen aan verbeterde leerresultaten, maar bekend is ook dat leraren alleen met grote moeite tot formatieve toetsing te bewegen zijn. De effectiviteit van summatieve toetsing is omstreden en deze vorm van toetsing brengt negatieve neveneffecten met zich mee.

De effectiviteit van formatieve toetsing

Precies tien jaar geleden publiceerden Black en Wiliam een omvattend overzicht van het onderzoek naar de effecten van toetsgebruik op de kwaliteit van het onderwijs (Black & Wiliam, 1998). Dit overzicht, gebaseerd op meer dan 250 afzonderlijke publicaties, heeft in de erop volgende jaren veel gezag verworven. Zo wordt er in latere literatuur veelvuldig naar verwezen. De review gaat met name in op effecten van *formative assessment* '... [which] is to be interpreted as encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged' (o.c., p. 8). In de review staan drie vragen centraal:

- a. Zijn er aanwijzingen dat formatieve toetsing leidt tot beter leren?
- b. Zijn er aanwijzingen dat de bestaande praktijk van toetsgebruik in de klas kan worden verbeterd?
- c. Is duidelijk welke maatregelen nodig zijn om tot een verbeterde formatieve toetsing te komen?

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Het antwoord op de beide eerste vragen is bevestigend: a. regelmatige feedback aan leerlingen heeft een positief effect op hun vorderingen, en b. de praktijk in klaslokalen laat nog veel ruimte voor verbeteringen. De auteurs signaleren de volgende problemen: veel toetsmethoden versterken oppervlakkig leren, ze leiden tot competitie tussen leerlingen (in plaats van *personal improvement*) en ze zijn meer gericht op het afleggen van verantwoording dan op verbeteren van het leren. Kortom, leraren toetsen wel, maar ze gebruiken de verkeerde toetsen, of ook de goede toetsen op de verkeerde manier. Ze toetsen summatief, terwijl ze formatief zouden moeten toetsen.

Black en Wiliam staan in deze conclusies niet alleen. In een andere review, gewijd aan de effectiviteit van feedback, wordt eveneens een negatieve conclusie getrokken over de manier waarop leraren toetsen en andere meetinstrumenten gebruiken (Hattie & Timperley, 2007). We citeren: 'Most current assessments provide minimal feedback, too often because they rely on recall and are used as external accountability thermometers rather than as feedback devices that are integral to the teaching and learning process' (p. 104). Hattie en Timperley merken op dat feedback in schoolklassen vaak persoonsgericht is in plaats van taakgericht en derhalve ineffectief. In hun model voor effectieve feedback staan drie vragen centraal: wat zijn de doelen, hoe ver ben ik al gevorderd, en wat moet ik doen om de doelen te realiseren. Effectieve feedback zou volgens dit model op alle drie de vragen antwoord moeten geven, zowel als het gaat om feedback aan leerlingen als om feedback aan leraren. Vergelijkbare inzichten zijn te vinden in de conclusies die Shute (2008) verbindt aan haar review over de effectiviteit van formatieve feedback.

De derde vraag van Black en Wiliam, over welke maatregelen er nodig zijn om tot een verbeterde formatieve toetsing te komen, valt buiten het bestek van deze sectie. Toch is hun antwoord op zichzelf genomen interessant. Zij staan op het standpunt dat verbeteringen niet gemakkelijk te realiseren zijn en langlopende verbeteringstrajecten vergen, waarin experts en leraren intensief samenwerken. De ineffectieve toetsingsmethoden zijn diep geworteld, zowel in de schoolorganisatie als in de leraar. In latere artikelen gaan ze in op hun ervaringen, opgedaan in projecten om leraren te helpen hun toetspraktijk te verbeteren (Black e.a., 2004; Black & Wiliam, 2003; Wiliam,

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

2008). Deze ervaringen bevestigen dat veranderingen in de toetspraktijk veel tijd en intensieve begeleiding vergen.

De effectiviteit van summatieve toetsing

Summatieve toetsing vindt vaak plaats aan het eind van een cursus of onderwijsprogramma, maar ook wel tussentijds als een programma meerdere jaren beslaat. De uitkomsten worden in hoofdzaak voor twee doeleinden aangewend, voor het certificeren van leerlingen en voor evaluatie van het betreffende onderwijsonderdeel en – in het verlengde daarvan – voor de evaluatie van de leraar en de school. In het Nederlandse basisonderwijs is certificering strikt genomen niet aan de orde, al kunnen scores op de Eindtoets Basisonderwijs en vergelijkbare toetsen wel een rol spelen bij de beslissing hoe de schoolloopbaan verder zal verlopen. In deze sectie beperken we ons tot de evaluatieve functie: leidt de regelmatige afname van summatieve toetsen en het gebruik daarvan met het oog op evaluatie van programma's en scholen tot onderwijsverbetering?

Over het antwoord op deze vraag wordt al decennialang gedebatteerd, ongeveer zo lang als dat er met dergelijke toetspraktijken ervaring is opgedaan. De discussie vindt de laatste jaren plaats onder de noemer van *standards-based education*, en eerder onder *outcome-based education*. In de Verenigde Staten, en overigens ook elders, zijn veel scholen onderworpen aan zogenoemde *accountability* programma's. Zulke programma's eisen dat scholen jaarlijks toetsen afnemen en over de resultaten daarvan rapporteren. Vaak zijn de af te nemen toetsen precies voorgeschreven en zijn aan de toetsen ook standaarden verbonden. Het gebruik van standaarden maakt het mogelijk een uitspraak te doen over goede en slechte scholen. Een slechte school is bijvoorbeeld een school waar relatief weinig leerlingen de standaard hebben gehaald. Het Amerikaanse programma *No Child Left Behind* (verder NCLB) uit 2002 is een recent voorbeeld van een *accountability* programma. NCLB verplicht de Amerikaanse staten tot het instellen van een verplicht toetsprogramma, jaarlijks af te nemen aan alle leerlingen in het basisonderwijs en het voortgezet onderwijs (*grade 3 – 8*). Verplichte toetsing moet leiden tot een hoger onderwijspeil, onder andere door de identificatie van leerlingen en scholen die onder de maat (de *standard*) presteren. Om scholen in beweging te krijgen, zijn aan negatieve uitkomsten sancties verbonden.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

De effectiviteit van de *standards-based reform movement* is omstreden. In de discussie gaat veel aandacht uit naar onbedoelde negatieve bijeffecten (Darling-Hammond, 2004; De Wolf & Janssens, 2007; Ingram, Louis & Schroeder, 2004; Kornhaber, 2004; McGill-Frantzen & Allington, 2006; Mehrens, 2002; Shepard, 2003; Wang, Beckett & Brown, 2006). We vatten ze samen onder vijf noemers.

a. Summatieve toetsing kan leraren en scholen tot onzuiver handelen brengen. De kans daarop is groter, naarmate de belangen voor de school groter zijn. Leraren kunnen bijvoorbeeld de zwakste leerlingen buiten de toetsing houden. Om verzekerd te blijven van correcte uitkomsten zijn controlemaatregelen nodig.

b. Summatieve toetsing kan leiden tot verarming van het curriculum. De kans daarop is groter naarmate de meetinstrumenten een slechtere afspiegeling vormen van het leerplan. In dat geval laten leraren zich verleiden tot een overaccentuering van de getoetste kennis en vaardigheden (*teaching to the test*).

c. Summatieve toetsing is onbillijk tegenover scholen met achterstandsleerlingen. Die onbillijkheid ontstaat doordat schoolvorderingen niet alleen afhankelijk zijn van de kwaliteit van de school, maar bijvoorbeeld ook van de thuissituatie van leerlingen. Door middel van het schatten van de zogenoemde toegevoegde waarde kan hiervoor worden gecorrigeerd, maar aan deze handelwijze kleven haken en ogen (Amrein-Beardsley, 2008; Visscher, 2001). De auteur onderscheidt technisch-analytische problemen, bruikbaarheidsproblemen en ethische en maatschappelijke problemen.

d. Summatieve toetsing brengt risico's met zich mee voor zwakke leerlingen. Scholen kunnen ertoe besluiten zwakke leerlingen te laten zitten of naar andere scholen te verwijzen om daarmee het schoolgemiddelde positief te beïnvloeden. Dit risico geldt in het bijzonder de leerlingen die de gestelde standaard niet halen.

e. Summatieve toetsing kan leiden tot onderstimulering van getalenteerde leerlingen. Zulke leerlingen halen de standaard toch wel en leraren kunnen daardoor in de verleiding komen hun aandacht te richten op leerlingen die 'op de wip' zitten, de wip van net wel of net niet de standaard halen.

Onlangs rapporteerde de Inspectie van het Onderwijs over een vergelijking tussen zogenoemde taalzwakke, taalsterke en gemiddelde taalscholen (Inspectie van het Onderwijs, 2008). Een van de vergelijkingscriteria was de vraag of de school jaarlijks

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

systematisch de kwaliteit van de leeropbrengsten evalueert. In het toetsingskader van de Inspectie maakt dit criterium onderdeel uit van de rubriek 'Kwaliteitszorg'. De uitkomst is dat de drie groepen scholen op dit criterium niet van elkaar afwijken (o.c., p. 34). Deze uitkomst is in feite een contra-indicatie voor de veronderstelling dat regelmatige summatieve toetsing tot een verbeterde leeropbrengst zou leiden. Hoe serieus die contra-indicatie genomen moet worden, is overigens moeilijk te beoordelen omdat de vergelijking alleen bivariaat is gedaan, zonder rekening te houden met eventuele multivariate relaties. Ook is het genoemde criterium (of de school de kwaliteit van de leeropbrengsten jaarlijks evalueert) een onvolkomen operationalisatie van de aandacht die een school aan summatieve toetsing schenkt.

2.5 Het gebruik van vorderingengegevens in het basisonderwijs

Wat is er bekend over het gebruik van vorderingenmetingen in het basisonderwijs? Is de werkwijze sterk meetgestuurd, dat wil zeggen dat leerlingen, leraren en managers zich sterk door vorderingenmetingen laten leiden? Of spelen metingen slechts een bescheiden rol? Bij de behandeling van deze vragen maken we gebruik van een tweetal Nederlandse bronnen. De oudste bron vormt een onderzoek naar leerlingvolgsystemen in het basisonderwijs rond de eeuwwisseling (Blok, Otter & Roeleveld, 2001, 2002). Aan dat onderzoek is deelgenomen door ongeveer 800 leraren uit het (speciaal) basisonderwijs en 200 schooldirecties. Zij vulden een uitgebreide vragenlijst in over hun evaluatiepraktijk. Van recenter datum is de zesde meting van het Primacohortonderzoek, die in 2004-2005 is verricht (Van der Veen, Van der Meijden & Ledoux, 2006). Informanten waren omstreeks 1400 leraren basisonderwijs en 350 schooldirecties. Van beide onderzoeken kunnen de gegevens bij benadering als landelijk representatief worden beschouwd. Een zwak punt is wel dat gegevens alleen via vragenlijsten verkregen zijn en betrekking hebben zelfrapportage. Het is daarom de vraag of de verkregen gegevens een juist beeld van de werkelijkheid geven. Een andere moeilijkheid is dat beide onderzoeken alleen op enkele hoofdlijnen uitkomsten bieden die relevant zijn voor de vragen die we hier aan de orde stellen. We structure-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

ren onze behandeling aan de hand van de drie groepen betrokkenen die we eerder hebben onderscheiden.

Leerlingen en hun gebruik van vorderingengegevens

Over het gebruik dat leerlingen van vorderingengegevens maken is nauwelijks iets bekend. In het Prima-cohortonderzoek zijn hierover helemaal geen vragen gesteld, in het onderzoek van Blok, Otter en Roeleveld slechts één vraag. Die vraag was in hoeverre leraren het leerlingvolgsysteem gebruiken om 'leerlingen inzicht te geven in de eigen vorderingen'. Dat blijkt vaker niet dan wel te gebeuren. Het gemiddelde bedraagt 2.6 op een schaal van 1 (helemaal niet van toepassing) tot en met 5 (helemaal van toepassing); zie Blok e.a. (2002, p. 14). Uit hetzelfde gegevensbestand valt af te leiden dat de leeftijd van de leerling een factor is die meespeelt. Naarmate leerlingen ouder worden, krijgen ze van hun leraren vaker feedback over hun vorderingen (Blok e.a., 2001, p. 97 e.v.). Deze twee uitkomsten laten nog veel te raden over, onder andere naar de aard van de informatie die leraren naar leerlingen terugkoppelen. Bijvoorbeeld, gaat het om schoolcijfers, om toetsuitslagen of om inhoudelijke feedback? Het zijn vragen die van onderzoekers, overigens ook van onderzoekers uit andere landen, nog nauwelijks aandacht gekregen hebben.

Leraren en hun gebruik van vorderingengegevens

Uit de beschikbare gegevens is een – tamelijk grove – reconstructie op te stellen van de manier waarop leraren de evaluatieve cyclus (zie par. 2.3) hanteren. De eerste stap, het vastleggen van doelen en standaarden, gebeurt slechts door een minderheid van de leraren, zo'n 25 tot 40 procent (Van der Veen e.a., 2006, p. 75 - 81). Het verzamelen en registreren van vorderingengegevens gebeurt daarentegen veel frequenter. Vierenveertig procent van de leraren doet dat voor begrijpend lezen minimaal vijf keer per jaar, en voor rekenen/wiskunde bedraagt het percentage zelfs 73 (Van der Veen e.a., 2006, p. 75). Over het interpreteren van de vorderingengegevens, de vierde stap van de cyclus, zijn in het Prima-cohortonderzoek geen vragen gesteld. Maar als het stellen van doelen en standaarden niet is gebeurd, kan er van een ordentelijke interpretatie geen sprake zijn. Er resteert tot slot dan nog de vijfde stap, waarin beslissingen wor-

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

den genomen. In het Prima-cohortonderzoek zijn zes terreinen onderscheiden waarop leraren of directies beslissingen kunnen nemen (Van der Veen, 2006, p. 26). We nemen de tabel die hierop betrekking heeft, integraal over (Tabel 2.1). De gegevens zijn overigens niet rechtstreeks afkomstig van leraren, maar van schooldirecties. Dat kan nadelig zijn voor de validiteit, indien directies minder goed zicht hebben op het evalueren door hun leraren. Uit de tabel blijkt dat leraren vorderingengegevens aanwenden voor alle onderscheiden terreinen: het vaststellen van het onderwijs voor de hele groep (functie a), het vaststellen van het onderwijs voor individuele leerlingen (functie b), het vaststellen van effecten van handelingsplannen, het maken van foutenanalyses (functie d) en het evalueren van het onderwijsaanbod¹ (functie e). De functie die het frequentst wordt onderkend, is functie b (het vaststellen en plannen van het onderwijs voor individuele leerlingen).

Tabel 2.1 Functies waarvoor het leerlingvolgsysteem wordt ingezet (1 = niet, 2 = enigszins, 3 = in belangrijke mate); percentages (n = 351 schooldirecties)

Functies	Antwoord		
	1	2	3
a. Om het onderwijs voor de groep vast te stellen of te plannen	14	49	37
b. Om het onderwijs voor individuele leerlingen vast te stellen of te plannen	3	20	77
c. Om het effect van handelingsplannen vast te stellen	8	37	55
d. Om foutenanalyses te maken	11	43	46
e. Om het onderwijsaanbod te evalueren	2	35	63
f. Voor het plannen van schoolbeleid	7	50	43

¹ We nemen hier de formulering uit de vragenlijst over, maar waarschijnlijk hebben de onderzoekers bedoeld 'het evalueren van het effect van het onderwijsaanbod op de vorderingen van de leerlingen'. We nemen aan dat de respondenten de vraag hebben opgevat zoals door de onderzoekers is bedoeld.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

De uitkomsten uit het onderzoek van Blok e.a. (2001, 2002) bieden in sommige opzichten een nuttige aanvulling op het voorafgaande. Leraren verzamelen en registreren veelvuldig gegevens over de leervorderingen, in elk geval voor de kernvakken. Ze gebruiken daarbij een breed samengesteld arsenaal aan meettechnieken: observaties, mondelinge beurten, schriftelijk werk, methodegebonden en methodeonafhankelijke toetsen. De belangrijkste functies zijn rapportagefuncties (aan ouders en bij een overdracht aan collega's) en didactische functies, zoals het signaleren van achterblijvers. Functies op het terrein van kwaliteitszorg, ook wel aangeduid als summatieve functies, worden door leraren van beduidend minder belang geacht.

Managers en hun gebruik van vorderingengegevens

Net als in het geval van de leerlingen is het beeld over het gebruik dat managers – we zullen eenvoudshalve verder over schooldirecties spreken – van vorderingengegevens maken, verre van compleet. Wel is er in het Prima-cohortonderzoek een aparte vraag opgenomen over het werken met vooraf bepaalde doelen per leerjaar. Dat blijkt veel voor te komen. Ruim 60 procent van de directies geeft aan leerjaardoelen vastgelegd te hebben in termen van te behandelen leerstof, ongeveer 30 procent relateert de doelen aan een methodeonafhankelijke toets of het leerlingvolgsysteem (Van der Veen e.a., 2006, p. 27). Waarschijnlijk zijn hierbij ook standaarden in het geding. Want ongeveer 65 procent van de directies geeft aan dat voor zwakke leerlingen minimumdoelen worden gehanteerd. Een aparte standaard voor (zeer) goede leerlingen komt met 35 procent veel minder voor.

In tabel 2.1 zagen we al dat het leerlingvolgsysteem veelvuldig wordt ingezet voor het plannen van schoolbeleid (functie f.), wat dat dan ook moge inhouden. Een systematische evaluatie van de effectiviteit van het onderwijsaanbod – één zo'n aspect van schoolbeleid – is in elk geval veel minder algemeen. Ongeveer 65 procent van de directies geeft aan zulks te doen (Van der Veen e.a., 2006, p. 27).

Uit het onderzoek van Blok e.a. (2001) komt naar voren dat schooldirecties vorderingengegevens in de eerste plaats benutten voor het opstellen van adviezen ten behoeve van de schoolkeuze (p. 100-101). Andere functies die met een zekere regelmaat wor-

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

den genoemd, zijn (tussen haakjes de gemiddelde scores op een schaal van 1 = helemaal niet van toepassing tot 5 = helemaal van toepassing):

- het schoolbeleid ten aanzien van de aanschaf van leermiddelen (3.3);
- het schoolbeleid ten aanzien van scholing of begeleiding (3.3.);
- het evalueren van de voortgang in relatie tot de kerndoelen (3.1);
- het ontwikkelen van het schoolplan (2.9).

2.6 Tussentijdse conclusies

In deze paragraaf vatten we de opbrengst van de literatuurverkenning samen. We doen dat aan de hand van de in paragraaf 2.1 onderscheiden onderzoeksvragen.

a. Wat valt er te verstaan onder data-driven teaching?

In het voorafgaande hebben wij voorgesteld de term *data-driven teaching* te vervangen door de term meetgestuurd onderwijs. We hebben meetgestuurd onderwijs als volgt omschreven: 'meetgestuurd onderwijs is onderwijs waarbij betrokkenen zich in hun taakuitvoering laten sturen door uitkomsten van metingen'. We hebben drie groepen betrokkenen onderscheiden: leerlingen, leraren en managers. Voor elk van deze groepen heeft meetgestuurd onderwijs een andere verschijningsvorm, afhankelijk van ieders specifieke positie. Een voorbeeld: voor leerlingen is leren de kerntaak. Meetgestuurd onderwijs betekent dus voor leerlingen in hoeverre zij zich in hun leergedrag door metingen laten sturen. Bij metingen dient men overigens niet alleen te denken aan uitkomsten van meerkeuzetoetsen, maar ook aan methodegebonden toetsen, beoordelingen van schriftelijk werk, mondelinge overhoringen en dergelijke.

b. Welke effecten worden er in de onderzoeksliteratuur van data-driven teaching beschreven?

We hebben onderscheid gemaakt tussen formatieve toetsing (gericht op bijsturing van het onderwijsleerproces) en summatieve toetsing (achteraf vaststellen welke de voorde- ringen zijn geweest). Van formatieve toetsing is bekend dat het bij kan dragen aan verbeterde leerresultaten. Van belang daarbij is vooral dat leerlingen inhoudelijke

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

feedback krijgen, feedback die taakgericht is en die leerlingen aanwijzingen geeft over wat ze in hun denken of gedrag moeten veranderen. Tegelijkertijd blijkt uit onderzoek, dat de meetpraktijk in klaslokalen nog veel te wensen overlaat. Leraren bepalen weliswaar regelmatig de vorderingen van hun leerlingen, maar hun feedback aan leerlingen is vaak persoonsgericht in plaats van taakgericht. Met andere woorden, ze informeren hun leerlingen over hoe goed ze iets doen, maar veelal niet over hoe ze zich kunnen verbeteren.

De effectiviteit van summatieve toetsing is omstreden. In de Verenigde Staten is veel ervaring opgedaan met summatieve toetsingsprogramma's, aan de uitkomsten waarvan voor leraren en scholen consequenties zijn verbonden. In zulke programma's worden leraren en scholen met tegenvallende toetsuitslagen onder druk gezet, bijvoorbeeld via openbaarmaking van de resultaten. In de discussie over de effectiviteit van zulke programma's ligt het accent op negatieve bijeffecten, waaronder onzuiver handelen van scholen en leraren en verarming van het curriculum tot alleen die onderwerpen die worden getoetst. Uit onderzoek van de Nederlandse Inspectie van het Onderwijs komt eveneens naar voren dat summatieve toetsing niet of nauwelijks samenhangt met verbeterde leerprestaties, in dit geval voor het leerdomein Nederlandse taal (Inspectie van het Onderwijs, 2008).

c. Over welke vorderingengegevens beschikken scholen voor basisonderwijs?

Leraren basisonderwijs beschikken – in elk geval voor de kernvakken Nederlandse taal en rekenen/wiskunde – over een grote hoeveelheid vorderingengegevens. De gegevens zijn van uiteenlopende aard. De verspreidingsgraad van het Cito-leerlingvolgsysteem bedraagt rond de 90 procent. Daaruit mag men afleiden dat de overgrote meerderheid van leraren gebruik maakt van de daarin opgenomen toetsen voor technisch en begrijpend lezen, spelling en rekenen/wiskunde. Verondersteld mag bovendien worden dat de meeste scholen zich houden aan het door het Cito aanbevolen afnameschema, dat grosso modo neerkomt op een halfjaarlijkse afname. Overigens is ook de verspreidingsgraad van de Eindtoets Basisonderwijs van het Cito hoog. Deze toets is in 2008 door 6300 van de bijna 7000 basisscholen afgenomen. Naast het gebruik van Cito-toetsen verzamelen leraren nog op andere manieren gegevens over

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

leervorderingen. Voor technisch lezen gaat het om leesbeurten, voor spelling en rekenen/wiskunde om methodegebonden toetsen en ook observeren tijdens lessen wordt veelvuldig toegepast. Steeds meer scholen kennen een centrale opslag van vorderingsgegevens, waardoor gegevens van afzonderlijke leraren ook voor managers toegankelijk zijn. In zo'n centraal systeem worden vooral uitslagen van methodeonafhankelijke toetsen opgenomen, zoals de toetsen uit het leerlingvolgsysteem. Het grootste deel van de anderssoortige gegevens is alleen aan de afzonderlijke leraren bekend. Leerlingen hebben in het algemeen geen toegang tot vorderingsgegevens, behoudens uitgedeelde toetsuitslagen, (rapport)cijfers, en dergelijke.

d. Hoe maken scholen voor basisonderwijs gebruik van de beschikbare vorderingsgegevens?

Ten aanzien van deze vraag hebben we geconstateerd dat er grote hiaten zijn in onze kennis. Weliswaar is via vragenlijstonderzoek vastgesteld voor welke doeleinden leraren en managers vorderingsmetingen inzetten, maar hoe ze dat doen – bijvoorbeeld, of ze dat doen met inachtneming van de evaluatieve cyclus – is onbekend. Leraren gebruiken meetuitkomsten in de eerste plaats om anderen, zoals ouders en collega's te informeren over de vorderingen van hun leerlingen. Van veel minder belang is in hun ogen de evaluatie van leeropbrengsten voor de groep als geheel. Ook managers geven geen groot gewicht aan een evaluatie van de leeropbrengsten op groeps- of schoolniveau. Of leraren en managers de evaluatieve cyclus als leidraad voor hun summatieve evalueren nemen, is strikt genomen onbekend. Maar op basis van informele observaties en gesprekken op scholen durven wij de veronderstelling aan dat leraren en managers bij hun evaluatie-activiteiten vaak weinig rationeel te werk gaan.

e. Welke belemmeringen zijn er bij de invoering van data-driven teaching?

Deze vraag is in het voorafgaande alleen zijdelings aan de orde geweest. Ons antwoord is daardoor misschien iets speculatiever dan de antwoorden op de eerdere vragen. We doen evenwel ons best om een tekort aan empirische gegevens te compenseren langs de weg van een beredeneerde analyse. We delen ons antwoord in aan de hand van de drie eerder genoemde groepen: leerlingen, leraren en managers.

Voor leerlingen is waarschijnlijk de grootste belemmering dat zij geen toegang hebben tot die vorm van feedback die, zoals uit onderzoek blijkt, hen het beste zou helpen, namelijk de taakgerichte.

Voor leraren gelden andere belemmeringen. We noemen de drie belangrijkste. We nemen daarbij als uitgangspunt dat leraren zowel formatief als summatief zouden moeten toetsen. De eerste belemmering is onwetendheid, die vervolgens tot een minder effectieve evaluatiepraktijk leidt. Leraren bepalen weliswaar veelvuldig hoe leerlingen vorderen, maar ze gebruiken hun meetuitkomsten niet goed. Ze bieden leerlingen vooral persoonsgerichte – geen taakgerichte – feedback en ze tonen onvoldoende belangstelling voor summatieve toetsing, de evaluatie achteraf. Daarbij komt dat leraren waarschijnlijk weinig weet hebben van de evaluatieve cyclus. Die onwetendheid komt niet uit de lucht vallen. Elders in het rapport (hoofdstuk 3) zal blijken dat lerarenopleidingen nauwelijks aandacht schenken aan een effectieve evaluatiepraktijk. Maar onwetendheid is misschien niet het hele verhaal. Er wordt op leraren een steeds grotere druk gelegd om summatieve toetsen, zoals die uit het leerlingvolgsysteem van het Cito, te gebruiken. Die druk, onder andere afkomstig van schooldirecties, schoolbesturen en de inspectie, draagt bij aan een toetspraktijk die steeds verder in onbalans geraakt. De derde belangrijke belemmering is een gebrek aan tijd en middelen. Met name de formatieve evaluatie en de daarbij behorende taakgerichte feedback stellen eisen aan leraren die moeilijk te realiseren zijn. Leraren ontbreekt gewoonweg de tijd om alle leerlingen van passende taakgerichte feedback te voorzien.

Tenslotte, voor managers gelden deels nog weer andere belemmeringen. Een belemmering is waarschijnlijk *niet* de beschikbaarheid van meetuitkomsten. Er wordt door scholen voldoende summatief getoetst. Of de uitkomsten ook in een hanteerbare vorm beschikbaar zijn, is misschien nog wel de vraag. Op basis van een in 2000 afgenomen vragenlijst constateerden wij, dat slechts twee op de drie basisscholen toetsgegevens centraal in een computer opslaat (Blok e.a., 2001). Waar managers net als leraren waarschijnlijk mee te kampen hebben, is een kennisachterstand. Ook voor managers geldt dat ze onvoldoende geschoold zijn in het toepassen van de evaluatieve cyclus. Dat was bijvoorbeeld een belangrijke uitkomst van een onderzoek naar de ervaringen van basisscholen met schoolzelfevaluatie (Blok, Slegers & Karsten, 2006). Verder

Van 'data-driven teaching' naar meetgestuurd onderwijs; verkenning van de literatuur

veronderstellen wij dat er bij managers een zekere terughoudendheid bestaat op basis van de angst voor misbruik. Die angst is niet onterecht. Hoe zwaarder het belang dat wordt gehecht aan rendementscijfers op basis van leervorderingen, des te groter is het risico van een onzuiver gebruik. Zo heeft de openbaring van schoolgemiddelden op de Eindtoets Basisonderwijs ertoe geleid, dat scholen terughoudend geworden zijn om hun zwakste leerlingen aan de Eindtoets laten deelnemen (Inspectie van het Onderwijs, 2006). Mogelijk speelt tenslotte ook nog een rol dat het toetsingskader van de Inspectie de nadruk legt op summatieve toetsing en weinig aandacht schenkt aan het belang van evalueren voor de ondersteuning van leerprocessen (zie ook Janssens, 2008, p. 6).

3 Opvattingen van experts over meetgestuurd onderwijs

3.1 Doel van het deelonderzoek

In aansluiting op het literatuuronderzoek is praktijkkennis verzameld over toepassingen van meetgestuurd onderwijs in het Nederlandse basisonderwijs. Daarvoor zijn groepsgesprekken en – deels telefonische – interviews gehouden met deskundigen op het gebied van kwaliteitszorg en schoolontwikkeling. In dit hoofdstuk doen wij verslag van de mening van deze experts over meetgestuurd onderwijs in Nederland. Daarnaast is een korte inventarisatieronde gehouden bij pabo's om na te gaan in hoeverre de opleidingen voor leraar basisonderwijs studenten voorbereiden op het benutten van data voor het verbeteren van hun eigen onderwijs in de latere beroepspraktijk. Ook de uitkomsten van deze ronde komen in dit hoofdstuk aan de orde.

De gesprekken met de experts zijn benut voor het beantwoorden van alle onderzoeksvragen waarvoor na de literatuurstudie nog aanvullende informatie gewenst is. Dat zijn alle onderzoeksvragen met uitzondering van vraag 1 (zie hoofdstuk 1) over de handelingen/stappen die bij meetgestuurd onderwijs aan de orde zijn. Die vraag achten we voldoende beantwoord vanuit de literatuurstudie. De resterende onderzoeksvragen zijn (we houden de nummering uit hoofdstuk 1 aan):

2. Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?
3. In hoeverre en hoe worden deze gegevens gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?
4. Welke handelingen/stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom?

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

5. Bestaat er draagvlak bij leerkrachten en schoolleiders voor meetgestuurd onderwijs?
6. Wat zijn risico- en succesfactoren bij meetgestuurd onderwijs? Zijn er eventueel onbedoelde effecten?
7. Wat hebben schoolleiders en leerkrachten nodig om meetgestuurd onderwijs te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

Doel van dit deelonderzoek was om via de gesprekken met experts een stand van zaken te schetsen van meetgestuurd onderwijs in de praktijk anno 2008: wat doen scholen al wel, wat nog niet, wat levert dat op, wat zijn de risico's? Wat zijn de voorwaarden waaronder meetgestuurd onderwijs kan worden toegepast op zo'n manier dat er ook verbetering uit voortkomt van de kwaliteit van het onderwijs, met name van de prestaties bij rekenen en taal/lezen?

3.2 Opzet van het deelonderzoek

Interviews en groepsgesprekken

Er zijn verschillende typen experts aan het woord gekomen. Twee maal gebeurde dat in een groepsgesprek, op de locatie van de betrokkenen zelf. Dat was het geval bij:

- a) *Cito*-medewerkers die ervaring hebben met het gebruik van het Cito-Leerlingvolgsysteem (LVS) en het LVS-computerprogramma door scholen en/of betrokken zijn bij pilots waarin nieuwe vormen van toetsen worden geprobeerd of leerstandaarden worden ontwikkeld;
- b) *Medewerkers van de Inspectie van het onderwijs*, die een goed zicht hebben op zowel beschikbaarheid als (typen) gebruik van gegevens over leervorderingen in scholen.

Verder zijn telefonische interviews gehouden met enkele andere experts:

Opvattingen van experts over meetgestuurd onderwijs

- c) *Schoolbegeleiders* die ervaring hebben met vormen van meetgestuurd onderwijs en scholen daarbij ondersteunen, bijvoorbeeld begeleiders die ervaring hebben met schoolverbeteringsprocessen (op zwakke scholen, onderwijskansenscholen).
- d) *Overige deskundigen* op het gebied van kwaliteitszorg, schoolontwikkeling en het gebruik van ken- en stuurgetallen

In totaal zijn negen experts uit deze twee laatste categorieën² geïnterviewd, in zeven gesprekken.

Korte inventarisatieronde onder pabo's

Omdat het tevens relevant is te weten in hoeverre de opleidingen voor leerkrachten in het basisonderwijs al dan niet actief aandacht besteden aan verschillende aspecten van meetgestuurd onderwijs is een korte inventarisatieronde (per e-mail en telefoon) gehouden onder coördinatoren van onderwijsprogramma's en docenten taal en rekenen van enkele pabo's. Gestreefd is naar een kleine, maar representatieve selectie wat betreft spreiding over het land en studentenaantallen. Uiteindelijk hebben we informatie ontvangen van tien respondenten, van de volgende zeven pabo's: IN-Holland Alkmaar, Groenewoud Arnhem (2x), Iselinge Hogeschool Doetinchem, Driestar Educatief Gouda, Hogeschool Domstad Utrecht (3x), Fontys-PABO Eindhoven en Hogeschool Zuyd Heerlen/Maastricht. Aan de respondenten is gevraagd wat in hun opleiding aan studenten wordt aangeboden op het gebied van het verzamelen, analyseren en interpreteren van gegevens over leerprestaties van leerlingen en het afstemmen van het onderwijsaanbod op de bevindingen.

Gespreksleidraden

Voor zowel de groepsgesprekken als de telefonische interviews met de overige experts zijn gespreksleidraden ontwikkeld. De accenten konden wisselen, al naar gelang de expertise van degenen die werden bevroegd, maar ze hadden alle in principe dezelfde opbouw en inhoud. De zeven kernvragen daarin luiden:

1. *Komt meetgestuurd onderwijs in de praktijk voor?*

² De meesten van hen zijn zowel ervaren in het begeleiden van scholen, als deskundige via coördinatie van grotere projecten. Vandaar dat er verder geen onderscheid meer is gemaakt.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

2. *Zijn er voorbeelden te geven van meetgestuurd onderwijs?*

Deze twee vragen samen zijn gericht op het beantwoorden van onderzoeksvraag 2, 3 en 4: welke gegevens worden verzameld en hoe worden ze gebruikt?

3. *Welke effecten zijn zichtbaar in de praktijk?*

Deze vraag sluit zowel aan bij onderzoeksvraag 4 (welke stappen worden uitgevoerd) als bij onderzoeksvraag 6 (wat zijn succesfactoren).

4. *Hoe staan betrokkenen er tegenover?*

Deze vraag heeft betrekking op onderzoeksvraag 5: bestaat er draagvlak bij de diverse betrokkenen.

5. *Wat zijn de belangrijkste condities?*

Deze vraag sluit aan op het eerste deel van onderzoeksvraag 7: wat hebben schoolleiders en leerkrachten nodig om meetgestuurd onderwijs te kunnen uitvoeren?

6. *Zijn er ook neveneffecten of risico's zichtbaar in de praktijk?*

Met deze vraag hebben we gezocht naar antwoorden op het tweede gedeelte van onderzoeksvraag 6: wat zijn eventuele risico's en ongewenste effecten van meetgestuurd onderwijs?

7. *Heeft u aanbevelingen?*

Deze vraag is gelijk aan het laatste gedeelte van onderzoeksvraag 7: tot welke aanbevelingen voor onderwijs en overheid leidt dit?

Bij de meeste vragen is, voor zover dat mogelijk en relevant was, ingegaan op de verschillende posities van schoolbestuurders, schoolleiders, intern begeleiders, leraren en leerlingen. Van alle groepsbijeenkomsten en telefonische interviews is een verslag gemaakt. De gespreksverslagen zijn ter aanvulling en correctie voorgelegd aan de respondenten.

Pabo-inventarisatie

Ook voor de korte inventarisatie bij de pabo-opleidingen, is gebruik gemaakt van een interviewleidraad. Daarin kwamen de volgende thema's aan de orde:

1. *Leren studenten op uw opleiding om toetsen af te nemen, of observaties te doen (of om eventueel 'leergesprekjes met leerlingen te voeren) die gericht zijn op het bepalen van het niveau van de leerlingen (bij taal of rekenen)?*
2. *Om welke toetsen gaat het dan bijvoorbeeld?*
3. *Leren studenten om zelf de toetsen na te kijken en te scoren volgens de handleiding, waarbij zij dan bijvoorbeeld ook moeten letten op het type fouten dat een leerling maakt?*
4. *Leren studenten iets over het analyseren en interpreteren van de uitkomsten van toetsgegevens?*
5. *Leren studenten iets over het afstemmen van het onderwijsaanbod op de gevonden resultaten (voor individuele leerlingen bijvoorbeeld in een individueel handlingsplan, of voor de hele groep)?*
6. *Op welke manier zit de kennis over het meten van vorderingen/niveau van de leerlingen in de opleiding?*
7. *Geldt dat voor alle studenten of voor een deel van de studenten?*
8. *Leren studenten iets over kwaliteitsbeleid in de school (inzicht in de sterke en zwakke kanten van het onderwijs en daarover gerichte gegevens verzamelen)?*

In paragraaf 3.3 staan de resultaten van de gesprekken met de experts. Paragraaf 3.4 bevat de uitkomst van de dataverzameling bij de pabo's. In 3.5 vatten we de uitkomsten samen aan de hand van de onderzoeksvragen.

3.3 Opvattingen van experts

3.3.1 Mate en wijze van voorkomen van meetgestuurd onderwijs

In deze paragraaf geven we een korte samenvatting van de antwoorden van de verschillende (groepen) deskundigen op twee van de vragen uit de interviewleidraad: komt meetgestuurd onderwijs in de praktijk voor en hoe ziet het eruit? We beginnen

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

met twee aspecten van meetgestuurd onderwijs: welke gegevens worden er zoal verzameld als het gaat om de prestaties en vorderingen van leerlingen, en hoe worden de gegevens vervolgens opgeslagen en geanalyseerd? Vervolgens komt in paragraaf 3.3.2 aan de orde hoe de verzamelde gegevens worden benut voor de afstemming van het onderwijs op het niveau en de behoeften van de leerlingen.

Welke gegevens worden er in het onderwijs verzameld over de leerprestaties?

De respondenten geven allen aan dat een belangrijk deel van de scholen gebruik maakt van het Cito-LVS. Cito-onderzoek uit 1995 wijst uit dat op dat moment al ongeveer driekwart van de scholen zich houdt aan de werkwijze zoals die in het Cito-leerlingvolgsysteem is ontwikkeld: signaleren, analyseren en didactisch handelen. Men heeft de indruk dat er sindsdien weinig is veranderd. Verder hebben 4500 van de 7000 scholen in Nederland een abonnement op het computerprogramma dat hoort bij het LVS. Zij hebben daarmee in principe de mogelijkheid in huis hun onderwijsopbrengsten efficiënt te analyseren.

De geïnterviewde schoolbegeleiders melden dat verreweg de meeste, zo niet alle scholen waarmee zij contact hebben, in elk geval objectieve gegevens verzamelen over het niveau van de leerlingen op het gebied van lezen/taal en rekenen, inderdaad meestal met toetsen uit het Cito-LVS. De toetsen die de meeste scholen afnemen zijn: Taal voor Kleuters, Ordenen, Leestechiek en tempo (AVI en DMT), Spelling, Leeswoordenschat, Begrijpend Lezen en Rekenen/wiskunde. Op sommige plaatsen is dat ook vrij strak geregeld. In Arnhem hebben de drie grootste schoolbesturen (die samen 90% van de Arnhemse scholen vertegenwoordigen) de Stichting PAS in het leven geroepen, die zorgdraagt voor de evaluatie en monitoring van de leerprestaties met behulp van een eigen centrale monitor. Stichting PAS stelt, vanaf het lopend schooljaar (2007-2008), voor alle 45 scholen een schoolrapportage op, die wordt besproken met de directie en het schoolbestuur.³ Vergelijkbaar zijn de Zeeuwse onderwijsmonitor (De Zeeuwse Toetssteen, 2006-2007) waaraan inmiddels ongeveer de helft (114) van

³ De afgelopen jaren is dat al gebeurd voor de 19 onderwijsachterstandscholen. De werkwijze lijkt overigens wel tamelijk uniek van Nederland. De Stichting PAS krijgt althans de laatste tijd regelmatig bezoek vanuit andere gemeentes die dezelfde systematiek willen gaan gebruiken.

Opvattingen van experts over meetgestuurd onderwijs

de scholen (220) in Zeeland deelneemt, en de Toetservice van het Centrum voor Educatieve Dienstverlening (CED) in Rotterdam, waarbij een belangrijk deel van de Rotterdamse basisscholen zich heeft aangesloten. Ook deelname aan grote projecten (zoals momenteel de pilots Taalbeleid, en enkele jaren geleden het Onderwijskansenproject) stimuleert het systematisch meten van de leerprestaties voor taal/lezen en rekenen.

We waren ook geïnteresseerd in (typen) scholen waar toetsgebruik mogelijk minder voorkomt, of waar men gereserveerder staat tegenover toetsafnames. Om die reden hebben we ook een interview gehouden met een deskundige uit het Montessori-onderwijs. Van de in totaal 160 Montessorischolen blijkt het merendeel eveneens het Cito-LVS te gebruiken, een ander deel werkt met toetsen van Eduforce voor de objectieve prestatie metingen. Zo'n 40 scholen gebruiken het Montessori-kindvolgsysteem dat uitgaat van systematische observaties op een aantal domeinen (zoals waarnemen, motoriek en de verschillende kennisgebieden). De observaties kunnen in dit systeem worden afgezet tegen ontwikkelingslijnen in een leerlingprofiel, waarmee zichtbaar wordt of het kind zich sneller of langzamer ontwikkelt dan leeftijdgenoten. Vanaf komend schooljaar zal een digitaal systeem beschikbaar zijn voor de opslag van de leerlingprofielen. Daarmee wordt het ook mogelijk om overzichten te maken per groep, per school of op het niveau van een schoolbestuur. Die overzichten kunnen vervolgens worden ingezet voor verdere schoolontwikkeling of bijsturing van het onderwijs op het niveau van de groep.

Naast de genormeerde, methodeonafhankelijke metingen, worden meestal ook nog methodegebonden toetsen afgenomen. Eén van de deskundigen geeft zelfs aan: *“Ik vind dat scholen ontzettend veel toetsen. Als je het uitreken kom je per schooljaar op een besteding van zo ongeveer drie weken onderwijstijd aan het afnemen van toetsen. Er wordt in die zin heel veel gemeten in het onderwijs, maar of scholen zich laten sturen door de uitkomsten van al die toetsen en metingen, dat is een tweede.”*

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Hoe worden de gegevens verzameld, opgeslagen en geanalyseerd?

Het afnemen van de verschillende toetsen gebeurt in de klas, onder leiding van de leerkracht. De resultaten worden meestal opgeslagen in een digitaal systeem, dat ook kan worden gekoppeld aan gegevens uit de leerlingenadministratie (ESIS-b bijvoorbeeld). Er zijn ook mogelijkheden om resultaten van methodegebonden toetsen op te nemen. Enkele scholen werken vooral nog van papier.

Het verzamelen en analyseren van de uitkomsten is vooral een taak die in de praktijk ligt bij de intern begeleider (of de zorgcoördinator). Hoewel in sommige projecten (zoals het Rotterdamse Delta+) ook leerkrachten zelf in behoorlijke mate gewend zijn om toetsgegevens bij te houden, is in andere projecten de opslag en analyse juist de verantwoordelijkheid van externe bureaus (zoals bij Stichting PAS).

3.3.2 Voorbeelden van meetgestuurd onderwijs

In deze paragraaf komt aan de orde op welke manier de verzamelde gegevens worden geanalyseerd en benut voor het bereiken van een goede afstemming van het onderwijs op dat wat de leerlingen al wel of nog niet beheersen.

Er zijn op het punt van het benutten van de beschikbare gegevens grote verschillen tussen de scholen, zo melden de experts. De indruk van de inspectie is dat managers (schoolleiders en intern begeleiders) nog weinig doen aan meetgestuurd onderwijs. Anno 2007 voldeed nog slechts de helft van de basisscholen aan indicator 1.3 'De school evalueert regelmatig het leren en onderwijzen'. Trendanalyses (longitudinale vergelijkingen) zijn op de meeste scholen nog volslagen onbekend, zowel op groeps- als op individueel niveau. Er lijkt bovendien weinig intrinsieke motivatie te bestaan bij scholen om dit type analyses te gaan doen en het komt zelden voor dat scholen gebruik maken van standaarden. Wel ziet men dat op zwakke scholen vaker sprake is van analyse van toetsresultaten, daar voelen de teams de urgentie sterker om de prestaties te verbeteren. Daar worden de resultaten ook vaker gevolgd aan de hand van vooraf vastgestelde doelen. Wel is het volgens de inspectie algemeen gebruikelijk om data te gebruiken voor het signaleren van zorgleerlingen, zowel aan de onderkant (zwakke leerlingen) als aan de bovenkant (extra goede leerlingen).

Opvattingen van experts over meetgestuurd onderwijs

Ook leraren maken naar de ervaring van de inspectie over het algemeen weinig gebruik van toetsgegevens. Niet alle leraren maken (systematisch) gebruik van methodegebonden toetsen, degenen die dat wel doen registreren meestal alleen voor de zorgleerlingen de toetsresultaten, en soms voor het bepalen van een rapportcijfer. Methodeonafhankelijke toetsen uit het leerlingvolgsysteem worden vooral benut voor de individuele leerlingenzorg, veel minder voor het uitvoeren van analyses op groepsniveau en nog minder voor trendanalyses. Leraren werken wel regelmatig met standaarden, vooral bij technisch lezen en rekenen/wiskunde.

Ook de andere respondenten zijn het erover eens dat de gegevens die scholen verzamelen over de leervorderingen van hun leerlingen in hoofdzaak worden gebruikt voor het bijhouden van die vorderingen op individueel leerlingenniveau en voor het inzichtelijk maken daarvan aan ouders.

Voor met name de zorgleerlingen komt daar nog diagnostiek bij, en worden ook aanpassingen in het onderwijs overwogen zoals preteaching, reteaching en verlengde instructie. De Deltaplusscholen in Rotterdam bijvoorbeeld gebruiken de gegevens tevens voor afstemming van het onderwijs in de groep door de leerlingen in te delen in niveaugroepen en gedifferentieerde instructieniveaus aan te bieden.

In veel mindere mate komt het voor dat scholen de gegevens gebruiken voor de evaluatie van het eigen onderwijs. Eén van de experts signaleert overigens wel dat leerkrachten de resultaten op methodegebonden toetsen intensiever gebruiken als richtlijn voor hun onderwijs, om te zien of de leerlingen de stof al beheersen en aan welke aspecten nog extra aandacht besteed moet worden. Verder zijn er enkele experts die aangeven wel scholen te kennen waar onderwijsevaluatie plaatsvindt aan de hand van leerdoelen, bijvoorbeeld een aantal scholen die momenteel deelnemen aan de Taalpilots. In andere projecten (zoals Deltaplus) bespreekt men in het voortgangsoverleg de knelpunten en de vorderingen van de leerlingen bij bijvoorbeeld rekenen op basis van de resultaten van methodegebonden toetsen en de resultaten uit het Cito-LOVS. Dat vraagt dan overigens wel dat de leerresultaten op gedetailleerd niveau zijn ingevoerd door de leerkracht. Bij die besprekingen gaat het om vragen als: zijn de resultaten zoals verwacht op basis van de eigen indruk van de leerkracht over de leerlingen, is eventuele uitval ergens uit te verklaren, zijn er nog manco's waaraan de komende tijd

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

aandacht besteed moet worden etc. Al deze zaken kan de leerkracht dan vervolgens oppakken in zijn/haar onderwijs.

De scholen die deelnemen aan overkoepelende monitorprogramma's ontvangen van de externe organisatie die de analyses uitvoert allerlei soorten informatie: analyse op klasniveau en schoolniveau, maar ook trendanalyses. Maar ook die overzichten worden vaak in sterkere mate gebruikt voor de rapportage en verantwoording dan voor onderwijsverbetering. "De lerende school en de onderzoekende cultuur zijn nog niet zo zichtbaar" constateert één van de experts.

Feedback aan leerlingen is nog niet zo erg gebruikelijk, afgezien van feedback in de vorm van een toets- of rapportcijfer, dan wel een krul bij hun goede antwoorden. Wanneer gewerkt wordt met het effectieve instructiemodel krijgen leerlingen via het nabespreken van de lessen en toetsen wel feedback op bijvoorbeeld veel voorkomende problemen die de leerkracht tegengekomen is. Leerlingen ontvangen meestal geen vergelijking van hun eigen prestaties met die van andere leerlingen. Wat overigens niet wil zeggen dat kinderen niet heel goed van elkaar weten wie er wel of niet goed is in bijvoorbeeld rekenen. Toch zijn er wel mogelijkheden om leerlingen meer zelfsturend te laten leren, deze worden ook in enkele methodes aangeboden. (Zelf een test doen, zelf nakijken, zelf beslissen: doorgaan of nog extra oefenen op...). Op enkele scholen (onder andere waar gewerkt wordt met het Montessori Kindvolgsysteem) krijgen leerlingen feedback via (evaluatie)gesprekjes met de leerkracht. De meeste experts geven echter aan dat er op dit punt 'nog een wereld te winnen is' in het basis-onderwijs.

3.3.3 Effecten

In de gesprekken is de experts de vraag voorgelegd welke effecten van meetgestuurd onderwijs zij in de praktijk zien optreden. Daarbij is meestal ingegaan op de tweedeling: zijn er effecten zichtbaar op het terrein van onderwijskwaliteit en schoolontwikkeling, en zijn er effecten als het gaat om de verbetering van de leerprestaties voor taal- en rekenen?

Opvattingen van experts over meetgestuurd onderwijs

De meeste experts die aan het woord kwamen, met name uit de kring van schoolbegeleiders, zien positieve effecten van meetgestuurd onderwijs op die scholen waar het wordt toegepast. Er lijkt daarbij sprake van een duidelijke link: het kijken naar de resultaten van leerlingen –op groeps- en schoolniveau, maar ook op individueel niveau– leidt tot meer bewustwording van (het belang van) de eigen onderwijskwaliteit. Die bewustwording kan aanleiding zijn tot verschillende vormen van schoolontwikkeling en dat leidt vervolgens tot verbetering van de prestaties van de leerlingen.

Deze experts melden dat meetgestuurd deze positieve effecten heeft/kan hebben doordat:

- scholen hierdoor nauwkeuriger kunnen signaleren waar in het onderwijs op de school nog zwakke plekken zitten,
- het scholen beschermt tegen ongewenst afzakken van het niveau, omdat zij hun eigen resultaten kunnen vergelijken met landelijke gegevens, en/of kunnen afzetten tegen een referentieniveau,
- scholen die merken dat zij zwak presteren, meer gerichte maatregelen zullen gaan nemen om hun onderwijskwaliteit te verbeteren.

De experts noemen ook een uitgebreide reeks voorbeelden van kwaliteitsverbeteringen die zij zien optreden wanneer scholen werken op basis van metingen:

- Er wordt meer onderwijstijd besteed aan de basisvaardigheden
- Inzicht in resultaten leidt tot een betere differentiatie in het onderwijsaanbod en geplande leertijd
- Keuzes voor nieuwe methodes worden beter gefundeerd, waarbij meer wordt gelet op aspecten als aansluiting bij de leerlingenpopulatie en de onderwijsvisie. Dat leidt bovendien tot betere invoeringscondities, omdat ook teamleden beschikken over kennis en inzicht over de nieuwe werkwijze en gemotiveerd zijn voor het bereiken van het doel: verbetering van de leerprestaties.
- Er ontstaat een bewuster beleid gericht op verbetering van de kwaliteit van het eigen onderwijs in de vorm van het inzetten van middelen, visieontwikkeling en deskundigheidsbevordering.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

- Wanneer er in een team duidelijkheid is over gezamenlijke doelen, leidt dat tot verbetering van de doorgaande lijn, tot reflectie op vaardigheden van de leerkrachten en uitwisseling daarover, zowel intern als extern.

De deskundigen zijn het erover eens dat voor het verbeteren van de leerprestaties vooral het primaire proces belangrijk is: leerkrachten moeten weten wat een leerling al wel of nog niet beheerst en het onderwijsaanbod daarop kunnen afstemmen. Een belangrijke voorwaarde daarbij is dat leerkrachten in staat zouden moeten zijn om zelf selecties te maken uit de leerstof, zowel voor de snelle als voor de zwakkere leerlingen. Over de vraag of leerkrachten in voldoende mate over de kennis en vaardigheden daarvoor beschikken, is men niet altijd zeker.

3.3.4 Draagvlak

In hoeverre schoolbestuur, schoolleiding/management en leerkrachten de waarde van meetgestuurd onderwijs inzien, lijkt nog erg per school te verschillen. Er zijn scholen die extreem veel toetsen, die je soms zelfs moet afremmen en moet laten zien dat niet alles nodig of zinvol is, volgens de experts. Er zijn ook scholen –bijvoorbeeld scholen die veel leerlingen hebben met wie alles ‘vanzelf’ al wel goed gaat- die zich liever meer richten op creativiteit en vrijheid. Voor ouders is dat dan een overweging om juist die school te kiezen, maar zwakke leerlingen hebben het er over het algemeen niet zo best. Ook het Cito signaleert dat de motivatie van scholen heel verschillend kan zijn. Het Cito biedt cursussen aan over het leren hanteren van het Cito-LVS en over schoolzelfevaluatie. Sommige cursisten zijn enthousiast en melden zichzelf aan omdat ze intrinsiek gemotiveerd zijn, andere worden gestuurd door hun bestuur, of komen ertoe vanwege een slecht inspectierapport.

Enkele van de deskundigen signaleren overigens een toename van de belangstelling en bereidheid: *“Het lijkt erop dat het draagvlak voor het bijhouden, monitoren en evalueren op basis van leerlingenvorderingen groter is geworden. Het meekijken in de keuken van het onderwijs wint langzaam terrein.”*

De indruk bij de inspectie is echter, anders dan die van de meeste andere experts, dat meetgestuurd onderwijs niet zo erg leeft op scholen. Regelmatig evalueren van de leerprestaties en het eigen onderwijs is volgens de inspectie zeker nog niet algemeen

gangbaar. Scholen zien nog in onvoldoende mate in dat ‘leren’ en ‘leuk’ in het onderwijs heel goed kunnen samengaan. In lijn met deze gedachte is ook de volgende observatie van twee onderwijsadviseurs: *“Meetgestuurd onderwijs vraagt een lerende organisatie, waarin de professional het beste wil bereiken, en de ambitie heeft alle leerlingen op een zo hoog mogelijk niveau te laten uitstromen. De houding op de meeste scholen is echter nog die van de ‘blue print professional’, die zo goed mogelijk alles uitvoert.”*

De houding van leerkrachten en schoolleiders lijkt vooral afhankelijk van een aantal condities (die in paragraaf 3.5 uitgebreider besproken zullen worden). De experts signaleren bij scholen die net beginnen met systematisch volgen van de vorderingen van de leerlingen klachten over dat het veel werk met zich meebrengt vanwege de organisatie van de toetsafnames en het invoeren van de gegevens. Ook is er soms sprake van weerstand die voorkomt uit angst om zo hard op papier te zetten wat het ontwikkelingsniveau van het kind is vergeleken met leeftijdgenoten. Wanneer echter voor leerlingen en leerkrachten duidelijk is ‘wat je eraan hebt’, als het kijken naar resultaten goed is ingebed in het schoolklimaat, als men eraan gewend is en ziet dat het bijdraagt tot betere leeropbrengsten, dan vergroot dat het draagvlak voor meetgestuurd onderwijs enorm. Eén van de schoolbegeleiders licht dat als volgt toe: *“Bij een goed klassenmanagement, en een goede manier om de vorderingen van de leerlingen te volgen, zie je vaak de leerlingen opfleuren. Dat geldt dan niet alleen voor hun taal- of rekenprestaties maar ook voor de zelfstandigheid, het zelfvertrouwen en gedrag van de leerlingen. Als het zover is dat dit soort dingen zichtbaar wordt, dan worden de betrokkenen vanzelf positief.”*

3.3.5 *Conditie*

In de interviews is de experts gevraagd naar wat zij de belangrijkste condities vinden voor (de invoering van) meetgestuurd onderwijs. Die condities kunnen betrekking hebben op verschillende aspecten van het onderwijs: de competenties van schoolleiders, intern begeleiders en leerkrachten, de schoolcultuur/het draagvlak binnen de school en de externe ondersteuning.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Voorop staat volgens de experts dat alle betrokkenen het belang van het goed kijken naar leeropbrengsten moeten inzien. Dat geldt ook voor de besturen: zij moeten interesse hebben in de resultaten van de scholen en de directies daarnaar vragen. Het is belangrijk dat het gehele team het ambitieniveau voor het bereiken van goede prestaties op het gebied van taal/lezen en rekenen onderschrijft, maar eveneens dat het bestuur dat ondersteunt via de besteding van middelen in personeelsbeleid en taakbelasting. (Het komt voor dat besturen liever inzetten op actuele thema's zoals de gezonde school of het nieuwe leren, en de kwaliteit van het onderwijs in de basisvakken overlaat aan de inspectie).

Bovendien vraagt meetgestuurd-onderwijs een duidelijke context: de visie en lange termijndoelen van de school moeten 'staan', die moeten de richtlijn vormen voor de interpretatie van de uitkomsten. Het werken met een kwaliteitsmodel is het meest ideaal. Daarin is een tijdsplan opgenomen voor vernieuwingen, doelstellingen en prioriteiten zijn vastgelegd, er is een vergaderstructuur waarin bepaalde onderwerpen systematisch aan de orde komen, en waar ook ruimte is voor uitwisseling van ervaringen, het trekken van conclusies en vaststellen van vervolgstappen.

De meest genoemde conditie is daarnaast de deskundigheid in het 'omgaan met cijfers'. De ib-er, maar ook de schoolleider moeten vaardig zijn in het analyseren van gegevens op groeps- en schoolniveau, in het interpreteren van de uitkomsten en daarbij de context in het oog kunnen houden. Zij moeten bovendien inzicht hebben in het 'vak' en dus ook de consequenties voor de inhoud van het onderwijs kunnen overzien. Volgens een enkele deskundige moeten ook de leerkrachten ter zake kundig zijn. Zij moeten vanzelfsprekend de methodes kennen en hun leerlingen, maar bovendien ook de toetsresultaten kunnen analyseren en interpreteren en hun onderwijs daarop kunnen aanpassen. Andere deskundigen geven aan dat vooral de ib-er (en eventueel de schoolleider) in staat moeten zijn de verzamelde gegevens en analyses inzichtelijk te presenteren.

Het overzichtelijk presenteren van de resultaten en het bespreken ervan door de ib-er (en/of schoolleider) vraagt vervolgens goede coachingscompetenties. De schoolleider of ib-er moet zonder waardeoordelen met de teamleden samen kunnen bekijken welke resultaten er liggen, en wat er beter kan. Dat is des te belangrijker omdat het bedrei-

Opvattingen van experts over meetgestuurd onderwijs

gend kan zijn voor een leerkracht wanneer een groep niet goed scoort, terwijl de oorzaak daarvoor niet altijd bij de leerkracht ligt.

Een open schoolcultuur en onderlinge communicatie binnen het team bevorderen het gezamenlijk nadenken over onderwijsverbeteringen. Ook daarin speelt de schoolleider een belangrijke rol. Hij/zij moet leidinggeven aan verandering en daarbij betrokkenheid van alle teamleden weten te bereiken.

Belangrijk, en vaak een bottleneck, voor meetgestuurd onderwijs, is verder dat leerkrachten in staat moeten zijn om te differentiëren in hun onderwijsaanbod op basis van de gemeten leeropbrengsten. Dat vraagt goede klassenmanagementvaardigheden. Maar voor een goede uitvoering van differentiatie komt de leerkracht soms tijd en handen te kort. Bovendien bevatten de onderwijsmethoden te weinig differentiatiemateriaal.

De ervaring leert dat (nog) lang niet op alle scholen bij directies, ib-ers of zorgcoördinatoren voldoende vaardigheid aanwezig is om gegevens te interpreteren en inzichtelijk te presenteren voor de teamleden. Om die reden is over het algemeen externe ondersteuning gewenst. Die externe ondersteuning zal naar de inschatting van de experts voorlopig nog wel nodig blijven. Op de pabo's is er volgens de experts weinig of geen aandacht voor analyse van onderwijsresultaten en het omzetten van de evaluatie daarvan naar beleid.

De inschatting van de inspectie en het Cito is dat schoolbegeleiders op dit vlak zelf ook (te) weinig expertise in huis hebben om de gesignaleerde tekorten aan kennis en vaardigheden van scholen te kunnen verhelpen. Eigenlijk lijken de Cito-cursussen het enige aanbod op dit vlak te vormen. De schoolbegeleiders en onderwijsadviseurs van landelijke instellingen die wij benaderden in het kader van deze rondvraag, bieden de genoemde vormen van begeleiding echter juist wel. Ook zij zijn van mening dat de rol van externe ondersteuners belangrijk is om het proces van schoolontwikkeling te ondersteunen en op gang te houden. De inhoud van die rol wordt als volgt door een van de experts verwoord: *“Het gaat dan om iemand die helpt te kiezen welke gegevens verzameld worden aansluitend op de doelstellingen, die de gegevens kan ordenen en die helpt te denken over de uitkomsten. Hij/zij moet ook voldoende overzicht hebben om een inhoudelijke push te geven wals de ontwikkelingen dreigen tot stil-*

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

stand te komen. Het is voor teamleiders soms lastig te zien waar het aan schort: als je zelf in de soep drijft, dan zie je de ballen niet.”

Verdere condities die genoemd worden kunnen ook worden opgevat als aanbevelingen: er moet meer aandacht komen voor statistiek en algemene onderzoeksvaardigheden op de initiële opleidingen en schoolleideropleidingen. Er moeten bovendien meer cursussen komen zoals het Cito die nu aanbiedt over het gebruik en de mogelijkheden van het leerlingvolgsysteem, bijvoorbeeld via scholing van de ib-netwerken via de samenwerkingsverbanden van WSNS. Ook scholing voor pabo-docenten, en schoolbegeleiders in meetgestuurd onderwijs wordt zinvol geacht. Daarnaast kan de mogelijkheid worden overwogen om te streven naar het aanstellen van een toetsspecialist per school. Het mogelijke nadeel daarvan is wel dat de basisdeskundigheid van leerkrachten op de achtergrond raakt en de verantwoordelijkheid te makkelijk wordt afgeschoven.

3.3.6 Neveneffecten en risico's

De experts zien ook enkele mogelijke risico's of ongewenste neveneffecten van meetgestuurd onderwijs:

- Het kan zijn dat scholen verkeerde conclusies trekken uit de verzamelde cijfers of blijven steken in een te mechanistische aanpak met te weinig aandacht voor de context die voor de leeropbrengsten van belang is. En wanneer er sprake is van gebrek aan inzicht bij leraren of schoolleiding in wat de toetsresultaten nu eigenlijk precies wel en niet zeggen, zal meten en evalueren niet bijdragen aan verbetering van het onderwijs.
- Er kan een sfeer ontstaan van 'teaching to the test', waarin scholen of leerkrachten geneigd raken zoveel mogelijk de hele methode erdoor te jagen om ervoor te zorgen dat in elk geval maar alle stof behandeld is. Een overwaardering van de toetsresultaten en een te sterke prestatiegerichtheid kan ten koste gaan van de aandacht voor het leerproces zelf.
- Teveel toetsen gaat ten koste van de onderwijstijd. Het risico van teveel toetsen ontstaat wanneer scholen oude systemen niet voldoende durven loslaten wanneer er nieuwe voor in de plaats komen, wanneer het alleen gebeurt 'omdat het

Opvattingen van experts over meetgestuurd onderwijs

moet van de inspectie', of wanneer er vanwege afspraken tussen schoolbesturen (bijvoorbeeld per gemeente) op alle scholen dezelfde toetsen moeten worden afgenomen. In die situaties zijn de gegevens namelijk niet voor alle scholen ook echt informatief, en zijn er dus geen effecten op de onderwijsopbrengst.

- Meetgestuurd onderwijs zou kunnen leiden tot een afrekencultuur: ofwel op bestuurlijk niveau van scholen die zwakker presteren, ofwel op schoolniveau van leerkrachten die minder goede resultaten weten te behalen. In dat geval raakt het doel, verhogen van de kwaliteit van het onderwijs, uit beeld.
- Een (te) sterke nadruk op toetsen en prestaties kan leiden tot pressie van ouders op leerkrachten om het kind naar een volgend niveau te laten doorgaan, en tot prestatiedruk bij kinderen, met als gevolg wedijver, faalangst, het verlies van plezier in leren of welbevinden op school.
- Voor het Montessori-onderwijs botst een werkwijze waarin op vaste momenten in het schooljaar toetsen moeten worden afgenomen (zoals met het Cito-LVS) met de onderwijsvisie. Het risico is dat scholen of leerkrachten de neiging krijgen om bepaalde leerstof dan toch maar alvast te gaan behandelen voordat de toetsperiode aanbreekt, ook als dat niet aansluit bij het verloop van de ontwikkeling van het kind op dat moment.

De genoemde risico's houden allemaal verband met de al in 3.3.5 besproken condities: meetgestuurd onderwijs vraagt deskundigheid en vaardigheden van schoolbesturen, schoolleiders, ict-ers en leerkrachten, die momenteel nog geen 'gemeengoed' zijn in het onderwijs. Vooral een goede (externe) begeleiding van scholen in het op deze manier werken aan onderwijsverbetering is daarom volgens de experts van belang. Geen van de experts geeft overigens aan in de praktijk te ervaren dat de genoemde risico's en neveneffecten in omvangrijke mate voorkomen.

3.3.7 Aanbevelingen van de experts

De slotvraag aan de experts luidde: heeft u nog aanbevelingen, voor het onderwijsveld en/of voor het beleid? Dat levert een uitgebreide lijst op met waardevolle ideeën, die hieronder kort zijn samengevat:

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

- Ondersteuning van scholen in meetgestuurd onderwijs en onderwijsontwikkeling is noodzakelijk: doelgericht verzamelen van gegevens, het goed analyseren en interpreteren ervan, kritische vragen leren stellen over het eigen onderwijs, het bieden van expertise waar dat onderwijsinhoudelijk nodig is, en veel tijd en inzet investeren in professionele ontwikkeling daar waar die het meest effectief is: op het niveau van de leerkracht in de groep.
- Schoolleiders (en ib-ers) moeten toegerust zijn of worden om hun teamleden te coachen bij het interpreteren van de gegevens, en het toepassen van vernieuwingen.
- De deskundigheid op het gebied van meetgestuurd onderwijs moet bevorderd worden, niet alleen in de scholen, maar ook daar omheen. Gedacht wordt aan cursussen gericht op pabo-docenten over opbrengstgericht werken, aandacht voor statistiek en evaluatietechnieken in de initiële opleidingen, en via de ib-netwerken en schoolleideropleidingen, deskundigheidsbevordering bij de schoolbesturen als verantwoordelijken voor de onderwijskwaliteit.
- In aansluiting hierop zou ook door het Cito een vernieuwd instructieboekje geschreven moeten worden met aanwijzingen en richtlijnen voor het gebruik en de interpretatie van toetsgegevens.
- Stimuleren van ontwikkeling van de mogelijkheden om leerprestaties te volgen en te meten. Er zijn alternatieven nodig voor enkele van de toetsen uit het Cito-LVS die verouderd zijn (leeswoordenschat), of die niet goed aansluiten bij wat je zou willen weten (Ordenen). Er moet meer ruimte komen om ook andere instrumenten te gebruiken dan alleen de Cito-toetsen, als dat beter aansluit bij het onderwijsconcept, of als langs een andere weg de vragen die voor de school relevant zijn beter of eenvoudiger beantwoord kunnen worden.
- Ook uitbreiding van de mogelijkheden om gegevens in te voeren in het leerlingvolgsysteem is daarbij van belang: niet alleen eindscores op de Cito-toetsen, maar ook methodegebonden toetsgegevens etc. en niet alleen ontwikkeling van cognitieve vaardigheden, maar ook de sociaal-emotionele ontwikkeling. Een digitaal leerlingvolgsysteem zou bovendien (zoals het MKVS nu

doet) didactische aanwijzingen moeten kunnen opleveren over extra uitleg, geschikte oefeningen voor de competenties waarop het kind nog achterloopt etc.

Kernpunt voor de experts is dat het van groot belang is aan te sluiten bij de vragen, en de aanpak van de school zelf, de dagelijkse praktijk. Daarnaast vraagt meetgestuurd onderwijs vooral een mentale omslag in de richting van professionele ontwikkeling, opbrengstgericht werken. Het is van belang dat de leerkrachten zelf de effectiviteit van meetgestuurd onderwijs ervaren. Het aanstellen van een (in- of externe) toetspecialist lijkt minder geschikt, omdat dan verantwoordelijkheden en competenties bij die persoon of instelling blijven liggen, en het benutten van data voor het bijsturen van onderwijs niet zonder meer zal plaatsvinden.

3.4 Resultaten inventarisatieronde pabo's

Het ging bij de bevraging van de pabo's over de aandacht die de initiële opleidingen besteden aan meetgestuurd onderwijs om een korte inventarisatieronde. Uiteindelijk is bij zeven pabo's informatie verkregen van in totaal tien respondenten. De vragenlijst is drie maal beantwoord via een telefonisch interview, de overige zeven vragenlijsten zijn door de respondenten in de digitale versie ingevuld en per e-mail teruggestuurd. De meeste respondenten zijn zowel docent als coördinator (vier keer op het gebied van taal/Nederlands, twee keer op het gebied van rekenen/wiskunde, twee keer op het gebied van pedagogiek/onderwijskunde. Twee respondenten vervullen een functie als programmamanager.⁴

De resultaten laten een aantal overeenkomsten zien tussen de verschillende opleidingen, daarnaast lijken er ook duidelijke verschillen te zijn. Dat betekent dat de oplei-

⁴ De letterlijke benamingen van de functies van de respondenten verschillen per opleiding en luiden: coördinator Nederlands, sectieleider en docent Taal, coördinator kernfase en docenten rekenen/wiskunde, programmamanager, vakgroepvoorzitter Nederlands, docent rekenen/wiskunde didactiek, domeinexpert Taal en Taaldidactiek, docent pedagogiek en onderwijskunde, docent O&P derde fase/schooladviseur, voorzitter programmacommissie.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

dingen waarover wij geen verdere informatie hebben andere accenten zouden kunnen leggen. De hieronder gepresenteerde resultaten moeten daarom vooral worden gelezen als een eerste verkenning.

Toetsen afnemen, observaties uitvoeren om het taal-rekenniveau van leerlingen te bepalen

Alle pabo's geven aan dat zij aandacht besteden aan (het afnemen van) toetsen en/of uitvoeren van observaties bij individuele leerlingen gericht op het bepalen van het taal/rekenniveau van de leerlingen. Waar de respondenten een toelichting op hun antwoord gaven, blijkt dat het thema toetsen vooral voorkomt vanaf het derde leerjaar van de opleiding: *“In het 3^e leerjaar leren studenten zowel voor rekenen als taal een handelingsplan op te stellen en uit te voeren. Dit bestaat uit het signaleren, diagnosticeren, remediëren en evalueren.”* Eén respondent geeft aan dat studenten weliswaar attent gemaakt worden op het bestaan van (taal)toetsen, maar er niet zelf mee oefenen. Op deze opleiding ligt het accent al vanaf het eerste jaar -volgens een collega-respondent van dezelfde pabo- op het observeren en het volgen van de ontwikkelingen van kinderen.

Negen van de tien respondenten geven aan dat in het studieprogramma *methodeonafhankelijke ofwel genormeerde toetsen* aan bod komen, meestal zijn dat tevens de toetsen uit het leerlingvolgsysteem. Als voorbeelden van de toetsen die studenten leren gebruiken, worden genoemd: PI-dictee, Brus, AVI, DMT, HOREB, LVS/CITO, PDO-formulieren.

Vaak maar niet altijd⁵, besteden de opleidingen tevens aandacht aan *methodegebonden toetsen* en aan het uitvoeren van observaties aan de hand van een registratieformulier. Daarbij gaat het vanzelfsprekend om een wisselwerking tussen de theorie op de opleiding en de uitvoering in de praktijk tijdens de stage. Eén respondent geeft daarover aan: *“De studenten leren omgaan met de toetsen uit het leerlingvolgsysteem. Ook toetsen de studenten leerlingen op de praktijkschool aan de hand van de Lees- en*

⁵ In elk geval op vijf van de zeven pabo's die deelnamen aan de inventarisatie.

spellingtoetsen van Van der Leij en Vieijra ten behoeve van het behandelingsplan dat ze gaan maken.”

Tot slot geven drie opleiders expliciet aan dat studenten leren *diagnostische gesprekken* te voeren. Eén van de respondenten licht dat als volgt toe: “*Studenten leren een interactieve nabespreking te houden onder andere om in beeld te krijgen welke oplossingsstrategieën leerlingen toepassen. Studenten leren hoe ze een beeld van een kind kunnen krijgen tijdens een leergesprek, of in het samen werken aan een gekozen activiteit/spel, of in het analyseren van zijn kladblaadje of uitwerking. Tevens door observaties tijdens de toetsen en het maken van rekenwerk.*”

Nakijken en scoren volgens richtlijnen

Zeven van de tien respondenten⁶ geven aan dat de studenten ook leren om zelf toetsen na te kijken en te scoren volgens de handleiding waarbij zij dan bijvoorbeeld ook moeten letten op het type fouten dat een leerling maakt. Enkele respondenten lichten hun antwoord toe. Dat levert de indruk op dat het oefenen in het nakijken van toetsen volgens scoringsrichtlijnen meestal gebeurt aan de hand van strak geformaliseerde, veelgebruikte instrumenten zoals de AVI-toetsen of het PI-dictee. Eén respondent geeft aan dat het weliswaar de bedoeling is dat studenten leren om toetsen na te kijken volgens de handleiding, maar dat de mate waarin dat gebeurt afhankelijk is van de student zelf én van de stageplek.

Soms oefenen studenten ook met observaties, een diagnostisch gesprek of authentiek schriftelijk werk van leerlingen. In het laatste geval geeft de respondent aan dat het erom gaat dat tweedejaars studenten zich proberen te verplaatsen in de oplossingsmanieren die kinderen hanteren bij hun rekenwerk. Daarbij lijkt het accent vooral te liggen op het ontwikkelen van inzicht in het denken en leren van kinderen, en wat minder op ‘meten’ of op scoren volgens vaste richtlijnen.

⁶ Deze respondenten vertegenwoordigen zes van de zeven in het onderzoek bevraagde pabo's. Van de twee opleidingen van wie we meerdere reacties ontvingen, geven de respondenten niet altijd hetzelfde antwoord. Er lijken binnen de opleidingen verschillen te zijn: wel bij taal, niet bij rekenen/wiskunde, of omgekeerd, dan wel afhankelijk van de gekozen minorrichting.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Analyseren en interpreteren

Op alle pabo's leren de studenten iets over het analyseren en interpreteren van toetsresultaten op het niveau van individuele leerlingen. Ook maken de studenten op elke opleiding kennis met het werken met een leerlingvolgsysteem.

Veel minder komt het analyseren en interpreteren van toetsresultaten op groepsniveau aan bod, dat lijkt op slechts drie van de zeven betrokken opleidingen opgenomen in het studieprogramma. Eén van de respondenten schrijft: *"Een belangrijk thema in de opleiding is dat studenten de opbrengst van hun onderwijs onderzoeken, kijken of doelen gehaald zijn, door wie en op welk niveau."* Een docent van een andere pabo geeft aan: *"Er wordt nagegaan wat er bij het leerlingvolgsysteem en bij methodegebonden toetsen wordt getoetst, wat die gegevens betekenen. Als ze bijvoorbeeld constateren dat er bij een toets over breuken veel fouten worden gemaakt, waar ligt dat dan aan? Komt dat door de methode, komt dat door hun onderwijs? Studenten leren een analyseplan maken en na te gaan wat de resultaten van de toetsen betekenen voor hun onderwijs."*

Afstemmen van het onderwijsaanbod

Het afstemmen van het onderwijsaanbod op de gevonden resultaten is op een bepaalde manier vanzelfsprekend: studenten leren dat hun onderwijsactiviteiten en leerdoelen altijd moeten aansluiten bij de beginsituatie van de leerlingen. *"Dat betekent dat zij die beginsituatie eerst moeten vaststellen. Dat kan via toetsgegevens, maar ook door in de warming up voor een activiteit 'oude kennis op te halen'"*, aldus één van de respondenten. *"In het begin van de opleiding stellen studenten nog weinig individuele verschillen vast tussen hun leerlingen, maar gaandeweg groeien zij toe naar differentiatie en onderwijs op maat."* De opleidingen geven aan dat studenten leren een groepshandelingsplan of een individueel handelingsplan te maken en ermee te werken, al is men daarover meestal ook tamelijk bescheiden. De studenten hebben immers nog niet heel veel ervaring en daardoor hebben zij nog weinig zicht op wat kinderen al wel en nog niet kunnen: *"Studenten moeten een handelingsplannetje voor een kind opstellen naar aanleiding van een diagnose."* Of, nog voorzichtiger: *"Het kan niet dat studenten wel eens even aan leerkrachten en ouders duidelijk zullen ma-*

Opvattingen van experts over meetgestuurd onderwijs

ken hoe een en ander bij een kind zou moeten worden aangepakt. In de praktijk adviseren we hen dan ook om hun handelingsplan te maken voor een kind dat zij goed kennen, en dat niet zulke ernstige problemen heeft.” Ook geeft één van de respondenten aan dat een (be)handelingsplan bijvoorbeeld in samenwerking met de mentor tot stand komt. Het afstemmen van het onderwijs op de onderwijsbehoeften van de kinderen, adaptief onderwijs, het ontwerpen van een gedifferentieerd onderwijsaanbod, lijken duidelijke aandachtspunten in de opleidingen (cq tijdens de stages). Daarbij leren de studenten -in elk geval op één van de opleidingen- ook expliciet om “de onderwijsbehoeften van kinderen te clusteren om een haalbaar groepsplan te maken.”

Plaats van meetgestuurd onderwijs binnen de opleidingen

Kennis over het meten van vorderingen en/of het niveau van leerlingen op het gebied van taal en rekenen leren studenten met name in hun praktijkstages. Er is daarbij meestal wel sprake van lesstof op de pabo en van stageopdrachten die daarbij aansluiten.

Het bepalen van de vorderingen en het niveau van de leerlingen is basisstof voor alle studenten. Een deel van de studenten gaat via hun minor, keuzeprofiel of afstudeeronderzoek dieper op deze stof in. Het gaat dan bijvoorbeeld om afstudeeronderzoek over de problematiek van toetsen (zo’n 3% op de betreffende opleiding), om een Minor ‘Special Education Needs’ (25%) waarin studenten leren hoe zij kunnen signaleren dat leerlingen uitvallen zowel omlaag als omhoog), of een minor ‘Omgaan met verschillen’ (ongeveer 50%) waarbinnen studenten zich verdiepen in individuele verschillen, onderwijs op maat, zorgverbreding, groepsplannen enzovoort.

Aandacht voor kwaliteitsbeleid binnen de opleidingen

Op één van de opleidingen wordt het onderwerp kwaliteitsbeleid “*niet meer dan even aangekaart*”, op enkele andere opleidingen geldt dat een deel van de studenten iets krijgt aangeboden op het gebied van kwaliteitsbeleid, afhankelijk van de gekozen minor en/of een eigen onderzoek. Het gaat dan bijvoorbeeld om het beschrijven van het taalbeleid van de school en het ontwikkelen van een taalbeleidsplan. Op vier opleidingen zijn er opdrachten voor alle studenten die liggen op het terrein van kwaliteitszorg, zoals: het schrijven van een soort schoolrapport aan de hand van de

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

inspectiecriteria om te leren naar kwaliteit te kijken, een module ‘Verder kijken’ over in- en externe kwaliteitszorg, het in kaart brengen van het zorgsysteem, of als onderdeel van een -tien weken durend- thema ‘schoolorganisatie’.

3.5 Tussentijdse conclusies

Terugkomend op de onderzoeksvragen leiden de gesprekken met experts tot de volgende conclusies.

Onderzoeksvraag 2

Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?

Vrijwel alle scholen beschikken over objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van hun leerlingen op basis van toetsen uit het Cito-leerlingvolgsysteem. Daarnaast maken leerkrachten vrij veel gebruik van methodegebonden toetsen. Op sommige Montessori-scholen is een volgsysteem in gebruik waarmee de ontwikkeling van leerlingen via systematische observaties kan worden gerelateerd aan een ontwikkelprofiel.

Onderzoeksvraag 3

In hoeverre en hoe worden deze gegevens gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?

De verzamelde gegevens gebruiken de scholen hoofdzakelijk voor het volgen van de prestaties op leerlingniveau, waarbij de focus ligt bij de leerlingen die in hun ontwikkeling achterblijven (soms ook: opvallend voorlopen) ten opzichte van hun leeftijdgenoten. Bij deze leerlingen worden ook toetsen afgenomen in het kader van diagnostiek en gericht op aanpassingen in het onderwijs. In veel mindere mate vindt

Opvattingen van experts over meetgestuurd onderwijs

op basis van de gegevens ook evaluatie plaats van het onderwijs voor de hele groep of specifieke subgroepen.

Dat dit evalueren op groepsniveau veel minder gebeurt, heeft veel te maken met het ontbreken van voldoende kennis van statistische technieken en analysemethodes in de scholen. Het zijn met name ib-ers en schoolleiders die deze evaluaties uitvoeren, maar de mate waarin dat gebeurt verschilt enorm per school. Het lijkt er wel op dat scholen inmiddels meer 'evaluatiegericht' gaan denken en handelen. Vooral deelname aan grote innovatieprojecten, zoals het Onderwijskansenproject of de Taalpilots bevordert een meer resultaatgerichte manier van werken en het vooraf bepalen van (leer)doelen – ook wel: ambitieniveaus – die bereikt moeten worden.

Onderzoeksvraag 4

Welke handelingen/stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom?

Er vindt op scholen wel veel dataverzameling plaats op basis van toetsen (naar de mening van enkele van de deskundigen zelfs wel eens te veel), maar de uitkomsten inzetten voor het verbeteren van het onderwijs is duidelijk een vervolgstap die veel scholen niet zetten.

De stappen in de evaluatieve cyclus worden over het algemeen allemaal goed doorlopen wanneer het gaat om zorgleerlingen. Op groepsniveau gebeurt dat echter in veel mindere mate: daar is weliswaar sprake van het verzamelen van gegevens over het ontwikkelingsniveau van de leerlingen, maar de uitkomsten worden nog weinig ingezet voor aanpassing van het onderwijs op groepsniveau. Ze worden nog vooral gezien als een constatering daarvan (dit kind zit op B-niveau).

Het kunnen inzetten van de gegevens ter optimalisering van het onderwijs, vraagt ten eerste vaardigheden van de leerkrachten in het differentiëren van hun onderwijs, zowel praktisch (via hun klassenmanagement), als inhoudelijk (via goed inzicht in de aangeboden leerstof), als op het niveau van doelstellingen die bij (bepaalde groepen) leerlingen bereikt moeten kunnen worden. In teams waar helderheid bestaat over een gezamenlijke doelstelling om de prestaties in taal en rekenen te verhogen, waar gewerkt wordt met een kwaliteitsmodel met vaste overlegstructuren, is meer deskundig-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

heid en meer uitwisseling over de mogelijkheden daartoe. Op die scholen leidt meetgestuurd onderwijs vervolgens inderdaad vaak tot verbetering van de prestaties. Ook trendanalyses om zwakke plekken in het eigen onderwijs op te sporen gebeuren nog weinig. Soms worden ze overigens wel uitgevoerd door externe bureaus, maar door de school verder niet actief bestudeerd en benut.

Onderzoeksvraag 5

Bestaat er draagvlak bij leerkrachten en schoolleiders voor meetgestuurd onderwijs?

Er is naar de indruk van degenen die het hele scholenveld overzien nog niet zoveel draagvlak voor meetgestuurd onderwijs. De experts echter die betrokken zijn bij projecten waarin systematisch met scholen wordt gewerkt aan onderwijsvernieuwing, ervaren dat bij schoolleiders en leerkrachten die goed op de hoogte zijn van de mogelijkheden en werkwijze van meetgestuurd onderwijs ook het enthousiasme toeneemt. Zeker wanneer de positieve effecten van de werkwijze zichtbaar zijn geworden voor de leerkrachten in de klas.

Onderzoeksvraag 6

Wat zijn risico- en succesfactoren bij meetgestuurd onderwijs? Zijn er eventueel onbedoelde effecten?

Risico's zijn vooral het ontstaan van een afrekencultuur, prestatiedruk, het toetsen zonder dat aan de uitkomsten consequenties voor de inrichting van het onderwijs worden verbonden, het mechanistisch toepassen en/of ondeskundig interpreteren van de gegevens, zonder dat er voldoende oog is voor het leerproces zelf.

Onderzoeksvraag 7

Wat hebben schoolleiders en leerkrachten nodig om meetgestuurd onderwijs te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

De experts ondersteunen het beleid van de overheid om meetgestuurd onderwijs te stimuleren, zij signaleren echter enkele serieuze invoeringsproblemen.

Opvattingen van experts over meetgestuurd onderwijs

Schoolleiders en leerkrachten, maar ook schoolbestuurders hebben een prikkel nodig om het belang in te zien van het goed kijken naar leeropbrengsten. Er moeten ambitieniveaus, doelstellingen of eventueel leerstandaarden zijn waartegen de resultaten vervolgens kunnen worden afgezet. Meetgestuurd onderwijs vraagt verder om een goed model van kwaliteitszorg waarin regelmatig wordt geëvalueerd. Dat vraagt van met name de schoolleider en/of ib-er ook deskundigheid in het omgaan met en overzichtelijk presenteren van gegevens, en in het voeren van coachende gesprekken met teamleden gericht op het oplossen van knelpunten en verbeteren van de onderwijskwaliteit. Een open schoolcultuur waarin sprake is van een goede, collegiale uitwisseling en leren van elkaar werkt bevorderend. Voorts vraagt meetgestuurd onderwijs vooral van leerkrachten dat zij beschikken over voldoende inzicht in hun leerlingen en in de leerstof, en over voldoende vaardigheden op het niveau van klassenmanagement om goed te kunnen differentiëren. Daartoe zouden overigens de methoden, en de leerlingvolgsystemen meer handvaten kunnen bieden dan momenteel het geval is. De aanbevelingen die hieruit voortvloeien volgens de experts zijn de volgende: in scholen is ondersteuning nodig in het ontwikkelen van voldoende expertise op het gebied van meetgestuurd onderwijs, er is in het algemeen meer scholing nodig van diverse betrokkenen als het gaat om statistiek en evaluatietechnieken. Men denkt dan aan: pabo-docenten, ib-ers, schoolleideropleidingen, schoolbesturen, maar ook aan een overzichtelijk instructieboekje met richtlijnen voor meetgestuurd onderwijs. Tot slot is het nodig om de beschikbare instrumenten zelf te verbeteren: nieuwe toetsen waar de veel gebruikte verouderd zijn, andere mogelijkheden als methodeonafhankelijke toetsing niet zinvol is, het ontwikkelen van leerlingvolgsystemen die ook didactische richtlijnen kunnen aangeven, en in het algemeen vergroting van het draagvlak voor meetgestuurd denken in de scholen.

Aandacht voor meetgestuurd onderwijs op de pabo's

De korte inventarisatieronde over meetgestuurd onderwijs onder zeven pabo's, levert het beeld op dat toekomstige leerkrachten voor het basisonderwijs in hun opleiding leren dat het van belang is inzicht te hebben in en aan te sluiten bij het niveau van hun leerlingen. Op hun opleiding worden zij geïnformeerd over toetsen en over andere methoden –zoals observaties of leergesprekken– om een indruk te krijgen van wat

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

leerlingen weten, kunnen en/of op welke manier zij te werk gaan bij het maken van hun taal- of rekenwerk. Met name via hun stages maken zij kennis met het afnemen van zowel methodegebonden als genormeerde toetsen, en met leerlingvolgsystemen.

De belangrijke rol van de stagescholen in het leren omgaan met toetsen, observaties en resultaten van leerlingen is vanuit de opleidingen wellicht vanzelfsprekend, maar plaatst wel enige vraagtekens bij wat de studenten uiteindelijk leren. Weliswaar worden op het grootste deel van de basisscholen regelmatig vorderingen van leerlingen gemeten en geregistreerd. De resultaten worden echter (nog lang) niet altijd terugvertaald naar het eigen onderwijs, tenzij het gaat om individuele leerlingen met (opvallende) achterstanden of leermoeilijkheden.

Expliciete aandacht voor het afnemen en scoren van toetsen volgens de richtlijnen van het instrument en het analyseren en diagnosticeren vindt meestal plaats in het derde leerjaar binnen een minor(specialisatie) of afstudeeronderzoek. Interpretatie van toets- of observatiegegevens, en het afstemmen van het onderwijs daarop, leren studenten in de vorm van opdrachten om een handelingsplan te ontwerpen en uit te voeren, ofwel voor een leerling, ofwel op groepsniveau. De docenten geven daarbij aan dat dit voor veel studenten een lastige opgave is en dat de mentor op de stageplek hierin een duidelijke rol heeft.

Ook hier past een kanttekening: in het derde leerjaar hebben studenten gekozen voor hun specialisatie. Dat houdt in dat de leerstof die aan de orde komt, niet voor alle studenten gelijk is. Niet alle toekomstige leerkrachten leren –in dezelfde mate– om toetsresultaten op groepsniveau te interpreteren, hun eigen onderwijs te evalueren en/of een handelingsplan te maken op basis van toetsresultaten.

Slechts een klein deel van de opleidingen lijkt voor alle studenten ook expliciet aandacht te besteden aan kwaliteitszorg, onderwijsdoelen, aan het nagaan of deze ook worden bereikt en aan factoren die daarbij een rol spelen (het gegeven onderwijs, de gebruikte methode etc.).

Samenvattend kan geconcludeerd worden dat pabo's wel aandacht besteden aan aspecten van meetgestuurd onderwijs, maar dat de mate waarin studenten worden voorbereid op het uitvoeren ervan in de praktijk erg samenhangt met de praktijk op de stagescholen waar zij terechtkomen en met hun gekozen specialisatie.

4 Opvattingen van scholen over meetgestuurd onderwijs

4.1 Doel van het deelonderzoek

Hoe denken basisscholen over meetgestuurd onderwijs? Zijn ze bijvoorbeeld bekend met de stappen van de evaluatiecyclus? Hoe denken ze over het nut van meten en evalueren? In hoeverre doen ze aan analyse en interpretatie van leerresultaten, en richten ze zich daarbij op vragen rond individuele leerlingen of (ook) op vragen die gaan over de kwaliteit van het onderwijs op groeps- en schoolniveau? Hoe tevreden zijn scholen over hun eigen praktijk; wat kan eventueel nog beter? Deze vragen hebben centraal gestaan in het deelonderzoek waarover we in dit hoofdstuk rapporteren. De antwoorden op deze vragen zijn van belang om een inschatting te kunnen maken van 'het klimaat' waarin het beleid van het Ministerie ten aanzien van opbrengstgericht werken gaat landen.

In dit hoofdstuk beperken we ons tot informatie over meetgestuurd onderwijs volgens managers, i.c. interne begeleiders en schooldirecteuren. Het perspectief van leerkrachten komt aan de orde in de hoofdstukken 5 en 6. Hier gaat het om de mate waarin managers zich in hun schoolbeleid laten leiden door kennisname en analyse van schoolvorderingengegevens. In hoofdstuk 2 hebben we voor wat betreft de stappen binnen meetgestuurd onderwijs de evaluatieve cyclus geïntroduceerd. Die cyclus bestaat uit vijf stappen: 1. het vastleggen van doelen en standaarden, 2. het verzamelen van informatie, 3. het registreren ervan, 4. het interpreteren en 5. het nemen van beslissingen. In de gebruikelijke werkwijze op scholen hebben managers vooral met de eerste, de vierde en de vijfde stap te maken. De toetsafname zelf (stap 2) en de registratie daarvan (stap 3) gebeurt veelal door groepsleraren, zij het dat interne begeleiders niet zelden aan zorgleerlingen toetsen afnemen. We zullen aan het eind van dit hoofdstuk deze stappen als uitgangspunt nemen om de resultaten samen te vatten.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Dit deelonderzoek is vooral gericht op beantwoording van de onderzoeksvragen 3, 4 en 5. Ze luiden als volgt.

3. In hoeverre en hoe worden gegevens over taal- en rekenvaardigheden van leerlingen gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?
4. Welke handelingen/stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom?
5. Bestaat er draagvlak bij leerkrachten en schoolleiders voor meetgestuurd onderwijs?

Hoewel feitelijk al beantwoord met de gegevens uit de literatuurstudie, hebben we verder ook nog een deel van onderzoeksvraag 2 meegenomen, namelijk:

2. Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen in het primair onderwijs tot hun beschikking?

Dit biedt de mogelijkheid om de gegevens hierover uit de literatuurstudie te actualiseren.

4.2 Opzet van het deelonderzoek

Voor dit deelonderzoek was het de wens van de opdrachtgever om een representatief beeld te krijgen van de stand van zaken in het basisonderwijs. Er is afgezien van een schriftelijk survey, omdat responspercentages bij schriftelijke surveys tegenwoordig vrij laag liggen en we verwachtten dat bij het nog vrij onbekende onderwerp 'data driven teaching', of, in onze termen, 'meetgestuurd onderwijs', een schriftelijke vragenlijst vaak terzijde zou worden gelegd. Voor dit deelonderzoek was het streven om een respons te halen van 70% of hoger. Er is daarom gekozen voor een telefonische vragenlijstafname, in twee rondes:

- a. een 'quick scan' bij enkele honderden basisscholen;
- b. een opvolgingsonderzoek bij circa 20 basisscholen die blijkens de quick scan al

Opvattingen van scholen over meetgestuurd onderwijs

wat verder gevorderd zijn met de invoering van meetgestuurd onderwijs.

Quick scan; populatie en steekproeftrekking

De populatie waaruit steekproeven getrokken zijn, bestaat uit alle scholen voor basisonderwijs (ongeveer 7000) en speciaal basisonderwijs (ongeveer 300). Uit de populatie zijn drie steekproeven getrokken:

- a. een aselecte steekproef van 250 basisscholen;
- b. een aselecte steekproef van 50 scholen voor speciaal basisonderwijs;
- c. een aselecte steekproef van 50 scholen voor (speciaal) basisonderwijs die volgens de inspectie 'taalzwak' zijn.

De taalzwakke scholen zijn opgenomen als aparte subgroep omdat het denkbaar is dat scholen met relatief zwakke scores voor de basisvaardigheden meer dan andere scholen aandacht zullen besteden aan meetgestuurd onderwijs, vanwege de noodzaak om de prestaties te verbeteren. Voor taalzwakke scholen is gekozen omdat op het moment van het onderzoek voor deze groep een bestand beschikbaar was bij de inspectie (voor rekenzwakke scholen was dat nog niet het geval). Dit bestand was het resultaat van een analyse die de inspectie in 2007 heeft uitgevoerd naar het niveau van de basisvaardigheden taal in het basisonderwijs (Inspectie van het onderwijs, 2008). Op basis van deze analyse heeft 12 procent van alle scholen het predicaat taalzwak gekregen. Dat zijn scholen waar de gemiddelde score op het onderdeel taal van de Eindtoets Basisonderwijs van het Cito gedurende twee of drie van de laatste drie schooljaren (2004-2007) meer dan een halve standaarddeviatie lager is dan het gemiddelde van scholen met een vergelijkbare leerlingpopulatie. Dit betekent dat scholen gedurende twee of drie jaar tot de ongeveer 25 procent laagst presterenden van hun schoolgroep⁷ horen.

Na trekking van de steekproeven bleek dat er een aanzienlijke overlap was tussen steekproef a (basisscholen algemeen) en steekproef c (taalzwakke scholen). De ver-

⁷ Schoolgroep: de categorie waar de school is ingedeeld op basis van de samenstelling van de leerlingenpopulatie. Er zijn zeven categorieën, categorie 1 bestaat uit scholen met vrijwel uitsluitend autochtone leerlingen van niet laagopgeleide ouders, categorie 7 bestaat uit scholen met vrijwel uitsluitend allochtone leerlingen waarvan de ouders wel laag opgeleid zijn.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

deling is vervolgens enigszins aangepast, namelijk tot 200 basisscholen/niet taalzwak, 100 basisscholen/wel taalzwak, en 50 scholen voor speciaal basisonderwijs. De scholen in deze steekproeven zijn gebeld met het verzoek tot deelname aan het onderzoek. Om de deelnamebereidheid zo hoog mogelijk te maken is de vragenlijst voor de quick scan zeer kort gehouden. Er is net zo lang gebeld tot het beoogde responspercentage van 70% gehaald was. Tabel 4.1 bevat de responsgegevens.

Tabel 4.1 Steekproeven en responsgroepen; aantallen scholen (tussen haakjes percentages)

Type scholen	Aantal in steekproef	Aantal in responsgroep
a. Basisscholen	200 (100)	147 (74)
b. Scholen voor speciaal basisonderwijs	50 (100)	39 (78)
c. Taalzwakke scholen	100 (100)	74 (74)

Instrumentontwikkeling

Bij de ontwikkeling van de vragenlijst voor de quick scan is een belangrijk gegeven geweest dat de afname via de telefoon zou plaatsvinden en dat de afnameduur niet meer dan circa tien minuten mocht bedragen. Vanwege de gevraagde snelheid hebben we de vragen beperkt tot twee ontwikkelingsdomeinen, namelijk lezen en rekenen. De vragenlijst telde vier onderwerpen:

- Hanteert de school leerstandaarden?
- Is er sprake van centrale opslag van vorderingengegevens?
- Worden vorderingengegevens centraal geanalyseerd; zo ja, met welk oogmerk?
- Hoe is de opinie over meetgestuurd (opbrengstgericht) werken?

Deze vier onderwerpen zijn uitgewerkt in een zestal vragen (zie paragraaf 4.3).

Opvattingen van scholen over meetgestuurd onderwijs

Instrument voor het opvolgingsonderzoek

Voor het opvolgingsonderzoek is een vragenlijst ontwikkeld die uit twee delen bestaat. In het eerste deel is gevraagd naar een verdere toelichting op motieven en werkwijze wat betreft analyse van gegevens. De onderwerpen daarin zijn:

- de aanleiding/de reden om er mee te beginnen;
- de gebruikte gegevens (welke toetsen, instrumenten) en de wijze van registreren;
- de uitvoering van de analyses (wie) en de deskundigheid die daarvoor nodig is;
- de onderwerpen van analyse (waar wordt naar gekeken?);
- het gebruik van doelen, standaarden, handelingsplannen, ontwikkelingsperspectieven;
- het gebruik van de uitkomsten van de analyse;
- procedures (is het een staande praktijk?).

Hierbij is onderscheid gemaakt tussen analyses op groepsniveau en op individueel niveau (zie onderzoeksvraag 3).

Het tweede deel gaat over wat meetgestuurd onderwijs oplevert voor de school. Gevraagd is naar tevredenheid over de werkwijze en over de opbrengsten en naar eventuele risico's of nadelen. Verder is in dit deel gevraagd naar wat eventueel nog verbeterd zou moeten worden en naar de opinie over het beleidsvoornemen om landelijke referentieniveaus in te voeren.

Gegevensverzameling

De gegevensverzameling voor de quick scan heeft plaatsgehad vanaf juni tot en met september 2008. Verzameling vond plaats via telefonische interviews. De vragen zijn gesteld aan de functionaris die goed op de hoogte was van de toetspraktijk op school. In 56 procent van de gevallen is het gesprek gevoerd met de directeur, in 38 procent met de intern begeleider. Vijf procent van de respondenten had een andere functie (meestal adjunct directeur).

Voor het opvolgingsonderzoek zijn de gegevens verzameld in de periode oktober-november 2008. Hiervoor zijn scholen geselecteerd die gezien hun antwoorden bij de quick scan al daadwerkelijk stappen van meetgestuurd onderwijs bleken toe te passen

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

en zich bovendien bereid hadden verklaard aan een vervolgggesprek deel te nemen. In totaal zijn aanvullende telefonische interviews gehouden met 20 scholen, waaronder drie scholen voor speciaal basisonderwijs en drie taalzwakke scholen. De gesprekken vonden plaats hetzij met de directeur (10x), hetzij met de adjunct-directeur (3x) of de intern begeleider (7x). De scholen kregen van te voren de interviewleidraad digitaal toegestuurd. De gesprekken namen ongeveer een half uur in beslag.

4.3 Resultaten quick scan

In de quick scan stelden we de volgende vraag over het gebruik van leerstandaarden (stap 1 in de evaluatieve cyclus):

Heeft u op uw school voor lezen en/of rekenen vastgelegde leerstandaarden die u wilt bereiken? (leerstandaard= een afgesproken beheersingsniveau, dus wat kinderen in ieder geval moeten kunnen op een bepaald moment). Zo ja, voor welke leeftijdsgroepen?

De resultaten waren als volgt.

Tabel 4.2 Percentage scholen dat leerstandaarden heeft vastgelegd (per jaargroep, per ontwikkelingsdomein)

	Basisscholen (n = 147)		Scholen voor sbo (n = 39)		Taalzwakke scholen (n = 74)	
	Lezen	Rekenen	Lezen	Rekenen	Lezen	Rekenen
Groep 1/2	46,2	50,3	7,6	10,2	43,2	40,5
Groep 3	85,7	79,6	28,2	23,0	79,7	66,2
Groep 4	85,0	80,9	28,2	23,0	78,3	66,2
Groep 5	81,6	79,6	25,6	20,5	78,3	66,2
Groep 6	78,2	79,6	25,6	20,5	74,3	66,2
Groep 7	67,3	77,5	25,6	20,5	68,9	64,9
Groep 8	68,0	77,5	25,6	20,5	68,9	66,2

Opvattingen van scholen over meetgestuurd onderwijs

De tabel laat een paar duidelijke patronen zien. In het reguliere basisonderwijs hebben de meeste scholen vastgelegde leerstandaarden, behalve in de groepen 1 en 2. Dit geldt voor zowel lezen als rekenen. Wel is er enig verschil tussen lezen en rekenen: leerstandaarden voor lezen heeft men in de onderbouw (vooral groep 3 en 4) iets vaker dan voor rekenen, in de bovenbouw is dat juist omgekeerd. Op taalzwakke scholen zijn er wat minder vaak leerstandaarden dan op niet taalzwakke scholen, voor rekenen geldt dat wat sterker dan voor taal. Het meest opvallende verschil is dat tussen de basisscholen (al dan niet taalzwak) en de scholen voor sbo. Van de sbo-scholen heeft maar ongeveer een kwart vastgestelde leerstandaarden, dit geldt voor zowel lezen als rekenen en voor alle groepen⁸.

Over het verzamelen van informatie en het registreren daarvan (stap 2 en 3 in de evaluatieve cyclus) stelden we in de quick scan de volgende vraag:

Worden op uw school in het leerlingvolgsysteem gegevens over leerprestaties/leervorderingen van leerlingen centraal opgeslagen? Zo ja voor welke vakken?

De resultaten staan in tabel 4.3.

⁸ Het extra lage percentage bij groep 1/2 in het sbo heeft weinig betekenis, want deze leeftijdsgroep is in het sbo nauwelijks aanwezig.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Tabel 4.3 Percentage scholen dat gegevens over leerprestaties/leervorderingen centraal opslaat in het leerlingvolgsysteem (per vak /ontwikkelingsdomein)

<i>vak/domein</i>	<i>Basisscholen (n = 147)</i>	<i>Scholen voor sbo (n = 39)</i>	<i>Taalzwakke scholen (n = 74)</i>
<i>lezen</i>	<i>98,6</i>	<i>100</i>	<i>99,2</i>
<i>rekenen</i>	<i>99,3</i>	<i>100</i>	<i>99,6</i>
<i>ander vak, namelijk:</i>	<i>95,2</i>	<i>94,9</i>	<i>95,4</i>
<i>taal (spelling, woordenschat)</i>	<i>85,0</i>	<i>94,8</i>	<i>90,5</i>
<i> sociaal-emotionele ontwikkeling</i>	<i>19,0</i>	<i>30,0</i>	<i>31,1</i>
<i>alle Cito-LVS toetsen</i>	<i>17,0</i>	<i>5,1</i>	<i>9,4</i>
<i>kleutertoetsen</i>	<i>10,2</i>	<i>2,5</i>	<i>13,5</i>
<i>overig</i>	<i>12,0</i>	<i>2,5</i>	<i>14,8</i>

Het blijkt dat vrijwel alle scholen beschikken over een leerlingvolgsysteem met opslag van gegevens over leervorderingen op verschillende vakken. Nagenoeg altijd voor lezen en rekenen, zeer vaak voor overige onderdelen van het taalonderwijs (zoals spelling, woordenschat) en nog relatief weinig voor andere leerdomeinen. In de rubriek ‘overig’ gaan verschillende antwoorden schuil, maar meestal gaat over wereldoriëntatie/zaakvakken.

We zien hier verder nagenoeg geen verschillen tussen de drie groepen scholen. Wat hooguit opvalt is dat zowel sbo-scholen als taalzwakke scholen vaker dan de niet-taalzwakke basisscholen gegevens registreren over sociaal-emotionele ontwikkeling.

De analyse en interpretatie van de verzamelde gegevens is stap 4 in de evaluatieve cyclus. We stelden daarover in de quick scan de volgende vragen:

Worden op uw school de gegevens over leerprestaties lezen en rekenen centraal geanalyseerd? (centraal geanalyseerd = op schoolniveau, met gebruik van meerjarige gegevens). Zo ja, welke analyses voert u uit?

Opvattingen van scholen over meetgestuurd onderwijs

Tabel 4.4 bevat de resultaten voor de vraag naar de centrale analyse.

Tabel 4.4 Percentage scholen dat gegevens over leerprestaties/leervorderingen centraal (op schoolniveau) analyseert

	Basisscholen (n = 147)	Scholen voor sbo (n = 39)	Taalzwakke scholen (n = 74)
ja	80,3	74,4	83,8
mee begonnen/aan het oriënteren	15,6	12,8	14,9
nee	4,1	12,8	1,4

Ook bij deze vraag zien we hoge percentages 'ja'. Sbo-scholen doen wat minder aan centrale analyse dan de andere twee groepen scholen, taalzwakke scholen doen het naar verhouding het meest (zeggen met name het minst vaak 'nee').

Als scholen centraal gegevens analyseren, welke analyses voeren ze dan uit? We legden aan de scholen in de quick scan de volgende mogelijkheden voor en vroegen of ze een of meer van deze analyses regelmatig uitvoeren:

- groepsgemiddelden van jaar tot jaar bekijken en interpreteren
- groepsgemiddelden vergelijken met landelijke gemiddelden
- ontwikkelingscurves van zorgleerlingen bekijken en interpreteren
- ontwikkelingscurves van alle leerlingen bekijken en interpreteren
- vaststellen of gestelde doelen voor individuele leerlingen zijn gehaald
- vaststellen of gestelde doelen voor groepen leerlingen zijn gehaald

De antwoorden staan in tabel 4.5.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Tabel 4.5 Percentage scholen dat verschillende soorten analyses uitvoert op gegevens over leerprestaties/leervorderingen

	Basisscholen (n = 147)	Scholen voor sbo (n = 39)	Taalzwakke scholen (n = 74)
groepsgemiddelden van jaar tot jaar bekijken en interpreteren	68,7	20,5	74,3
groepsgemiddelden vergelijken met landelijke gemiddelden	60,5	5,1	66,2
ontwikkelingscurves van zorgleerlingen bekijken en interpreteren	73,5	66,6	75,7
ontwikkelingscurves van alle leerlingen bekijken en interpreteren	44,8	64,1	45,9
vaststellen of gestelde doelen voor individuele leerlingen zijn gehaald	74,8	64,1	66,2
vaststellen of gestelde doelen voor groepen leerlingen zijn gehaald	61,2	10,2	60,8

Ook deze tabel laat een paar patronen zien. In de eerste plaats geldt voor elk van de drie groepen scholen dat er vaker analyses plaatsvinden van de resultaten van zorgleerlingen dan van overige leerlingen (alle leerlingen, groepen leerlingen). In het sbo analyseert men naar verhouding vaak de ontwikkelingscurves van alle leerlingen, vergeleken met de overige scholen. Dat hoeft echter niet te verbazen: in het sbo zijn immers alle leerlingen zorgleerlingen. Daarentegen vindt in het sbo aanmerkelijk minder vaak analyse plaats van resultaten op groepsniveau; dit komt in het sbo zelfs nauwelijks voor. Ook is men in het sbo niet geïnteresseerd in vergelijking met landelijke gemiddelden. De taalzwakke en niet taalzwakke scholen laten onderling weinig verschillen zien.

Over het geheel genomen zijn de percentages vrij hoog: een meerderheid van de basisscholen zegt analyses te plegen op de resultaten van zorgleerlingen en op de resul-

Opvattingen van scholen over meetgestuurd onderwijs

taten van groepen, en een meerderheid van de sbo-scholen zegt analyses te plegen op de leervorderingen van alle individuele leerlingen.

Er zijn in totaal 89 scholen (34%) die nog andere analyses noemen die ze regelmatig uitvoeren. Het meest genoemd zijn ‘trendanalyses’. Dit gebeurde relatief vaak bij taalzwakke scholen. Hoewel trendanalyses geïmpliceerd zijn in de door ons voorgeslede categorieën (gemiddelden van jaar tot jaar bekijken, ontwikkelingscurves bestuderen zijn immers voorbeelden van trendanalyses) vond men het kennelijk nodig deze nog apart te benoemen. De reden daarvoor is vermoedelijk dat de inspectie de laatste tijd scholen aanmoedigt om trendanalyses uit te voeren.

Ook noemde men vaak analyses in het kader van de sociaal emotionele ontwikkeling van de leerlingen, overigens vooral in het basisonderwijs en niet in het sbo. Binnen het sbo noemt men als andere analyses vooral analyses aan de hand van ontwikkelingsperspectieven of didactische leeftijdsequivalenten (DLE). Dit zijn beide manieren om inzicht te krijgen in de mate waarin een leerling voorsprong/achterstand heeft, ten opzichte van de eigen prognose (ontwikkelingsperspectief) of ten opzichte van leeftijdgenoten (DLE).

Tenslotte worden analyses op methodegebonden toetsen nog genoemd als ‘andere analyse’, en analyses die geen betrekking hebben op leervorderingen maar op andere aspecten van schoolkwaliteit en schoolbeleid, zoals evaluaties met ouders.

Stap 5 in de evaluatieve cyclus is het nemen van beslissingen. Hierover hebben we in de quick scan geen vragen voorgelegd, omdat antwoorden hierop moeilijk te categoriseren zijn en er geen ruimte en tijd was voor open vragen. In de vervol ginterviews is dit wel aan bod geweest, zie 4.4.

We hebben in de quick scan nog wel gevraagd naar opinies over meetgestuurd onderwijs. Daartoe legden we de respondenten enkele stellingen voor. Er konden antwoorden gegeven worden op de range ‘helemaal niet mee eens’ tot ‘helemaal mee eens, op een vijf puntsschaal. De uitkomsten staan in tabel 4.6.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Tabel 4.6 Meningen van scholen over meetgestuurd onderwijs (gemiddelde score op vijfpuntschaal, 1= helemaal oneens..... 5= helemaal eens, en standaarddeviatie)

	Basisscholen (n = 147)		Scholen voor sbo (n = 39)		Taalzwakke scholen (n = 74)	
	score	sd	score	sd	score	sd
Verbetering van het onderwijs in taal en rekenen kan niet zonder regelmatig meten van je onderwijsopbrengsten	4,4	0,7	4,4	0,7	4,7	0,6
Op onze school is er voldoende openheid in het team om met elkaar de leeropbrengsten in ieders groep te bespreken	4,4	0,8	4,3	0,8	4,6	0,7
Op onze school is er voldoende deskundigheid om analyses te plegen op de gegevens in het leerlingvolgsysteem	4,1	0,7	4,1	0,7	4,0	0,7
Jezelf vergelijken met andere scholen met hetzelfde type leerlingenpubliek is een goede manier om te evalueren of de leerlingen presteren wat verwacht mag worden	3,8	0,9	3,4	1,4	3,8	1,0

We zien hier dat men het gemiddeld eens is met de voorgelegde stellingen. Alleen de vierde stelling krijgt wat minder steun en daarover zijn de meningen ook meer verdeeld (zie de spreiding), vooral in het sbo.

Bij de eindopmerkingen vermelden enkele zeer kleine en enkele vrije scholen en montessorischolen dat zij minder met groepsanalyses werken. De zeer kleine scholen,

omdat dat weinig zin heeft, de vrije scholen en montessorischolen, omdat dit minder bij hun uitgangspunten zou passen.

Verder merkt een aantal respondenten op dat zij bij analyses steun vanuit het bestuur krijgen.

4.4 Resultaten opvolgingsonderzoek

In de vervolginterviews hebben we groepsanalyses als eerste aan de orde gesteld. Dit deden we omdat in het licht van de doelstellingen van de Kwaliteitsagenda het analyseren op groepsniveau een belangrijk element is en uit de quick scan was gebleken dat analyses op dit niveau wat minder vaak worden uitgevoerd dan analyses rond individuele leerlingen. Twee van de drie sbo-scholen merkten echter meteen al op dat in het speciaal basisonderwijs analyses op groepsniveau niet aan de orde zijn en dat zij daar dus weinig aandacht aan besteden.. De derde sbo-school meldde deze analyses op verzoek van de inspectie wel uit te voeren, maar daar eigenlijk niet goed raad mee te weten.

Van de 17 basisscholen gaf het merendeel (11) aan dat zij weliswaar ook op groepsniveau analyseerden, maar dat toch duidelijk de nadruk lag op het analyseren van individuele leerprestaties. Voor een minderheid van de scholen lag de balans in het midden (5) of waren de groepsgegevens het belangrijkste (1).

In paragraaf 4.5 besteden wij apart aandacht aan de interviews met de scholen voor sbo, omdat hun praktijk rond meetgestuurd onderwijs in een aantal opzichten duidelijk afwijkt van die van de basisscholen. Paragraaf 4.4 bevat dus de resultaten voor de 17 basisscholen. Waar dat relevant is, worden daar de taalzwakke scholen apart vermeld.

We memoreren dat voor het opvolgingsonderzoek scholen zijn geselecteerd die blijkens de resultaten van de quick scan al daadwerkelijk bezig waren met stappen van meetgestuurd onderwijs. Het gaat om scholen die bij de vraag naar de soorten analyses tenminste vier keer hebben aangegeven dat zij het betreffende type analyse uit-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

voeren (zie de soorten analyses in tabel 4.5). Dit betekent dat alleen scholen zijn geselecteerd die zowel analyses op individueel niveau als op groepsniveau zeggen uit te voeren. Het doel van de vervolggesprekken met deze scholen was niet om opnieuw (zoals in de quick scan) een representatief beeld van voorkomende praktijken te schetsen, maar om meer in de diepte informatie te krijgen over wat scholen dan feitelijk doen, wat hun ervaringen daarmee zijn en welke condities een rol spelen bij meetgestuurd onderwijs. Om die reden is er voor gekozen om bij de presentatie van de resultaten niet het accent te leggen op kwantitatieve uitkomsten, maar op het inzicht bieden in de verschillende keuzes die scholen maken en hun motieven daarvoor. We maken daarbij gebruik van citaten van respondenten, ter verduidelijking en illustratie. Deze citaten zijn over het algemeen letterlijk en geven steeds de mening of ervaring van één respondent weer. Wel zijn ze soms enigszins stilistisch aangepast. Daar waar antwoorden goed te categoriseren waren presenteren we wel tabellen, om toch enig inzicht te geven in de mate van voorkomen van bepaalde antwoorden.

4.4.1 De analyses op groepsniveau

Aanleiding voor het analyseren van groepsgegevens

De meeste scholen volgen al zo'n jaar of tien de individuele leerlingresultaten, meestal met behulp van het leerlingvolgsysteem van het Cito. Het analyseren van groepsgegevens is veelal van recentere datum. De inspectie speelt in die ontwikkeling een sturende rol, zoals ook blijkt uit het volgende interviewfragment. *'Geregistreerd wordt er al lang, maar sinds ongeveer 1 jaar kijken we ook wat preciezer naar de groepsgemiddeldes. De Inspectie heeft daarop aangedrongen: zij misten bij ons trendanalyses en dwarsdoorsneden'*. Ook schoolbesturen spelen soms een aansturende rol, al dan niet op advies van de Inspectie.

Achterliggend doel is dan het terugdringen van achterstand. Het blijkt dat vooral scholen met relatief zwakke leeropbrengsten geneigd zijn analyses op groepsniveau uit te voeren. Scholen die presteren volgens de norm, lijken minder overtuigd van de noodzaak daarvan. De volgende uitspraak is wat dat betreft veelzeggend: *'Eerlijk gezegd doen we niet zoveel met de groepsgemiddelden, wel veel meer met de individuele gegevens. Twee jaar geleden was een slechte beoordeling van de Inspectie aanleiding om ook naar de groepsgemiddelden te kijken. Er heeft in die periode ook een direc-*

Opvattingen van scholen over meetgestuurd onderwijs

tiewisseling plaatsgevonden. Ondertussen zijn we weer een goedpresterende school geworden, en gaat er nog maar weinig aandacht uit naar de groepsmiddelen’.

Bij ongeveer de helft van de geïnterviewde scholen lag het initiatief om groepsanalyses te gaan uitvoeren bij de school zelf en niet bij Inspectie of bestuur. Dat kwam dan bijvoorbeeld voort uit de behoefte het zorgbeleid te evalueren. In die gevallen gaat het niet zozeer om het wegwerken van achterstanden, maar om het verbeteren van het rendement. *‘Het ging er niet om dat we het als school als geheel niet goed deden. Wij presteren gemiddeld voor onze categorie’.* Niet altijd was het bij deze scholen overigens duidelijk of genoemd motief sloeg op het analyseren op groepsniveau, of op het volgen van leerlingprestaties in het algemeen.

Tabel 4.7 Overzicht van de motieven voor het analyseren van groepsgegevens (het totaal aantal bedraagt 17)

Motieven	Aantal
a. Aandringen van de inspectie	6
b. Opdracht van het bestuur, onderdeel van verandering bij aantreden	2
c. Behoefte aan evaluatie zorgbeleid	5
d. Behoeften aan meer inzicht in het functioneren van de groepen	3
e. Streven naar meer kwaliteit in het algemeen	1

Gebruikte gegevens en wijze van registreren

Voor de analyse van de groepsgegevens worden in eerste instantie de toetsen van het leerlingvolgsysteem van het Cito genoemd, veelal aangevuld met toetsen voor technisch lezen. Ook instrumenten voor sociaal-emotionele ontwikkeling worden genoemd.

In het algemeen worden de gegevens digitaal geregistreerd, soms door een administratieve kracht of ICT-er, vaker door de interne begeleider en in toenemende mate ook door de leerkrachten zelf. De scholen noemen de volgende voordelen van regi-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

stratie door leerkrachten zelf: *'het is veel werk voor één persoon'*; *'de leerkrachten kunnen dan zelf de resultaten beter volgen'*.

Methodegebonden toetsen worden meestal op papier bewaard en worden als aanvullende informatie gebruikt. Op een enkele school worden ook de methodegebonden toetsen digitaal ingevoerd (in ESIS) of gaat dat in de nabije toekomst gebeuren.

Deskundigheid bij het analyseren

We vroegen de scholen wie de analyses uitvoert en of degene die de analyses uitvoert daar een speciale deskundigheid voor heeft. Op vrijwel alle scholen is analyseren de taak van de intern begeleider(s), al dan niet in samenwerking met de directie of (een enkele keer) de leerkrachten.

Op ongeveer de helft van de scholen hebben de intern begeleiders een cursus gevolgd, bijvoorbeeld een aantal studiemiddagen bij het CITO. Eén school meldt dat de intern begeleider nog een vervolgcursus heeft gevolgd, gericht op het leren omgaan met trendanalyses en zelfevaluatie.

Op andere scholen acht men speciale deskundigheid niet nodig. De analyses *'rollen er gewoon uit'*, is nogal eens de mening. *'Er worden rapporten opgevraagd op individueel- en groepsniveau en er is nog niet zo'n behoefte aan uitgebreidere statistische analyses'*. De term *'analyseren'* zorgde overigens regelmatig voor verwarring. Analyseren werd dan opgevat als *'interpreteren'* en als *'bespreken met de leerkrachten'* en leek niet als aparte stap te worden onderscheiden.

Welke vergelijkingen worden gemaakt?

Op de meeste scholen worden de gegevens binnen het schooljaar vergeleken, maar ook (longitudinaal) over de jaren. In mindere mate worden vergelijkingen gemaakt met dezelfde groep van het vorige leerjaar. Een enkele keer worden de resultaten vergeleken met die van andere scholen van hetzelfde bestuur of samenwerkingsverband.

Er is aandacht voor vragen als *'hoe ontwikkelt zich deze groep'*, *'zijn er opvallende verschillen (of overeenkomsten) met hetzelfde leerjaar vorig jaar'*. *'Zijn er vakken of specifieke onderdelen waarop een groep (of de school als geheel) onderpresteert?'* Deze vergelijkingen lijken in het algemeen vrij globaal plaats te vinden, met vooral

Opvattingen van scholen over meetgestuurd onderwijs

aandacht voor opvallende afwijkingen ten opzichte van het verwachtingspatroon. Een aantal scholen meldt preciezer te kijken naar het ene vak (of vakonderdeel) dan naar het andere, afhankelijk van bekende zwakke plekken.

Tabel 4.8 Welke vergelijkingen worden gemaakt? (het totaal aantal scholen bedraagt 17, meer soorten vergelijkingen mogelijk)

Vergelijkingen waarvan de scholen melding maken	Aantal
a. zelfde groep/zelfde leerlingen, longitudinaal (over jaren heen)	12
b. vergelijking tussen vakken	6
c. vergelijking met andere groepen (zelfde groep ander jaar, parallelgroepen)	7
d. vergelijking met andere scholen	4
e. globaal in de gaten houden/ alleen vergelijkingen als daar aanleiding voor is	4
f. alleen gegevens zorgleerlingen worden nader bekeken	3

Op de vraag of er geëvalueerd wordt aan de hand van doelen of standaarden, wordt vrijwel unaniem gemeld dat de CITO-normen worden gebruikt. De meeste scholen houden globaal in de gaten of ze niet onder het gemiddelde scoren van de categorie scholen waartoe ze behoren. Sommige scholen hebben explicietere doelstellingen: *'we gaan ervan uit dat leerlingen minimaal een C-niveau moeten halen en de groepen minimaal op B-niveau'*.

De Inspectie blijkt niet altijd genoeg te nemen met de gebruikte gegevens. Zo meldde een directeur: *'Er was een verschil tussen de LVS-uitkomsten in termen van A-E niveaus en de vaardigheidsscores waar de Inspectie naar kijkt. Daardoor dachten we dat we goed zaten, gemiddeld voor de categorie scholen waar we bij horen, terwijl we er wat onder zaten als je preciezer keek. Met name voor rekenen/wiskunde. We zijn nu alerter'*.

Niet altijd worden alle leerlingen meegerekend bij het bepalen van het groepsgemiddelde. *De leerlingen die een eigen zorglijn hebben, vallen (om pedagogisch didactische redenen) buiten de groepsanalyse. Dat zijn er 1 of een enkele keer 2 per klas'*.

Op de meeste scholen (11 van de 17 betrokken basisscholen) worden echter ook deze leerlingen in de groepsanalyse betrokken: *'Alle leerlingen worden meegenomen in de*

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

berekening van de groepsgemiddelden, ook leerlingen die een toets maken van een lager niveau. Die gegevens worden automatisch omgerekend naar het niveau van de eigen jaargroep'.

Wat gebeurt er met de analyseresultaten?

De scholen verschillen sterk in de wijze waarop zij met de resultaten omgaan. In de meeste gevallen worden de resultaten 2 (of 3) keer per jaar met de leerkrachten besproken, waarbij bespreking van de groepsresultaten en de individuele resultaten veelal hand in hand gaan. *'De uitkomsten worden besproken met de leerkrachten en zo nodig wordt het onderwijs bijgesteld. Het zijn groepsbesprekingen waarbij ook de individuele handelingsplannen een rol hebben: is het zo uitvoerbaar binnen de groep? Moet er extra formatie worden ingezet?'*

In een enkel geval bespreken de intern begeleider en de directie de resultaten samen en worden de leerkrachten er alleen bij geroepen als er onder de maat wordt gepresteerd. *'Voor zover het de groepsgemiddelden betreft, worden die alleen besproken wanneer een groep er zwak uitkomt. Door de directeur en intern begeleider met de betreffende leerkracht. Maar in het algemeen worden gewoon de leerlingresultaten besproken en zien de leerkrachten zelf wel aan de rapporten hoe het met het gemiddelde gesteld staat'.*

Een deel van de scholen meldt expliciet de groepsanalyses te gebruiken bij de evaluaties van het onderwijs, inclusief de lesmethodes. *'Nagegaan wordt op verschillende niveaus of er aanpassingen moeten komen in het aanbod. Niet alleen op individueel en groepsniveau, maar soms ook op schoolniveau. Zo komen we er langzamerhand achter dat over de hele linie de werkwoordspelling wat onder de maat is. Daar voorziet de methode kennelijk onvoldoende in. We voegen nu extra materiaal toe voor groep 7 en 8 op het gebied van werkwoordspelling'.*

Niet altijd wordt het aanbod aangepast. Soms neemt men de resultaten voor lief omdat (verdere) aanpassing niet haalbaar is of op de langere termijn niet nodig. Zo meldt een school met veel allochtone leerlingen uit achterstandsgroepen de normen te hebben bijgesteld en niet meer te streven naar AVI-3 niveau aan het eind van groep 3. Een andere school weet dat de scores voor begrijpend lezen aan de lage kant zijn en

Opvattingen van scholen over meetgestuurd onderwijs

wijkt dat aan de discrepantie tussen de toetsvragen van het Cito en die van de leesmethode. *‘Dat geeft niet. Dat betekent soms nog wat extra oefening in groep 7 en 8 tot de leerlingen gewend zijn aan de manier van vragen van het Cito. De meesten pikken dat snel op’.*

Een van de taalzwakke scholen verbindt nogal wat consequenties aan de groepsanalyses. *‘ Op directieniveau nemen we de gegevens op in het ontwikkelingsplan. Dat houdt ook in dat leerkrachten worden bijgeschoold. Op dit moment worden vier leerkrachten gecoacht en krijgen uitgebreide klasbezoeken met video-opnamen en nabespreking. Alle leerkrachten leren het directe instructiemodel en klasmanagement toe te passen om meer tijd over te houden voor leerlingen die extra instructie nodig hebben.’*

Tabel 4.9 Wat gebeurt er met de analyseresultaten? (het totaal aantal bedraagt 17, meer antwoorden mogelijk)

<i>Wat gebeurt er met de resultaten?</i>	<i>Aantal</i>
<i>a. De resultaten worden besproken met de groepsleerkrachten</i>	<i>9</i>
<i>b. De resultaten worden alleen besproken met leerkrachten bij wie sprake is van problemen</i>	<i>3</i>
<i>c. De resultaten worden (eerst) besproken met de directie</i>	<i>5</i>
<i>d. De resultaten worden (ook) besproken in het team</i>	<i>4</i>
<i>e. De resultaten worden (ook) besproken met het bovenschools management, gerapporteerd aan inspectie</i>	<i>3</i>
<i>f. De resultaten worden (ook) besproken met MR</i>	<i>2</i>
<i>g. Het onderwijs wordt aangepast (bijvoorbeeld extra aandacht voor onderdelen, aanvullend materiaal)</i>	<i>6</i>
<i>h. Leerkrachten worden bijgeschoold</i>	<i>2</i>

4.4.2 De analyse van de individuele leerprestaties

Voor de individuele analyses worden dezelfde gegevens gebruikt als voor de groepsanalyses, maar deze worden ook wel aangevuld door de methodegebonden toetsen.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Dat gebeurt om te checken of de resultaten van de methodeonafhankelijke toetsen kloppen met de dagelijkse praktijk, maar ook om na te gaan waar precies bij die leerling de hiaten zitten.

De analyses van de individuele gegevens hebben op een groot deel van de betrokken basisscholen alleen betrekking op de zorgleerlingen. Dat zijn dan meestal niet alleen de leerlingen aan de onderkant, maar ook de uitschieters naar boven. Het aantal 'zorgleerlingen' kan overigens sterk verschillen. Sommige scholen melden dat het om een paar leerlingen per klas gaat, op andere scholen kan dat oplopen tot 30% en meer. Een van de taalzwakke scholen legt de nadruk van de inspanningen om achterstand in te halen op groep 3, 4 en 5. *'In die groepen wordt voor zo'n 40 -60% van de leerlingen een handelingsplan of groepsplan gemaakt'*.

Een minderheid van de scholen meldt alle leerlingen individueel te analyseren. Dat wil niet zeggen dat vervolgens bij de besprekingen alle leerlingen ook daadwerkelijk aan bod komen. Slechts 2 scholen melden expliciet alle leerlingen ook bij de besprekingen te betrekken, zij het dat zorgleerlingen meer tijd kosten dan anderen. *'4 x per jaar passeren alle leerlingen de revue in een bespreking tussen de interne begeleider en de leerkracht. De gemiddelde leerlingen kosten weinig tijd, de zwakkere leerlingen en de echte uitblinkers wat meer'*.

Tabel 4.10 Welke individuele resultaten worden geanalyseerd en besproken? (het taal aantal bedraagt 17)

<i>Van welke leerlingen worden individuele leercurves geanalyseerd?</i>	<i>Aantal</i>
<i>a. Alleen van de zorgleerlingen</i>	<i>11</i>
<i>c. Van alle leerlingen (maar zij worden niet allemaal besproken)</i>	<i>4</i>
<i>d. Van alle leerlingen en zij worden ook allemaal besproken</i>	<i>2</i>

Consequenties

De groepsanalyses en de individuele analyses leiden tot verschillende soorten consequenties. Aan de groepsanalyses worden algemene maatregelen verbonden als meer

Opvattingen van scholen over meetgestuurd onderwijs

aandacht voor specifieke leerstofonderdelen in de klas, meer aandacht voor klasmanagement, eventueel bijscholing van leerkrachten en aanschaffen van aanvullend materiaal bij zwakke punten in de gebruikte methodes.

De individuele analyses hebben consequenties voor het toekennen van extra instructie en extra aanbod aan individuele (en groepjes) leerlingen. *‘Voor de zwakke leerlingen wordt ondersteuning aangevraagd van een speciale leerkracht. Deze geeft dan extra instructie aan individuele leerlingen of groepjes leerlingen, of geeft in de klas extra voorinstructie’.*

In de interviews kwamen de uitblinkers wat minder expliciet aan de orde. Veel scholen melden wel dat zij ook tot de zorgleerlingen horen, maar gingen minder in op de maatregelen die zij voor deze leerlingen namen. We kregen de indruk dat deze maatregelen meer van algemene aard zijn. *‘Voor de goede leerlingen is er verrijkingstof die de leerkracht zo uit de kast trekt’.* Op één school was sprake van een plusklas *‘met allerlei aanbod op het gebied van wereldoriëntatie’* gedurende een beperkt deel van de lestijd. Aanleiding daarvoor was overigens niet zozeer de analyse van de leerprestaties als wel de observatie van de interne begeleider dat een deel van de kinderen van een drukke klas zich verveelde.

Standaarden, normen en handelingsplannen

In het algemeen worden de Cito-normen gehanteerd bij het bepalen of een leerling een zorgleerling is. Scholen hanteren daarbij niet allemaal dezelfde grens. Veel scholen rekenen alleen leerlingen uit de D en E categorie tot zorgleerlingen, bij anderen worden ook de ‘lage C’s’ daartoe gerekend en bij een enkele school alle leerlingen die een lagere waardering hebben dan B. Aan de bovenkant rekenen de meeste scholen alleen de leerlingen met een ‘A+’ mee.

In het algemeen worden voor de zorgleerlingen plannen opgesteld. Het kan daarbij gaan om handelingsplannen (voor individuele leerlingen) en om groepsplannen (voor leerlingen met vergelijkbare problemen)⁹. Soms vermelden scholen dat deze handelingsplannen en groepsplannen ook in het groepsoverzicht worden opgenomen zodat

⁹ Een van de scholen meldt dat externe instanties groepsplannen niet accepteren, maar individuele handelingsplannen eisen.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

de leerkracht het algemene aanbod en de verschillende extra activiteiten gemakkelijker in elkaar kan schuiven.

Bij het opstellen van handelingsplannen, lijkt in principe de groepsnorm uitgangspunt te blijven. Maar als dat niet meer haalbaar is, wordt overgestapt op het volgen van de individuele ontwikkeling. *‘Handelingsplannen kunnen twee kanten op’* stelt een van de geïnterviewde directeuren. *‘Enerzijds zijn zij bedoeld om achterstand in te halen. In dat geval blijft de groepsnorm richtinggevend. Anderzijds zijn zij ook bedoeld om daar wat dat niet (meer) kan, ook de normen aan te passen’*. Een andere school meldt bij deelproblemen te werken met handelingsplannen (waar de vergelijking met de groepsnorm blijft) en bij meeromvattende problemen over te stappen op leerlijnen. Men is zich daarbij wel bewust van de gevaren van de aanpassing van de norm. *‘Je moet met zorgleerlingen wel uitkijken met het stellen van doelen, dat ze niet te stigmatiserend werken of niet te laag uitpakken. Van de andere kant moeten doelen wel haalbaar blijven’*. De plannen worden bij de volgende bespreking geëvalueerd. Sommige scholen melden expliciet dat de hele cyclus wordt gevolgd van opstellen van een plan, evalueren en bijstellen.

4.4.3 Uitvoeringsproblemen, tevredenheid en nadelen

De meeste scholen werken met vaste procedures ten aanzien van het registreren en analyseren van de leerprestaties. Wel zit er ontwikkeling in de procedures. *‘De individuele ontwikkelingscurves volgen we al langer, vanaf de invoering van het LVS (zo’n 10 jaar geleden). Maar 10 jaar geleden hield het allemaal wel minder in dan de laatste jaren’*. Zoals vermeld komen de groepsanalyses minder vaak voor en zijn meestal wat recenter, maar ook daarvoor zijn de procedures meestal vastgelegd. De bewaking van de procedures valt onder de verantwoordelijkheid van de interne begeleider of van de interne begeleider en directeur samen.

Uitvoeringsproblemen

We vroegen de scholen of er sprake was van uitvoeringsproblemen. In het algemeen zijn er wat de technische aspecten betreft weinig klachten. Zoals eerder vermeld, nemen op sommige scholen de interne begeleiders of een administratieve kracht het in-

Opvattingen van scholen over meetgestuurd onderwijs

voeren van de gegevens op zich. Maar ook waar de leerkrachten dat zelf doen, lijken er weinig problemen te zijn. *'Er zijn geen uitvoeringsproblemen. Het digitaal invoeren gebeurt al zo lang door de leerkrachten zelf dat ook degenen met weinig IT-affiniteit er geen problemen mee hebben'. 'Er zijn alleen uitvoeringsproblemen op het moment dat er problemen zijn met de IT. Dan moet worden overgestapt op papieren versies. Niet alle leerkrachten zijn daar dan even vlot mee. Is ook dubbel werk'.*

Uiteraard zijn er in het begin wel wat meer problemen. *'Dit gebeurt dus pas een jaar. Er zijn wel vaste procedures ingesteld, die soms stuiten op uitvoeringsproblemen. Zo lukt het moeilijk om de toetskalender te halen. Door ziekte en dergelijke komt het nogal eens voor dat degene die zou moeten toetsen invalt voor een ander. En natuurlijk is er weerstand bij een deel (ongeveer 1/3) van de leerkrachten.*

In het algemeen melden de scholen weinig weerstand bij het in de gaten houden van de ontwikkeling van de leerprestaties. *'De leerkrachten vinden het wel prettig om op deze concrete manier feedback te krijgen'. Anders ligt het wanneer de resultaten alleen besproken worden met de enkele leerkrachten bij wie de resultaten tegenvallen: 'Dat kan wel bedreigend zijn. De laatste keer ging het om twee leerkrachten van wie de ene zich defensief opstelde en de ander zeer geïnteresseerd was en meedacht over waar het aan kon liggen en wat er aan te doen viel'.*

Weerstand wordt (uiteraard) ook opgeroepen wanneer de uitkomsten aanleiding geven tot substantiële veranderingen in het onderwijs. Zo gingen de vrij ingrijpende veranderingen op een van de (taalzwakke) scholen niet zonder slag of stoot. *'Uitvoeringsproblemen waren er aanvankelijk ten aanzien van de consequenties (van de groepsanalyses) voor het onderwijs. De leerkrachten moesten zich bewust worden van het feit dat de manier waarop ze onderwijs geven er toe doet. Ze moeten zich openstellen voor begeleiding, erkennen dat ze zelf moeten veranderen. Aanvankelijk werd de rol van de interne begeleider ook niet van harte aanvaard en werden bijvoorbeeld afspraken nog wel eens afgezegd. Nu werken de leerkrachten en de interne begeleiders echt samen.*

Ook ten aanzien van de consequenties van individuele analyses meldt deze school weerstand: *'De uitvoeringsproblemen hadden ook hier te maken met de attitude van de leerkrachten. Zij moeten inzien dat zorgleerlingen niet de verantwoordelijkheid*

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

zijn van een remedial teacher maar van de groepsleerkracht zelf. De cultuuromslag die leerkrachten hebben moeten maken is nu naar schatting voor 80% geslaagd, maar gaat nog door'.

Een deel van de scholen meldt uitvoeringsproblemen in de zin van gebrek aan mankracht bij het uitvoeren van de handelings- en groepsplannen. Een paar citaten:

- *'Uitvoeringsproblemen doen zich alleen voor in zoverre er natuurlijk altijd te weinig mankracht is om de handelingsplannen die uit de besprekingen voortvloeien volledig uit te voeren'.*
- *'De uitvoeringsproblemen bevinden zich voornamelijk op het vlak van de consequenties. Die geven vaak organisatorische problemen. We kijken dan waar we extra formatie kunnen inzetten'.*

Tevredenheid

De scholen zijn redelijk tevreden over de wijze waarop zij de gegevens analyseren. Dat geldt zowel voor de scholen waar de nadruk ligt op de individuele analyses als voor de scholen waar de groepsanalyses even belangrijk zijn. Ook scholen waar de analyses niet zoveel om het lijf hebben, zijn niet ontevreden. Er is misschien wel meer mogelijk, maar dat kost ook weer tijd. Het is ook een kwestie van prioriteit, een school die goede resultaten boekt, heeft minder behoefte aan uitgebreide analyses. *'We beschikken over heel veel gegevens en doen er niet zoveel mee. Dat roept wel eens vragen op. Maar we kiezen er toch voor om wel de vinger aan de pols te houden, maar niet door te analyseren zolang de resultaten zo goed zijn'.* *'Trendanalyses zouden kunnen, maar de directe aanleiding ontbreekt en je kunt je tijd maar een keer uitgeven'.*

Op één school is de geïnterviewde interne begeleider niet tevreden over het feit dat er te weinig aandacht is voor de gemiddelde leerling *'Zoals wij dat op dit moment doen (met veel nadruk op de onderkant en de bovenkant), verdwijnt de aandacht voor de middenmoot een beetje. En dat is niet terecht. Als je ook een aanmeldingsformulier zou maken voor de middenmoot, kun je zien dat ook daar af en toe extra aandacht nodig is'.*

Opvattingen van scholen over meetgestuurd onderwijs

Op een andere school meldt de geïnterviewde adjunct-directeur niet tevreden te zijn met de wijze waarop geanalyseerd wordt, omdat er te weinig wordt gekeken naar *'wat de kinderen echt nodig hebben'*. Deze school is echter in een overgangsfase en krijgt in de nabije toekomst de beschikking over een bovenschoolse interne begeleider.

Ook over de consequenties die aan de analyses worden verbonden is men in het algemeen tevreden. De antwoorden op de vraag naar de tevredenheid over de opbrengsten van meetgestuurd onderwijs laten zien dat de analyses vaak worden gebruikt om veranderingen in het onderwijs te onderbouwen.

- *Ja, je hebt zo de resultaten voldoende in beeld. Voldoende om je onderwijs mee te kunnen evalueren en bij te stellen.*
- *Ja, nu wel, omdat er consequenties aan de uitkomsten worden verbonden. De leerkrachten zijn beter gericht op wat ze doen en wat dat uitricht. Het komt in feite neer op planmatig werken waarbij je de controlerende (evaluerende) stap handen en voeten geeft. Die check kun je niet missen.*
- *Ja, er is meer aandacht voor differentiatie en ook meer aandacht voor onderlinge afstemming van de groepen en de leerjaren. Leerkrachten beginnen daar nu toch wel de voordelen van te zien.*
- *De globale analyses van de groepsgegevens hebben geleid tot het bewustzijn dat er iets moest veranderen. Deze veranderingen hebben duidelijk resultaat gehad in de zin van meer aandacht voor en bedrevenheid in klassenmanagement. Met daardoor ook meer aandacht binnen de klas voor zorgleerlingen.*
- *Het levert zeker op met betrekking tot de leerkrachten. Leerkrachten staan open voor tips van elkaar, school is een lerende organisatie, men doet veel aan collegiale consultatie.*
- *Ja, men krijgt een steeds duidelijker beeld van de leerlingen al zou men de aanpak van hen nog wel willen verbeteren.*
- *Het verschil met vóór 2001 is wel dat we toch beter zicht hebben, ook op de zwakke plekken en daar vervolgens wel wat aan doen. Vroeger was toch meer de tendens: 'als het maar gezellig is'.*

Dat wil niet zeggen dat het niet (nog) beter kan. Vaak is er sprake van een proces, maar heeft men het gevoel op de goede weg te zijn.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Zien de scholen ook nadelen?

Als belangrijkste nadeel wordt toch wel genoemd dat het tijd kost. Zowel de analyses als de besprekingen en het uitvoeren van de consequenties. Het is ook een bezwaar dat leerkrachten in het begin vaak noemen. *‘Er was eerst veel gesputter van collega’s dat al dat toetsen maar afhoudt van waar het omgaat, maar dat wordt duidelijk minder. Er zit beweging in. Mensen beginnen nu zelf ook te vragen om afstemmingsoverleg of extra programma’s.’* De algemene tendens is toch: *‘het kost tijd, maar het levert ook op’*.

Nog een nadeel dat regelmatig wordt genoemd is dat de analyses kunnen leiden tot een vorm van stigmatiseren: *‘Soms kunnen analyse-uitkomsten stigmatiserend werken naar leerlingen, leerkrachten of methodes. Je moet realiseren dat je dit gebruikt voor verbetering van het leerproces’*.

Een ander nadeel is volgens een van de geïnterviewden dat er verwachtingen worden gewekt die de school niet altijd kan waarmaken. *‘Als je signaleert dat een leerling of een groep toch iets onder de maat presteert, wil dat nog niet zeggen dat je dat ook kunt veranderen.’*

Dat de toetsen het onderwijs teveel gaan bepalen, zien scholen in het algemeen niet als reëel gevaar, in ieder geval niet op de manier waarop het op het ogenblik gebeurt. *‘Geen gevaar voor teaching to the test, we schieten heus niet door’*.

Wel signaleerde een (vrije) school dat leerkrachten soms wat sneller nieuwe leerstof aanpakken dan de bedoeling is. *‘Leerkrachten, die de klassen meer jaren houden, zijn soms geneigd eerder met lesstof te beginnen dan vanuit de school bedoeld wordt (bijvoorbeeld breuken een jaar eerder behandelen om zo hogere testresultaten te halen).’*

Opvattingen van scholen over meetgestuurd onderwijs

Tabel 4.11 Overzicht nadelen (het totaal aantal bedraagt 17)

	Aantal
a. Kost tijd, is veel werk, te weinig mankracht voor consequenties (maar...)	8
b. Kan stigmatiseren/ te weinig rekening met achtergrond leerlingen	3
c. Leerkrachten zijn soms geneigd te oefenen/ eerder met bepaalde lesstof te beginnen	2
d. Het is lastig leerkrachten aan te spreken als ze slecht scoren	1
e. Er is op deze manier te weinig aandacht voor de middenmoot	1
f. Je schept verwachtingen die je niet na kan komen	1
g. Geen nadelen	1

4.4.4 Voorwaarden voor verdere verbetering

We legden de scholen ook de vraag voor of de wijze van analyseren van de leerprestaties (zowel op groepsniveau als op individueel niveau) nog beter kan en wat daar dan voor nodig is. Daarvoor werden de volgende suggesties gedaan:

- Meer bekendheid over wat verdere analyses zouden moeten inhouden, inclusief het maken van prognoses.
- Scholing van de leerkrachten in het analyseren van resultaten van de eigen groep, zodat de interne begeleider zich kan richten op gecompliceerdere en meer overstijgende analyses.
- Digitaliseren van de methodegebonden toetsen en opname in één overkoepelend systeem.
- Duidelijker handvatten voor het interpreteren van de scores voor gedrag en houding.

Veel voorwaarden voor verbetering betreffen niet zozeer de analyses, maar het natriject. Men krijgt, naar eigen zeggen, steeds duidelijker een beeld van de leerlingen, maar dat wil niet zeggen dat daarmee ook de aanpak voldoende verbetert. Het is niet altijd duidelijk welke handvatten achterblijvende leerkrachten nodig hebben om hun aanpak te verbeteren. Leerkrachten moeten ook open staan voor veranderingen. Een van de geïnterviewde interne begeleiders vond dat oudere leerkrachten het daar moei-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

lijker mee hebben. *'Ik ben een voorstander van teams die qua leeftijdsopbouw gemengd zijn. Maar ouderen hebben het er wel moeilijker mee. Zij laten zich moeilijker coachen en zijn minder gewend naar zichzelf te kijken'*. Zij vond het een verademing te werken met jonge leerkrachten die op de opleiding gewend zijn te reflecteren en ook zelf om coaching vragen.

Vooral komt bij deze vraag ook weer het gebrek aan tijd en mankracht naar voren om de consequenties van de individuele analyses ten volle uit te voeren. Scholen proberen de extra mankracht op allerlei manieren te mobiliseren:

- *De school heeft zelfstandig werken ingevoerd om wat extra instructietijd vrij te maken van de leerkrachten.*
- *We hebben de groepen relatief groot gemaakt, zodat er tijd voor extra instructie kan worden vrijgemaakt.*
- *Ieder vrij moment van de leerkrachten wordt benut voor extra instructie. Zo nemen onderbouwleerkrachten (groepjes) leerlingen onder hun hoede tussen 11:30 en 12:00, worden de gymlessen benut voor extra instructie door de leerkracht die dan geen les hoeft te geven etc..*
- *Bij groep 1-2 kunnen de leerkrachten een beroep doen op de tutor (van piramide). Verder is er wat groeiformatie die het mogelijk maakt de groep te splitsen.*
- *Soms wordt ook de extra begeleiding van een rugzakleerling ingezet voor een groepje leerlingen en niet alleen voor de betreffende leerling*

Niettemin lopen scholen tegen de grenzen van hun mogelijkheden aan en blijft er een duidelijk gevoel van gebrek aan handen in de klas.

Op de vraag of men op de school wel eens een sterkte/zwakte analyse heeft uitgevoerd, waarbij ook analyse van leerprestaties is gebruikt, kregen we in het algemeen een ontkennend antwoord. Drie scholen hebben er wel ervaring mee. In alle drie de gevallen leidde dat tot ingrijpende veranderingen.

- *Ja zeker (en je wilt niet weten hoeveel tijd dat kost). Als nieuwe directeur samen met de interne begeleider. Naar aanleiding daarvan is een missie-visie stuk opgesteld. Het onderwijs moet eigentijdser worden met meer plaats voor differentiatie (en dus meer aandacht voor individuele- en groepsprestaties)*

Opvattingen van scholen over meetgestuurd onderwijs

- *Ja, naar aanleiding van het ongunstige inspectierapport is er een SWOT-analyse geweest, uitgevoerd door het CPS waarin uiteraard de leerlingprestaties een belangrijke rol hebben gespeeld. Dat heeft een heel pakket aan veranderingen opgeleverd*
- *Ja en toen heeft men meer aandacht gekregen voor een beter klassenmanagement. De school is nu een BAS-school (Bouwen aan een Adaptieve School).*

Landelijke referentieniveaus?

We vroegen de scholen ook hoe ze dachten over de invoering van landelijke referentieniveaus. De meningen hierover waren verdeeld. Enkele scholen hebben daar geen bedenkingen tegen: *'Prima om te vergelijken, je moet toch ergens vanuit gaan'*. Een aantal scholen is van mening dat we daarvoor de Cito-normen al hebben. *'De Cito-normen zijn prima. Je kunt de school tenminste ook vergelijken met scholen met een vergelijkbare populatie'*. Men vraagt zich ook af wat er dan nog meer zou moeten: *'Het LVS levert dat eigenlijk al. Als het meer moet zijn dan dat, dan zou ik graag weten wat dat dan in zou moeten houden'*.

Een vrij groot deel van de scholen heeft bedenkingen en voorwaarden die te maken hebben met de onvergelijkbaarheid van scholen. Enkele citaten:

- *'Niet zozeer van landelijke referentieniveaus, wel van regionale. Wij kunnen nooit het niveau halen van een Gooise school. Wij willen ons wel graag vergelijken met scholen met een vergelijkbare leerlingpopulatie.'*
- *'Wel positief, maar dan wel met inachtneming van de (regionale) context. Dat gebeurt nu onvoldoende, daar moet je maar eens een gesprek met de inspectie over hebben.'*
- *'Daar heeft de school wel moeite mee. De school heeft veel zwakke leerlingen en daar houdt de inspectie geen rekening mee. Inspectie stelt normen niet bij als veel leerlingen een laag IQ hebben.'*

In het algemeen is men van mening dat er voorzichtig met landelijke normen moet worden omgegaan en dat daar niet teveel op moet worden gefocust. Nog twee citaten:

- *Lastig in te schatten wat je daarvan mag verwachten. Ons doel is op basis van analyse er uit halen wat er inzit, maximaal rendement. Het is wel prettig daarbij*

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

een meetlat te hebben, maar het moet ook niet meer zijn dan dat. En er moet rekening worden gehouden met de context.

- *Je moet er heel voorzichtig mee zijn. Normen moet je gebruiken als signaal, niet als een heilig doel. Gras gaat niet groeien als je gaat trekken. Individuele leerlingen moeten extra hulp krijgen, maar moeten wel in hun waarde gelaten, niet gezien worden als spelbedervers bij het halen van een norm.*

Tabel 4.12 Standpunt over landelijke referentieniveaus (het totaal aantal bedraagt 17)

	<u>Aantal</u>
a. Ja, daar ben ik vóór	3
b. Die hebben we toch al (cito)?	5
c. Ja, maar onder voorwaarden (regionaal/ rekening houdend met de context etc.)	6
d. Nee (want eerlijke vergelijkingen zijn niet mogelijk)	3

4.5 De scholen voor speciaal basisonderwijs (sbo)

De drie sbo-scholen uit de telefonische ronde lijken niet goed raad te weten met de toenemende vraag van de inspectie om meer groepsanalyses. Zij zijn van mening dat er zoveel verschillen zijn tussen de leerlingen dat dergelijke analyses weinig zinvol zijn. Dat wil niet zeggen dat zij geheel afwijzend staan tegenover een ontwikkeling richting meer meetgestuurd onderwijs. Wij vatten hieronder hun commentaar en suggesties samen.

4.5.1 Bezwaren tegen groepsanalyses

De geïnterviewde sbo-scholen vinden groepsanalyses niet relevant. Op één van deze scholen vinden ze niettemin plaats, op aandringen van de inspectie. De school heeft een ‘denktank’ geïnstalleerd (interne begeleider, bouwcoördinatoren, directie), die onderzoekt hoe met groepsanalyses gewerkt zou kunnen worden. Op dit moment is men uiterst kritisch, omdat de analyses wel zinvol moeten zijn en men die zin nu nog niet ziet.

Ook de beide andere scholen zijn van mening dat het in het sbo lastig is gegevens met elkaar te vergelijken. Daarvoor zijn de individuele verschillen te groot en bovendien stromen leerlingen op verschillende tijdstippen het sbo in. *‘Wat zegt het bijvoorbeeld als je groep 4 van dit jaar met die van vorig jaar vergelijkt? Wat zie je dan? Per groep zijn er zoveel verschillen (bijvoorbeeld in kindkenmerken, IQ, leeftijd instroom in sbo) dat het lastig is hiermee analyses te maken’.*

4.5.2 Ontwikkelingen bij de individuele analyses

In het sbo worden de ontwikkelingscurves van alle leerlingen geanalyseerd. Op alle drie de betrokken sbo-scholen wordt gewerkt met uitstroomprofielen¹⁰ voor het bepalen en evalueren van het onderwijsaanbod. Dit gebeurt aan de hand van gegevens als IQ-tests, leerlingkenmerken (werkhouding etc.), diagnostisch onderzoek, thuissituatie, onderwijsleeftijd (DLE), Cito-uitkomsten (rekenen, spelling) en soms AVI-uitkomsten.

Dit lijkt een verandering te zijn ten opzichte van vroeger. *‘Men is pas een jaar met deze werkwijze bezig, maar het geeft wel meer inzicht in zowel cognitieve als sociaal emotionele ontwikkeling van de leerlingen’*, stelt een van de geïnterviewden. Men is echter van mening dat wetenschappelijke criteria ontbreken en dat het vooral voor zeer jonge kinderen lastig is een (uitstroom)perspectief te maken. Verder is men ook beducht voor stigmatisering: *‘Als er een praktijkonderwijsperspectief wordt uitgebracht, kan dit stigmatisering van de leerling betekenen doordat de leerkrachten eerder tevreden zijn met de resultaten van betrokken leerling’.*

Een van deze scholen lijkt een stap verder te gaan. Op deze school nam de inspectie geen genoegen meer met de wijze waarop de leerprestaties werden bijgehouden. *‘De inspectie vond dat we de handelingsplannen teveel achteraf opstelden. We zijn nu proactiever geworden’.* *We koppelen de gegevens nu aan een ontwikkelingsperspectief en aan voorspellingen over het niveau dat de leerling gaat halen (DLE’s, eisen voortgezet onderwijs).* Sinds de nieuwe aanpak worden de toetsresultaten met de leerkracht besproken. De procedure is als volgt.

¹⁰ Uitstroom: leerwegondersteunend, praktijkonderwijs of anders.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

- Aan de hand van de uitkomsten worden handelingsplannen opgesteld.
- Handelingsplannen én ontwikkelingsperspectieven, worden verwerkt in de plannen van aanpak.
- De plannen van aanpak zijn individueler dan de handelingsplannen. Er wordt preciezer gekeken naar de hiaten en naar middelen om deze op te vullen.
- Er zijn geen groepsanalyses. Wel wordt bij het opstellen van de handelingsplannen nagegaan hoe de plannen passen binnen de instructiegroepen.
- Er wordt net zo lang getoetst tot een kind op C-niveau zit.

Deze manier van werken is nieuw. De school heeft lang gestoeid met de plannen van aanpak, wat vooral in het begin veel tijd kostte en ook veel weerstand. Er begint nu meer lijn in te komen.

De geïnterviewde interne begeleider is wel tevreden over deze werkwijze. *'Je krijgt meer zicht op hoe een kind zich ontwikkelt. Je ziet bijvoorbeeld ook beter wanneer een kind zich sneller dan de prognose ontwikkelt. Dat kan consequenties hebben voor het programma.'* Zij vindt het echter nog te vroeg om te zien of deze manier van werken ook een positieve invloed heeft op de houding en didactische vaardigheden van de leerkracht. *'Bij sommige zie je wel dat ze de uitkomsten vertalen naar hun wijze van lesgeven, anderen maken die koppeling veel moeilijker of niet'* Op de betreffende school wordt ook nog niet systematisch aandacht geschonken aan de consequenties voor de leerkrachten. Dat hangt ook samen met de komst van een nieuwe directie waardoor het proces wordt vertraagd.

4.5.3 Voorwaarden voor meetgestuurd onderwijs in het SBAO

De sbo-scholen hebben behoefte aan een LVS dat meer is toegesneden op het sbo en waarmee leerlingen op maat kunnen worden getoetst. Ook voor het regulier basisonderwijs zouden zulke toetsen er moeten zijn voor leerlingen met ontwikkelingsachterstand. Een van de scholen heeft meegedaan met een pilot van het Cito voor aanpassing van de toetsen, maar op dit moment zijn er nog geen goede normen voor sbo's. *'Er zouden toetsen moeten komen, die met alle aspecten van de leerlingen rekening houden. Wat doe je met een Roma-kind dat een groot deel van de schooltijd verzuimt, omdat het rondtrekt. Hoe test je sociaal emotionele ontwikkeling?'*

De scholen staan niet afwijzend ten opzichte van een landelijk referentiekader, maar ook dan wel in aangepaste vorm. *‘Als er een landelijke meetlat voor sbo-leerlingen zou zijn, zou dat wel een spiegel voor de scholen kunnen vormen’*. De scholen zijn beducht voor afrekening op hun eindresultaten. *‘In het regulier worden zorgleerlingen vaak buiten de eindresultaten van een school gehouden. Dat kan in het SBAO niet, alle leerlingen zijn zorgleerlingen. Ook is het erg lastig om van groepsgemiddelden van SBAO-leerlingen uit te rekenen. Eén leerling kan de uitkomsten significant beïnvloeden’*.

De school die de meeste stappen richting meetgestuurd onderwijs lijkt te hebben gezet, stelt dat ook algemene condities als een goed ingevoerde directie van belang zijn: *‘het is nu met de nieuwe directie een beetje pas op de plaats voor deze ontwikkeling’*. Een bestuur dat de ontwikkelingen stimuleert en bijvoorbeeld stichtingsbreed zou invoeren, zou ook een steun in de rug zijn.

En net als ook in het regulier onderwijs het geval was, lijkt een jong team gemakkelijker voor veranderingen te porren dan een wat ouder team: *‘De ene leerkracht heeft meer problemen dan de ander met de nieuwe manier van werken. Er is hier toch wel sprake van een duidelijke scheiding tussen oud en jong (zeker wanneer er ict bij komt kijken).’*

4.6 Tussentijdse conclusies

In dit deelonderzoek hebben we verkend wat de stand van zaken is in het primair onderwijs ten aanzien van meetgestuurd onderwijs. Daarvoor gebruikten we een quick scan (korte telefonische ronde) bij een representatieve groep schoolleiders en/of interne begeleiders en een kwalitatief vervolgonderzoek bij scholen die in de quick scan hebben aangaven dat zij relatief veel doen aan meetgestuurd onderwijs. Het accent lag op de onderzoeksvragen 3, 4 en 5. We vatten de resultaten samen met als startpunt vraag 4.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Welke handelingen/stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom? (vraag 4)

We hebben de volgende stappen onderscheiden: 1. het vastleggen van doelen en standaarden, 2. het verzamelen van informatie, 3. het registreren ervan, 4. het interpreteren en 5. het nemen van beslissingen.

(1) Vastleggen van doelen en standaarden

In het reguliere basisonderwijs hebben de meeste scholen (circa 80%) vastgelegde leerstandaarden, behalve in de groepen 1 en 2. Dit geldt voor zowel lezen als rekenen. Wel is er enig verschil tussen lezen en rekenen: leerstandaarden voor lezen heeft men in de onderbouw (vooral groep 3 en 4) iets vaker dan voor lezen, in de bovenbouw is dat juist omgekeerd. Op taalzwakke scholen zijn er wat minder vaak leerstandaarden dan op niet taalzwakke scholen, vooral voor rekenen (ca. 66%). Van de sbo-scholen heeft maar ongeveer een kwart vastgestelde leerstandaarden, dit geldt voor zowel lezen als rekenen en voor alle groepen.

(2) Verzamelen van informatie en (3) registreren

Vrijwel alle scholen beschikken over een leerlingvolgsysteem met opslag van gegevens over leervorderingen op verschillende vakken. Nagenoeg altijd voor lezen en rekenen, zeer vaak voor overige onderdelen van het taalonderwijs (zoals spelling, woordenschat) en nog relatief weinig voor andere leerdomeinen. Dit geldt zowel voor het regulier basisonderwijs als voor het speciaal basisonderwijs. We zien hier dus een bevestiging van de uitkomst van de literatuurstudie (zie hoofdstuk 2) dat het gebruik van leerlingvolgsystemen in het primair onderwijs wijd verbreid is.

(4) Interpreteren

Een groot deel van de scholen geeft aan de gegevens die men invoert in het leerlingvolgsysteem ook te analyseren. Dit geldt voor 80 % van de niet-taalzwakke basisscholen, 84% van de taalzwakke basisscholen en 74% van de sbo-scholen.

Op groepsniveau kijkt men vooral naar de prestaties van een groep op een bepaald moment, in vergelijking met een landelijke normgroep (ontleend aan de Cito-normering), en naar de ontwikkeling van de betreffende groep leerlingen in de tijd (dus vergelijkingen met resultaten uit voorgaande jaren). Veel minder vaak vinden vergelijkingen plaats tussen parallelgroepen, tussen groepen leerlingen uit hetzelfde leerjaar (bijvoorbeeld resultaten groep 4 van jaar x met die van groep 4 uit jaar y) en

Opvattingen van scholen over meetgestuurd onderwijs

met andere scholen. Op een klein deel van de scholen gaat analyseren niet verder dan 'globaal in de gaten houden'. Leerkrachten worden niet altijd bij de groepsanalyses betrokken. Dan worden resultaten alleen met hen besproken als er 'onder de maat' wordt gepresteerd.

Op individueel niveau kijkt men eveneens naar prestaties op een bepaald moment, in vergelijking met een landelijke normgroep, en naar de prestatieontwikkeling van het kind in de tijd. In de regel gebeurt dit alleen voor zorgleerlingen.

(5) Beslissingen nemen

De analyse van de gegevens leidt tot beslissingen van verschillende aard, afhankelijk van het type analyse. Ligt het accent op analyse en interpretatie van leerresultaten van individuele (zorg)leerlingen, dan gaan de beslissingen over bijstelling van handelingsplannen en aanpassing van instructie. Ligt het accent op analyse van groepsresultaten, dan gaan de beslissingen over het nemen van maatregelen als meer aandacht voor specifieke leerstofonderdelen in de klas, meer aandacht voor klasmanagement, eventueel bijscholing van leerkrachten en aanschaffen van aanvullend materiaal bij zwakke punten in de gebruikte methodes of keuze van nieuwe methoden. Een discrepantie tussen een landelijke norm en eigen resultaten is niet altijd aanleiding tot actie. Soms vindt men de norm voor de eigen school te hoog gegrepen of heeft men het vertrouwen/de ervaring dat een geconstateerde achterstand nog wel bijtrekt. Slechts een enkele (taalzwakke) school maakt melding van acties ter verbetering van de kwaliteit van de leerkrachten (coaching, scholing).

In hoeverre en hoe worden gegevens over taal- en rekenvaardigheden van leerlingen gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen? (vraag 3)

Het meest kijkt men naar ontwikkelingscurves van zorgleerlingen (75% van de basisscholen doet dit en 64% van de sbo-scholen). Daarbij bekijkt men ook of individuele (zorg)leerlingen gestelde doelen halen. Hoewel niet specifiek gevraagd, is het aannemelijk dat dit gebeurt aan de hand van handelingsplannen voor zorgleerlingen. Analyses op groepsniveau komen wat minder voor in het basisonderwijs, wel meer op taalzwakke scholen (69% van de niet-taalzwakke basisscholen doet dit en 74% van de

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

taalzwakke scholen). In het sbo gebeurt dit echter nauwelijks (20% van de scholen). Vooral analyses die vergelijking met landelijke gemiddelden impliceren zijn in het sbo nagenoeg afwezig. In het sbo analyseert men wel meer dan het reguliere basisonderwijs ontwikkelingscurves van alle leerlingen (64% vs 45%). Maar dat is niet verwonderlijk: in het sbo zijn alle leerlingen zorgleerlingen.

Analysen op groepsniveau gebeurt vaker op scholen met zwakkere leeropbrengsten. Goed presterende scholen (scholen waar de inspectie de leeropbrengsten als voldoende kwalificeert) vinden het minder nodig.

Voor de beoordeling van de leerresultaten gebruiken de scholen vooral de normen uit het Cito LVS, met name de verdeling over de niveaus A t/m E. De vraag die men zich daarbij stelt is of de school ten opzichte van de landelijke meer/minder leerlingen heeft in een bepaalde niveaugroep. De meeste scholen houden bij de groepsanalyses globaal in de gaten of ze niet onder het gemiddelde scoren van de categorie scholen waartoe ze behoren (scholen met een vergelijkbaar leerlingenpubliek). Minder vaak komt het voor dat de school zelf vooraf specifieke doelen stelt en aan de hand daarvan de leerresultaten evalueert. Scholen die dat doen, gebruiken daarbij vaak ook weer de niveau-indeling A t/m E. Een doelstelling is dan bijvoorbeeld: 80% van onze leerlingen haalt tenminste het C-niveau.

Voor individuele zorgleerlingen stelt men individuele leerstandaarden op in de handelingsplannen. Het kan van school tot school verschillen welke definitie men hierbij hanteert voor zorgleerling. Meestal leerlingen in niveaugroep E en D, maar op sommige scholen ook 'lage C's' of alle C's. En soms ook A+ leerlingen (extra begaafde leerlingen). Bij het opstellen van doelen in handelingsplannen, blijft in eerste instantie de groepsnorm het uitgangspunt. Maar als dat niet meer haalbaar is, wordt overgestapt op het volgen van de individuele ontwikkeling in een eigen leerlijn.

Bestaat er draagvlak bij leerkrachten en schoolleiders voor meetgestuurd onderwijs? (vraag 5)

De respondenten in de quick scan waren het overwegend eens met de volgende stellingen:

- verbetering van het onderwijs in taal en rekenen kan niet zonder regelmatig meten van je onderwijsopbrengsten

Opvattingen van scholen over meetgestuurd onderwijs

- op onze school is er voldoende openheid in het team om met elkaar de leeropbrengsten in ieders groep te bespreken
- op onze school is er voldoende deskundigheid om analyses te plegen op de gegevens in het leerlingvolgsysteem
- jezelf vergelijken met andere scholen met hetzelfde type leerlingenpubliek is een goede manier om te evalueren of de leerlingen presteren wat verwacht mag worden

Dat wijst op een behoorlijk draagvlak voor de in de stellingen vervatte aspecten van meetgestuurd onderwijs. Alleen over de vierde stelling waren de meningen meer verdeeld; met name in het sbo wordt dit minder onderschreven.

Uit de vervolginterviews is gebleken dat scholen die al wat langer bezig zijn met meetgestuurd onderwijs daarvoor inmiddels goed ingeburgerde procedures hebben waar leerkrachten zich in kunnen vinden. Beginnende scholen melden wel dat ze nog bij een deel van de leerkrachten wat weerstand moeten overwinnen. Vooral als de resultaten tegenvallen kan dat bedreigend zijn, een enkele maal wordt melding gemaakt van leerkrachten die zich defensief opstellen in het gesprek hierover. Ook wanneer de uitkomsten aanleiding geven tot substantiële veranderingen in het onderwijs kan dit in eerste instantie weerstand oproepen.

Niet elke school wil evenveel diepgang op het gebied van meetgestuurd onderwijs. Scholen die goede resultaten boekten, zeggen minder behoefte te hebben aan uitgebreide analyses.

De sbo-scholen hebben behoefte aan een leerlingvolgsysteem dat meer is toegesneden op de eigen populatie. De wensen die men heeft op dit gebied gaan soms tamelijk ver ('voor elke doelgroep toetsen waarmee leerlingen op maat kunnen worden getoetst').

In het kader van de verkenning van draagvlak is in de vervolginterviews ook gevraagd naar de mening over de invoering van referentieniveaus. Daarover wordt wisselend gedacht. Een minderheid vindt dat een goede ontwikkeling. Eveneens een minderheid is er beslist tegen. De meeste schoolleiders/interne begeleiders vinden het of overbodig ('we hebben de normering van de Cito LVS toetsen toch al?') of wel

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

goed maar onder de voorwaarde dat er bij het opstellen van referentieniveaus rekening wordt gehouden met verschillen tussen scholen in leerlingenpubliek.

5 Leraren over referentieniveaus

5.1 Achtergrond en onderzoeksvragen

Hoe denken leraren over het werken met referentieniveaus? Dat is de centrale vraag van dit hoofdstuk. We hebben deze vraag beantwoord via een vragenlijst-onderzoek onder een specifieke groep leraren, namelijk leraren die deelgenomen hebben aan het Cito-pilotproject 'Referentieniveaus medio jaargroep 5'. Vanzelfsprekend vormen deze leraren geen representatieve selectie van alle leraren uit het land. Maar in dit geval vormt dat geen nadeel. Het werken met referentieniveaus is voor veel leraren een nog onbekende werkwijze. Bijgevolg hebben ze zich vaak nog niet een mening kunnen vormen. Onze interesse betreft met name de mening van leraren die met het onderwerp al enige bekendheid hebben. De deelnemers aan het Cito-pilotproject komen aan deze omschrijving tegemoet.

In dit hoofdstuk staat de term 'referentieniveau' centraal. *Een referentieniveau is een aanduiding van het vorderingenniveau dat men op een nader omschreven moment bij leerlingen wil bereiken.* Men kan één referentieniveau hanteren, dat dan noodzakelijkerwijs samenvalt met de grens tussen voldoende en onvoldoende. Men kan – rekening houdend met verschillen tussen leerlingen – voor het betreffende moment ook méér referentieniveaus vaststellen, bijvoorbeeld een referentieniveau voor snelle leerlingen en een referentieniveau voor zwakke leerlingen. In het Cito-project zijn steeds twee referentieniveaus bepaald, en wel voor drie domeinen (technisch lezen, begrijpend lezen en rekenen/wiskunde). Omdat het een pilotproject betreft, is dit alleen nog gedaan voor medio jaargroep 5. De gekozen referentieniveaus zijn het minimumniveau en het voldoende-niveau.

Het *minimumniveau* is het laagste referentieniveau. Het geeft het niveau aan dat voor vrijwel alle leerlingen – zeg 90 tot 95 procent – haalbaar is.

Het *voldoendeniveau* ligt hoger. Het geeft het niveau aan dat voor de meeste leerlingen – zeg 70 tot 75 procent – haalbaar is.

Deelnemers aan het project hebben de niveaus bepaald op basis van toetsitems die zijn ontleend aan toetsen van het Cito-leerlingvolgsysteem. Voor afzonderlijke

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

items hebben leraren aangegeven of deze al dan niet beheerst zouden moeten worden om van een minimum of van een voldoende niveau te kunnen spreken. Om inconsistenties in oordelen te voorkomen, zijn de items gerangschikt naar moeilijkheidsgraad. Het beoordelingsproces is in drie fasen uitgevoerd, analoog aan de werkwijze die in het project Periodieke Peiling van het Onderwijsniveau is ontwikkeld. In de eerste fase oordeelden leraren individueel, in de tweede fase overlegden de leraren in groepjes over hun oordelen, in de derde fase zijn ze geïnformeerd over de vaardigheidsverdeling in de leerlingenpopulatie. Deze informatie stelde de leraren in staat zich een oordeel te vormen over de referentieniveaus in relatie tot de niveaus die door leerlingen daadwerkelijk worden gerealiseerd. Desgewenst hebben ze op grond hiervan hun oordelen nog kunnen bijstellen.

Wij zullen hier niet verder ingaan op de werkwijze van het Cito-project en ook niet op de resultaten. Zie hiervoor de rapportage door het Cito zelf¹¹. We beperken ons hier tot de vragenlijst die we hebben voorgelegd aan de deelnemende leraren enkele maanden na afloop van het project. Via de vragenlijst wilden we een antwoord krijgen op de volgende vragen.

- a. Maken de leraren op de eigen school gebruik van referentieniveaus; zo ja, welke?
- b. Hoe staan de leraren tegenover het gebruik van landelijke referentieniveaus?
- c. Welke voor- en nadelen zien de leraren ten aanzien van het gebruik van landelijke referentieniveaus?
- d. Verwachten de leraren dat het gebruik van referentieniveaus leidt tot verbeterde leerresultaten?
- e. Menen de leraren dat de in het Cito-project ontwikkelde referentieniveaus op landelijke schaal ingevoerd zouden moeten worden?

5.2 Opzet van het deelonderzoek

De onderzoekspopulatie is omschreven als alle deelnemers aan het Cito-project 'Referentieniveaus medio jaargroep 5'. Het betreft 65 leraren met leservaring in

¹¹ Frank van der Schoot (2008). *Referentieniveaus medio jaargroep 5; een pilotstudie in opdracht van het Ministerie van OCenW onder basisscholen in Arnhem en Den Haag*. Arnhem: Cito.

jaargroep 5, werkzaam op een school in Arnhem (n = 34) dan wel in Den Haag (n = 31). Over de leraren is verder weinig bekend. Ze hebben zich, na een lokale werving, vrijwillig opgegeven als projectdeelnemer.

De in de vorige paragraaf aangeduide onderzoeksvragen zijn omgezet in een schriftelijke vragenlijst. De vragenlijst is op 5 juni 2008 via e-mail toegezonden aan alle 56 leraren waarvan ons het e-mailadres bekend was. We hebben twee rappels verstuurd, op 23 juni (vlak voor de zomervakantie) en nogmaals op 26 september (vlak na de zomervakantie). Eindresultaat is als volgt geweest: 25 leraren hebben de vragenlijst geheel of gedeeltelijk ingevuld en teruggestuurd, 11 leraren hebben deelname geweigerd – veelal door gebrek aan tijd – en 20 leraren hebben niets van zich laten horen. Ten opzichte van het totaal aantal aangeschreven leraren bedraagt de respons 45 procent. Het is gezien de eenvoud van de onderzoeksopzet niet mogelijk om na te gaan of de responsgroep representatief is voor de populatie van alle leraren die aan de Cito-pilot hebben meegedaan.

Tegen de verwachting in bleek dat de responsgroep toch drie leraren bevatte zonder leservaring in groep 5. Deze leraren hebben we uit het bestand verwijderd. De resterende 22 leraren zijn gemiddeld inmiddels 21 jaar als groepsleraar werkzaam (standaarddeviatie 10 jaar). Het gemiddelde aantal ervaringsjaren in groep 5 bedraagt 6 (standaarddeviatie 4). Naar het ons derhalve voorkomt, gaat het om een groep leraren die gekenmerkt wordt door een over het algemeen langdurige onderwijservaring.

5.3 Resultaten

We hebben de leraren gevraagd of ze – in de eigen lespraktijk – gebruik maken of hebben gemaakt van een of meer referentieniveaus voor drie met name genoemde domeinen, namelijk technisch lezen, begrijpend lezen en rekenen/wiskunde. Ongeveer de helft van de leraren beantwoordde deze vraag bevestigend, ongeacht het domein (tabel 5.1). Het blijkt dat leraren ofwel helemaal geen referentieniveaus gebruiken, ofwel dat voor alle drie de domeinen inderdaad wel doen. Sommige leraren verschaften een toelichting op de gebruikte toetsen, de hoogte van het niveau en het aantal leerlingen dat het niveau zou moeten halen. Zes leraren verwij-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

zen voor hun referentieniveaus naar toetsen uit het leerlingvolgsysteem van het Cito. Leraren noemen verschillende standaarden. Genoemd zijn onder andere: alle leerlingen minimaal het C-niveau, 95 procent van de leerlingen minimaal het D-niveau, 75 procent van de leerlingen minimaal het B-niveau. Vier andere leraren hanteren referentieniveaus die gerelateerd zijn aan methodegebonden toetsen. Meestal wordt in deze gevallen als standaard aangehouden dat alle leerlingen tenminste 80 procent van de opgaven goed moeten hebben. Ook de Avi-toetsen zijn genoemd: de in dat geval door leraren genoemde standaard is dat medio jaargroep 5 75 procent van de leerlingen op Avi-9 leest.

Tabel 5.1 Het aantal leraren dat één of meer referentieniveaus gebruikt

<i>Domein</i>	<i>Huidig gebruik van referentieniveau</i>		<i>Totaal</i>
	<i>Ja</i>	<i>Nee</i>	
<i>Technisch lezen</i>	11	11	22
<i>Begrijpend lezen</i>	12	10	22
<i>Rekenen/wiskunde</i>	12	10	22

Wij hebben de leraren gevraagd naar hun houding tegenover de voorgenomen invoering van landelijke referentieniveaus. Die houding is bij vijftien van de 22 leraren positief. Slechts twee staan er negatief tegenover (tabel 5.2). De tegenstanders hebben niet de moeite genomen hun standpunt toe te lichten, veel voorstanders hebben dat wel gedaan. Negen leraren geven aan dat zij de twee door het Cito voorgelegde niveaus – voor voldoende en minimum – een goede keus vinden. Zeven leraren zouden graag meer dan twee niveaus zien, tot vijf aan toe. Genoemd is onder andere een niveau voor het gemiddelde – derhalve door 50 procent van de leerlingen te realiseren – en een referentieniveau 'hoog' voor de beste twintig procent.

Tabel 5.2 Houding tegenover de invoering van landelijke referentieniveaus

Houding t.o. landelijke referentieniveaus	Aantal
Ik sta hier negatief tegenover	2
Ik heb hierover geen mening	0
Ik sta hier aarzelend maar niet onwelwillend tegenover	5
Ik sta hier positief tegenover	15
Totaal	22

Wij hebben de leraren gevraagd welke voor- en nadelen zij zien in verband met de invoering van landelijke referentieniveaus. Ondanks dat de leraren in meerderheid positief zijn over de invoering van landelijke referentieniveaus, brengen ze een veelheid aan nadelen naar voren (zie tabel 5.3). Acht leraren zijn bevreesd dat scholen met veel achterstandsleerlingen benadeeld zullen worden (nadeel a.). Scholen zullen er niet in slagen verschillen die er bij de aanvang van de schoolloopbaan zijn, geheel weg te werken. Drie leraren vrezen dat getallen de overhand zullen krijgen en dat de context waarin scholen en leerlingen functioneren naar de achtergrond verdwijnt (nadeel b.). Een voorbeeld daarvan is dat de gebruikte toetsen onzuiver zijn voor leerlingen die het Nederlands niet als eerste taal hebben (nadeel c.). Leraren zijn verder bang dat leerdoelen die niet gerepresenteerd zijn in de referentieniveaus, minder aandacht zullen krijgen (nadeel d.). En ook vrezen ze dat er problemen kunnen rijzen wanneer scholen een andere indeling van de leerstof hanteren. Het kan gebeuren dat de toets kennis en vaardigheden aan de orde stelt die in de door de school gevolgde methode nog niet aan de orde zijn geweest (nadeel e.). Twee leraren menen dat de representativiteit van de referentieniveaus tekort schiet (nadeel f.). En tot slot, één leraar signaleert een spanning met het beleidsprogramma passend onderwijs (nadeel g.).

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Tabel 5.3 Overzicht van de nadelen die leraren zien; aantal leraren (het totaal aantal bedraagt 22)

Nadelen	Aantal
a. Landelijke referentieniveaus doen onvoldoende recht aan verschillen tussen scholen wat betreft instroom ('scholen hebben verschillende input, dus verschillende niveaus'; 'scholen in achterstandswijken zullen een lager resultaat halen'; 'ik heb te maken met zij-instromers van allochtone afkomst').	8
b. Getallen krijgen de overhand ('het beeld van een school wordt snel bepaald door getallen'; 'geen kind is hetzelfde'; 'de score op toetsen gaat het oordeel van de leraar overvleugelen').	3
c. De gebruikte toetsen zijn niet eerlijk voor allochtone leerlingen ('leerlingen verschillen in taalniveau en begrijpen soms vragen niet, terwijl ze de kennis wel hebben'; 'voor allochtonen is het vaak te moeilijk en te hoog gegrepen'; 'achterstand in taal leidt ook tot slechtere prestaties bij rekentoetsen').	3
d. Niet getoetste doelen worden minder belangrijk ('belangrijke andere ontwikkelingsdoelen zijn niet in toetsen te vangen, zoals zelfstandigheid, zelfvertrouwen'; 'wat zegt een resultaat op een toets als een kind andere vaardigheden heeft?').	2
e. Er ontstaan problemen als scholen een methode gebruiken die een afwijkende leerstofindeling hanteert ('als de methode niet alle getoetste onderwerpen behandelt'; 'sluit niet aan bij de methode').	2
f. De landelijke representativiteit schiet tekort ('het onderzoek is niet breed genoeg geweest'; 'de vaststelling moet zorgvuldig gebeuren, geen nattevingerwerk zoals nu').	2
g. Er is spanning met het streven passend onderwijs te bieden ('de kans bestaat dat je de leerstof weer centraal gaat stellen en bij passend onderwijs stel je juist de leerlingen centraal').	1

Anderzijds merken we op dat de leraren zeker ook voordelen aangeven (tabel 5.4). Tien leraren waarderen de duidelijkheid die landelijke referentieniveaus opleveren positief (voordeel a.). Andere leraren zien de objectiviteit van de normen als een voordeel (voordeel b.), of wijzen op de vergelijkingsmogelijkheden met

andere scholen (voordeel c.). Eén leraar denkt dat het gebruik van referentieniveaus helpt bij het opsporen van hiaten in de leerstofaanbieding (voordeel d.). Opvallend is dat sommige leraren voordelen noemen die eigenlijk onterecht zijn. Ze verwarren de doelen van de landelijke referentieniveaus met die van een leerlingvolgsysteem. De invoering van landelijke referentieniveaus is natuurlijk niet bedoeld om 'uitvallers te signaleren' of 'groei voor ouders zichtbaar te maken', zoals respondenten menen.

Tabel 5.4 Overzicht van de voordelen die leraren zien; aantal leraren (het totaal aantal bedraagt 22)

Voordelen	Aantal
a. Het geeft duidelijkheid ten aanzien van de landelijke leerdoelen ('iedereen weet waar hij aan toe is'; 'het vergemakkelijkt het overstappen naar een andere school of naar het vervolgonderwijs'; 'uitvallers zie je niet over het hoofd'; 'grafiek maakt groei zichtbaar voor ouders').	10
b. Het is goed om objectieve normen te hebben ('onafhankelijke vaststelling van leerniveau van kinderen').	3
c. Het biedt de school de mogelijkheid tot zelfreflectie of -evaluatie ('kan de resultaten vergelijken met andere scholen').	3
d. Het helpt bij het opsporen van hiaten in de aangeboden leerstof.	1

Verwachten leraren dat de invoering van een of meer referentieniveaus bijdraagt aan verbeterde resultaten? We legden de respondenten deze vraag voor en we vroegen ook naar hun argumenten. Het blijkt dat leraren hierover verschillend denken. Ongeveer de helft is positief of voorwaardelijk positief. Eén op de drie leraren daarentegen weet het niet of denkt dat het niet het geval zal zijn (tabel 5.5). De aangedragen argumenten laten overigens ook hier weer zien dat niet alle leraren de oogmerken van de referentieniveaus goed voor ogen hebben.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Tabel 5.5 Dragen landelijke referentieniveaus bij aan verbeterde leerresultaten?

Stellingname (en argumenten)	Aantal
Ja ('want het stimuleert het geven van onderwijs op maat'; 'de eventuele achterstand kan beter begeleid worden'; 'uitvallers worden tijdig opgespoord')	5
Ja, mits ... ('het is afhankelijk van de kwaliteiten van de leerkracht'; 'alleen het weergeven van de resultaten is niet genoeg, handelingsadviezen zijn óók nodig')	6
Weet het niet ('dat vraag ik me af!!??')	2
Nee ('hier geloof ik niet in; de kennis van de kinderen zal hetzelfde blijven'; 'ik denk dat er al optimaal aandacht wordt gegeven aan kinderen die uitvallen of achterblijven')	5
Geen antwoord	4
Totaal	22

De slotvraag was of leraren menen dat de referentieniveaus die in de Cito-pilot voor medio jaargroep 5 zijn vastgesteld daadwerkelijk op landelijke schaal zouden moeten worden ingevoerd. Met andere woorden, is de gevolgde werkwijze goed genoeg en is de representativiteit van de twee groepen leraren voldoende gegarandeerd om tot landelijke invoering over te gaan? Elf leraren menen van wel, zeven leraren menen van niet en vier onthouden zich van een oordeel (tabel 5.6). Hierover lopen de meningen dus uiteen. Zie verder de aangedragen argumenten in de tabel.

Tabel 5.6 Resultaten uit de Cito-pilot landelijk invoeren?

Stellingname (en argumenten)	Aantal
Ja ('we zijn toch op een professionele manier ermee aan de slag geweest'; 'er is drie middagen een aantal ervaren leerkrachten aan het werk geweest met deze niveaubepaling'; 'Cito heeft veel ervaring en weet ook dat de gestelde doelen niet door iedereen gehaald gaan worden'; 'wij werken in het veld en weten waar de problemen liggen, door erover te praten kom je tot betere resultaten')	11
Geen mening ('het is moeilijk om daar een uitgesproken mening over te geven, het bleek dat de leerkrachten meer van de kinderen verwachten'; 'op zogenoemde zwarte scholen zal dit tot frustratie kunnen leiden omdat het referentieniveau hoger ligt')	4
Nee ('ik heb niet de illusie dat wat wij in die middag hebben gedaan, enige invloed heeft'; 'ik vraag me af of er wel een voldoende groot aantal mensen meegewerkt heeft'; 'ik heb het idee dat we tijdens de bijeenkomsten erg uitgegaan zijn van onze normering die al helemaal is ingeslepen op grond van de huidige Cito-normering'; 'de referentieniveaus stonden al vast en er werd alleen gecontroleerd of ze overeen kwamen met onze inschatting'; 'het vaststellen is niet zorgvuldig gebeurd ... ik noem het nattevingerwerk')	7
Totaal	22

5.4 Tussentijdse conclusies

In deze paragraaf vatten we de lessen uit dit hoofdstuk samen. We doen dat aan de hand van de in paragraaf 5.1 onderscheiden onderzoeksvragen.

a. Maken de leraren op de eigen school gebruik van referentieniveaus; zo ja, welke?

Ongeveer de helft van de leraren maakt inderdaad al gebruik van referentieniveaus voor de kernvakken technisch en begrijpend lezen en voor rekenen/wiskunde. Het betreft per vak veelal één niveau, waar 75 tot 100 procent van de leerlingen aan dient te voldoen. De referentieniveaus zijn vaak gedefinieerd in termen van scores op toetsen uit het Cito-leerlingvolgsysteem of op methodegebonden toetsen.

b. Hoe staan de leraren tegenover het gebruik van landelijke referentieniveaus?

De leraren zijn overwegend positief over het gebruik van landelijke referentieniveaus. Iets minder dan de helft van hen kiest voor het hanteren van twee niveaus, een minimum- en een voldoende niveau zoals in de Cito-pilot. Eén op de drie leraren hanteert liever meer dan twee niveaus om beter recht te kunnen doen aan verschillen tussen leerlingen.

c. Welke voor- en nadelen zien de leraren ten aanzien van het gebruik van landelijke referentieniveaus?

Ondanks de positieve houding van leraren tegenover de invoering van landelijke referentieniveaus, zien ze toch ook bezwaren. Een op de drie leraren is bevreesd dat de niveaus nadelig uitpakken voor scholen met veel achterstandsleerlingen. Zulke scholen hebben het moeilijker de afgesproken niveaus te halen. Aan de andere kant noemen leraren ook voordelen. Ongeveer de helft van hen ziet als voordeel dat landelijke referentieniveaus klaarheid scheppen die er nu nog niet is.

d. Verwachten de leraren dat het gebruik van referentieniveaus leidt tot verbeterde leerresultaten?

Ongeveer de helft van de leraren onderschrijft deze verwachting. Leraren beseffen dat alleen het vastleggen van landelijke referentieniveaus niet genoeg is. Het gaat om de follow-up: wat doet de school als resultaten tegenvallen? Een deel van de leraren is sceptisch, omdat men meent dat de marges voor onderwijsverbetering smal zijn. De kwaliteit is al heel hoog.

e. Menen de leraren dat de in het Cito-project ontwikkelde referentieniveaus op landelijke schaal ingevoerd zouden moeten worden?

Ongeveer de helft van de leraren beantwoordt deze vraag met ja. Een ander deel is kritisch over de door het Cito gevolgde procedure voor de bepaling van de niveaus. Men vindt de basis te smal of heeft twijfel over de gevolgde driestapsprocedure.

Tot slot, de in dit hoofdstuk gerapporteerde uitkomsten hebben zoals we in paragraaf 5.1 aangaven, betrekking op een specifieke groep leraren, namelijk leraren die betrokken zijn geweest bij de Cito-pilot om voor medio jaargroep 5 ervaring

Leraren over referentieniveaus

op te doen met het vastleggen van referentieniveaus. Het gaat om leraren die ten aanzien van referentieniveaus waarschijnlijk inhoudelijk redelijk goed op de hoogte zijn en een positieve grondhouding hebben. We moeten constateren dat de 'support' in deze groep niet overtuigend groot is. Weliswaar stellen de leraren zich welwillend op tegenover de invoering van landelijke referentieniveaus voor de kernvakken. Maar tegelijkertijd vrezen ze niet onbelangrijke nadelen en zijn ze sceptisch over het effect ervan op onderwijsverbetering. Ook zijn ze kritisch over de Cito-pilot en de daarin gevolgde werkwijze.

6 Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

In dit hoofdstuk staan de ervaringen centraal van vijf scholen uit de pilots taalbeleid. Deze scholen hebben, onder andere via hun deelname aan de taalpilots, relatief veel ervaring opgedaan met meetgestuurd onderwijs. Hoe zij precies hun metingen uitvoeren en benutten voor onderwijsverbetering verschilt, maar kan voor andere scholen mogelijk aanknopingspunten bieden voor het inrichten van een (meer) opbrengstgerichte werkwijze.

6.1 Achtergrond en onderzoeksvragen

De focus van het beleid voor het primair onderwijs ligt momenteel bij het verbeteren van het onderwijs in de basisvaardigheden. De verwachting is dat die verbetering mogelijk is, als scholen meer aandacht krijgen voor opbrengstgericht werken, als de zorg voor zwakke leerlingen wordt opgevoerd en als het onderwijs beter wordt afgestemd op het omgaan met verschillen. Middelen die scholen daarvoor kunnen inzetten zijn met name het werken met leerstandaarden of referentieniveaus voor taal en rekenen, het systematisch en evaluatief benutten van het leerlingvolgsysteem, en het kiezen voor goede methoden die bij voorkeur ‘evidence based’ zijn.

Duidelijk is dat de meeste scholen inmiddels beschikken over gegevens over het prestatieniveau en de leervorderingen van leerlingen, via hun leerlingvolgsysteem (Blok, Otter & Roeleveld, 2001). Dit type gegevens wordt het meest gebruikt voor rapportages aan ouders, bij overdracht van een groep aan een volgende leerkracht, en voor het opsporen van leerlingen die extra zorg nodig hebben. Het lijkt er echter op dat de beschikbare gegevens nog weinig gebruikt (konden) worden voor de evaluatie van prestaties op groepsniveau of voor vergelijking tussen groepen, met andere scholen of het landelijk gemiddelde als ingang voor een evaluatie van het eigen onderwijs. Toch zou dit type gebruik van metingen een goede

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

mogelijkheid bieden voor gerichte aanpassingen en interventies, en voor het formuleren van doelen ten aanzien van de leeropbrengsten.

Taalpilots

In 2007 zijn in een aantal gemeenten en regio's de Pilots Taalbeleid Onderwijs-achterstanden gestart, in opdracht van het Ministerie van OCW, en uitgevoerd door het Projectbureau Kwaliteit van het PO-Raad. In het voorbereidingsjaar 2006-2007 ging het om zeven pilots, in de G4 (Amsterdam, Rotterdam, Den Haag en Utrecht), Enschede, Oost-Groningen en Zeeland. Doel van de pilots is scholen een extra impuls te geven voor de uitvoering van hun taalbeleid om taalachterstanden te bestrijden. In 2007-2008 zijn er nog veertien nieuwe pilots bijgekomen. In totaal zijn er nu 350 scholen die binnen de taalpilots werken aan het verbeteren van hun taal- en leesresultaten door onder andere het vergroten van leerkrachtcompetenties, het inzetten van een taalcoördinator en gerichte bovenschoolse activiteiten.

Uitgangspunt is dat de invulling van de pilots moet aansluiten op wat scholen al doen. Per pilot zijn dan ook de accenten en de werkwijze enigszins verschillend. In de aanpak zijn echter wel steeds drie principes van effectieve schoolontwikkeling gehanteerd: leren van eigen ervaringen, leren van elkaar en leren van experts. Ook het gebruik van datafeedback en monitoring als onderdeel van het innovatieproces, maakt deel uit van het projectplan van de Taalpilots, met als onderbouwing: "De beschikbare literatuur laat zien dat datafeedback weliswaar een nuttige functie kan vervullen en resultaatgerichtheid bevordert in school, maar ook dat scholen er niet zomaar mee overweg kunnen. Het vaststellen van wat men wil weten (formuleren van een vraag), daarvoor gerichte informatie verzamelen, de verkregen data interpreteren en daaraan conclusies verbinden voor het schoolbeleid/de voortgang van het innovatieproces zijn stappen die veel schoolleiders, laat staan leraren, niet zonder begeleiding kunnen zetten. De ervaringen tot nu toe wijzen er op dat de inzet van onderzoekers nodig is om het proces goed te laten verlopen, niet alleen om de school te voorzien van betrouwbare data, maar ook om ze te helpen die data te begrijpen en in het juiste perspectief te zien. Wel is het zo dat veel scholen al beschikken over tal van bronnen die daarvoor benut kunnen worden. De kunst is echter om die ook daadwerkelijk te gebruiken, en om dat doelgericht te doen." (Zie Projectplan Taalpilots, www.taalpilots.nl). Vanwege de relatieve onervarenheid die men verwachtte bij de scholen in het

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

omgaan met ‘datafeedback’ voorziet het project in monitoring en evaluatie door externe onderzoekers. Eén van de voorwaarden die is gesteld aan scholen die deelnemen was echter dat zij bij aanvang zicht hadden op de taal- en leesprestaties van hun leerlingen, en binnen de Taalpilots is er expliciet aandacht voor het verspreiden van kennis over datafeedback door experts en onderzoekers.

Gezien deze context leek het te verwachten dat de scholen die deelnemen aan de pilots Taalbeleid bekend zijn met het uitvoeren van meetgestuurd onderwijs.

Onderzoeksvragen in de casestudies

In het kader van het onderzoek zijn vijf taalpilotscholen benaderd die ook binnen de groep van taalpilotscholen verhoudingsgewijs ver zijn in het uitvoeren van meetgestuurd onderwijs. Doel van de casestudies die op deze vijf scholen zijn uitgevoerd, is het in beeld brengen van hun werkwijze en ervaringen. Het accent ligt daarbij op beantwoording van de onderzoeksvragen 4, 5, 6 en 7 uit deze studie:

4. Welke handelingen/stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom?
5. Bestaat er draagvlak bij leerkrachten en schoolleiders voor meetgestuurd onderwijs?
6. Wat zijn risico- en succesfactoren bij meetgestuurd onderwijs? Zijn er eventueel onbedoelde effecten?
7. Wat hebben schoolleiders en leerkrachten nodig om meetgestuurd onderwijs te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

Vanzelfsprekend geven we ook een beschrijving van de stand van zaken op deze scholen in het licht van de twee voorafgaande vragen:

2. Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?
3. In hoeverre en hoe worden deze gegevens gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?

6.2 Opzet van het deelonderzoek: casestudies pilots taalbeleid

In deze paragraaf bespreken we de selectie en werving van de casestudiescholen, de aanpak van de dataverzameling via gesprekken met leerkrachten en schoolleiding, en de werkwijze bij de kwalitatieve analyse van de verzamelde gegevens.

Selectie en werving van de casestudiescholen

De selectie van de vijf scholen heeft plaatsgevonden in overleg met de landelijk coördinator van de pilots taalbeleid en enkele van de lokale coördinatoren. Het belangrijkste selectie criterium was de wens om voorbeelden te vinden van scholen die ervaringen hebben met meetgestuurd werken waarvan anderen kunnen leren. Daarbij waren zowel de positieve, als de minder positieve ervaringen van belang: een belangrijk leerpunt voor anderen is immers ook het vermijden van valkuilen. Verder is gestreefd naar spreiding over scholen met vrijwel uitsluitend (allochtone en/of autochtone) gewichtenleerlingen en scholen met een meer gemengd publiek, en naar het uitvoeren van één van de casestudies in een school voor speciaal (basis)onderwijs.

De werkwijze hield in dat de landelijk coördinator van de pilots taalbeleid vijf scholen voorstelde die voldeden aan de opgegeven criteria, en die wij voor de casestudies zouden kunnen benaderen. De scholen kregen een telefonische uitnodiging en via e-mail verdere achtergrondinformatie over het onderzoek. Met de scholen die positief reageerden werd vervolgens een afspraak gemaakt voor een bezoek van één van de onderzoekers aan de school. Wanneer een school afviel - wat overwegend gebeurde vanwege tijdgebrek - werd opnieuw via de landelijk coördinator van de pilots taalbeleid, een school gezocht die aan de criteria voldeed.

Langs deze weg zijn de vijf casestudiescholen gevonden, volgens de gewenste verdeling: twee van de scholen die deelnamen aan het onderzoek zijn te typeren als school met veel allochtone gewichtenleerlingen (in respectievelijk Amsterdam en Terneuzen (school 1 en 2), is er één school met een gemengde populatie van 50% allochtone leerlingen en 50% autochtone leerlingen (Rotterdam, school 3), is er één school met overwegend autochtone leerlingen (Enschede, school 4) en één SBO-school (Enschede, school 5).

Gesprekken met leerkrachten en schoolleiders

Kern van de casestudies was het voeren van gesprekken met leerkrachten. Leerkrachten zijn de sleutelpersonen vanuit het perspectief van de doelstellingen van de beleidsagenda: zij moeten immers zorgen voor de gewenste prestatieverbetering. Op alle scholen was de dataverzameling opgebouwd uit de volgende elementen:

- een verkenning van documenten over dataverzameling en -gebruik op de school: welke gegevens zijn er, hoe zijn ze verzameld en weergegeven, welke analyses voert de school uit, hoe komen interpretaties tot stand, en wat zijn daarbij leidende vragen?
- twee groepsgesprekken, met in totaal circa vier leerkrachten verdeeld over kleuterbouw/onderbouw en middenbouw/bovenbouw
- een groepsgesprek met de schoolleider, interne begeleider(s) en/of de taalcoördinator

Voor de gesprekken zijn twee gespreksleidraden ontwikkeld: één voor de gesprekken met leerkrachten en één voor de gesprekken met de schoolleiding, interne en/of taalcoördinator. Gespreksonderwerpen waren: de toepassing van de verschillende stappen bij meetgestuurd onderwijs, het draagvlak, de opbrengsten, de succes- en risicofactoren en de vereiste kwaliteiten en houdingen bij leerkrachten, schoolleiders en intern begeleiders. Ook is gevraagd naar de rol van het bestuur/bovenschools management. Deze onderwerpen sluiten aan op de in paragraaf 6.1 genoemde onderzoeksvragen.

De gesprekken zijn uitgevoerd door twee onderzoekers. Met één van de scholen zijn, om redenen van efficiëntie de interviews telefonisch gedaan. Voor de andere vier scholen hebben de gesprekken op de locatie van de school zelf plaatsgevonden. De onderzoekers hebben van de gesprekken een verslag geschreven dat ter autorisatie aan de betrokkenen is gestuurd. De geïnterviewden hebben in de verslagen hun correcties en aanvullingen aangegeven die zijn verwerkt voordat verdere analyse van de informatie plaatsvond¹².

De beschrijvingen van de vijf scholen zijn steeds gebaseerd op zowel de door de scholen aangeleverde documenten, als de uitgewerkte verslagen van de gesprek-

¹² Over het algemeen waren er geen of slechts kleine aanpassingen nodig.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

ken met leerkrachten uit onder- en bovenbouw, schoolleiding, ib-ers en/of taalcoördinatoren.

6.3 Resultaten: vijf voorbeelden van meetgestuurde onderwijspraktijk

De casestudiescholen hebben elk hun eigen invulling van meetgestuurd onderwijs. Om de voorbeelden goed tot hun recht te laten komen, volgt hierna een beschrijving per school van de verschillende aspecten van meetgestuurd onderwijs. Elk 'portret' heeft als opbouw: een korte typering van de school, en vervolgens de beantwoording van de onderstaande vragen:

- (1) Welke leerstandaarden en/of streefdoelen hanteert de school?
- (2) Welke gegevens verzamelt de school over de taal- en rekenprestaties van de leerlingen?
- (3) Hoe worden die gegevens geanalyseerd en besproken?
- (4) Hoe worden de resultaten teruggekoppeld en ingezet voor eventuele onderwijsverbeteringen?
- (5) Zijn er effecten zichtbaar van meetgestuurd werken op de taal- en rekenprestaties? Zijn er daarnaast nog andere effecten?
- (6) Welke rol ziet men voor het schoolbestuur en het bovenschools management?

De portretten geven een beeld van de *praktijk* van meetgestuurd onderwijs. De *opvattingen* van de geïnterviewden van de verschillende scholen over het draagvlak, de risico- en succesfactoren en de benodigde competenties vertoonden zoveel overeenkomsten dat zij zijn samengenomen. De resultaten daarvan zijn bij elkaar gezet in paragraaf 6.4.

6.3.1 School 1

School 1 is een openbare school in Amsterdam Zuid-Oost. De school is het resultaat van een recente fusie tussen twee scholen. Uit de taalbeleidsplannen van beide scholen, die zijn opgesteld in het schooljaar 2006-2007 komt naar voren dat de scholen samen ongeveer 500 leerlingen tellen. Ruim 80% van deze leerlingen is allochtoon en gewichtenleerling (0.9). Het merendeel van de leerlingen heeft een Surinaamse, Ghanese, Antilliaanse of Dominicaanse achtergrond. Met name één

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

van de scholen kende een hoog leerling-mutatiecijfer (per jaar zo'n 60 nieuwe en 60 vertrekkende leerlingen), als gevolg van afbraak van woningen in de buurt rond de school. Voor de hele school geldt dat de 4-jarigen die instromen veelal nog geen of weinig Nederlands spreken.

In Amsterdam is de Taalpiloot ingebed in een integraal, stedelijk taalbeleidsplan met als primair doel het verhogen van de taal-/leesresultaten. Kern van de aanpak in het voorbereidingsjaar was deelname van de scholen aan een training rond woordenschatontwikkeling (Met woorden in de weer), gericht op het vergroten van leerkrachtvaardigheden, een coachingstraject met klassenconsultaties en een traject voor directeuren om het taalbeleid ook daadwerkelijk vorm te geven. (Projectplan Pilots Taalbeleid 2007-2010).

Bij de start van de deelname aan de pilots taalbeleid heeft de school in kaart gebracht welk percentage van de leerlingen per leerjaar een D- of E-score haalt op de toetsen uit het Cito-leerlingvolgsysteem¹³ die de school gebruikt. Ook is een overzicht gemaakt van de resultaten op de AVI-toetsen¹⁴ en zijn de resultaten op de Entree- en Eindtoetsen van het Cito geanalyseerd. Op basis van die gegevens heeft de school verbeterpunten, activiteiten en gewenste resultaten geformuleerd. Een belangrijke interventie die is ingezet binnen de pilot Taalbeleid is het trainen van de leerkrachten in de methodiek van Met woorden in de weer¹⁵, daarnaast is ingezet op verbetering van het voortgezet technisch lezen via de implementatie van de methode Estafette, en is het VVE-programma Piramide dat al in groep 1

¹³ De uitslagen van de toetsen uit het Cito-leerlingvolgsysteem worden omgezet naar letterscores A t/m E. Deze normeringen worden landelijk gebruikt en betekenen het volgende:

A-score = goed tot zeer goed: de 25% hoogst scorende leerlingen

B-score = voldoende tot goed: de 25% net boven het landelijk gemiddelde scorende leerlingen

C-score = matig tot voldoende: de 25% net onder het landelijk gemiddelde scorende leerlingen

D-score = zwak tot matig: de 15% ruim onder het landelijk gemiddelde scorende leerlingen

E-score = zeer zwak tot zwak: de 10% laagst scorende leerlingen

¹⁴ AVI-toetsen meten het niveau van technische leesvaardigheid van de leerlingen. Er worden 9 AVI-niveaus onderscheiden, gemiddeld genomen bereiken leerlingen de volgende niveaus:

juni groep 3: AVI-2

juni groep 4: AVI-4

juni groep 5: AVI-8

november groep 6: AVI-9

Zie verder: www.cito.nl, avi, of www.taalsite.nl

¹⁵ Voor meer informatie zie: www.metwoordenindeweer.com

Opbrengricht werken; over de waarde van meetgestuurd onderwijs

en 2 werd gebruikt, nu ook doorgevoerd in groep 3. Er wordt bij de kleuters en in groep 3 specifieke aandacht besteed aan de tien aspecten uit de leerlijn beginnende geletterdheid van het Expertisecentrum Nederlands¹⁶. Inmiddels is ook een start gemaakt met interactief taalonderwijs op teamniveau en invoering van het protocol leesproblemen en dyslexie.¹⁷

Gesproken is met: twee leerkrachten uit de onderbouw (groep 2 en 3), twee leerkrachten uit de midden/bovenbouw (groep 6 en 8), twee ib-ers, de taalcoördinator en de directrice.

Leerstandaarden en streefdoelen

De school werkt met de normeringen uit het Cito-LVS voor rekenen en taal (spelling, woordenschat en begrijpend lezen) en formuleert als streefdoel: maximaal 10% leerlingen op E-niveau en 15% op D-niveau: "de rest moet gewoon hoger" vinden de schoolleider, ib-ers en taalcoördinator. Ook in het Taalbeleidsplan is voor het schooljaar 2008-2009 een doelstelling voor de toetsen uit het Cito-LVS geformuleerd: gemiddeld B-niveau, en maximaal 10% op D/E-niveau). Voor het technisch lezen is per leerjaar vastgesteld wat het streef- en het minimumniveau is dat de leerlingen op de AVI-toetsen zouden moeten bereiken. Voor eind groep 3 is het streefdoel dat de leerlingen lezen op AVI-2 niveau, met als minimumbeheersingsniveau AVI-1.

Voor de kleuters gelden naast de normen van de Cito-LVS-toetsen (i.c. Taal voor Kleuters en Ordenen) de doelen zoals Piramide die aangeeft, de tussendoelen uit de leerlijn beginnende geletterdheid, en de woordclusters die zijn behandeld via de methodiek van Met woorden in de weer. Van de tussendoelen beginnende geletterdheid is een eigen vertaling gemaakt naar streefdoelen, in groep 2 geldt bijvoorbeeld als streefdoel: de leerling beheerst 10 letters passief.

De school heeft veel anderstalige leerlingen, maar kiest niet voor het toepassen van een soepeler normering. De resultaten worden altijd afgezet tegen de landelijke normen, ook voor de AVI-toetsen: "We hanteren hoge streefdoelen met daarnaast een minimumdoelstelling voor de leerlingen aan de onderkant."

¹⁶ Voor meer informatie zie: www.expertisecentrumnederlands.nl, producten, tussendoelen beginnende geletterdheid.

¹⁷ Eveneens ontwikkeld door het Expertisecentrum Nederlands. Meer informatie is te vinden via de bovengenoemde website, bij producten: leesproblemen en dyslexie. Het gaat om een stappenplan voor observatie en toetsing van leesvaardigheden en vaardigheden die daaraan voorafgaan.

Gegevens over taal- en rekenprestaties

De toetsen uit het Cito-LVS zijn een belangrijk instrument om zicht te krijgen op de vorderingen van de leerlingen. Voor de kleuters zijn er, naast de toetsen Taal voor Kleuters en Ordenen, ook minimaal tweemaal per jaar leerkrachtobservaties van alle leerlingen met eigen observatie-instrumenten: een verkorte versie van het observatie-instrument van Piramide, aangevuld met een zelfontwikkelde lijst op basis van de tussendoelen beginnende geletterdheid. Alle informatie, dus ook leerkrachtobservaties als: 'kan een schaar hanteren' en 'kan een verhaaltje vertellen', en de gegevens over woordenschatontwikkeling op basis van toetsen die aansluiten op de thema's waaraan is gewerkt -eens per drie of vier weken- worden bewaard in het leerlingvolgsysteem.

In groep 3 en hoger wordt eveneens gewerkt met een combinatie van methodeonafhankelijke toetsen uit het Cito-LVS (Begrijpend Lezen, Woordenschat en Rekenen), AVI en DMT¹⁸, en methodegebonden toetsen. Voor technisch lezen gaat het om de methodetoetsen die horen bij de kernen van Veilig Leren Lezen. De school maakt daarvoor gebruik van de toetsite van Veilig Leren Lezen. Voor het voortgezet technisch lezen werkt de school met Estafette en de daarbij behorende toetsen. Voor de hogere groepen gebruikt men de methodetoetsen uit Taalleeland. De herfst- en lentesignaleringen uit het protocol Leesproblemen en Dyslexie worden uitgevoerd. Twee keer per jaar zijn er de officiële AVI-momenten, maar als leerlingen er sneller aan toe zijn, wordt de AVI-toets ook wel tussendoor afgenomen om te voorkomen dat het tempo voor hen te traag wordt.

Voor rekenen gebruikt de school, naast de methodeonafhankelijke toetsen, ook de toetsen uit de methode Wis&Reken. Sinds dit schooljaar is er in groep 6 een schakelklas ingericht, de leerlingen in die groep maken nu ook computergestuurde, interactieve opdrachten voor rekenen waarvan de resultaten worden besproken in het schakelklassenoverleg. Bovendien observeren de leerkrachten de motorische en sociaal-emotionele ontwikkeling van alle leerlingen. In groep 7 wordt de Cito-Entreetoets afgenomen, en in groep 8 de Cito-Eindtoets.

Alle toetsresultaten worden, per leerling, opgeslagen in een groepsmap. In de groepsmap zijn bovendien de leerlingenobservaties opgenomen en worden ook de gegevens van andere instanties bewaard, zoals verslagen van de logopedist, GGD,

¹⁸ DMT = drieminutentoets, een toets voor leestehnik en leestempo

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

de schoolmaatschappelijk werkster. Tot nu toe is het een papieren dossier dat mee gaat met de leerlingen naar elk volgend leerjaar. Elke leerkracht heeft zo dus voor elk kind een goed overzicht over het gehele ontwikkelingsverloop en eventuele problemen.

De digitale versie van het LVS is nog niet toegankelijk voor iedereen. De ib-er voert alle gegevens in het digitale systeem in en maakt de overzichten. De leerkrachten tekenen wel zelf de grafieken van de ontwikkeling van hun leerlingen op basis van de Cito-LVS-toetsen, in het papieren dossier. Daarvoor is bewust gekozen omdat zij op die manier veel meer informatie onder ogen krijgen dan wanneer zij van een ib-er alleen een overzicht ontvangen van de leerlingen met een D- of E-score. De leerkrachten kunnen bijvoorbeeld meteen zien of er bij een leerling misschien toch sprake was van vooruitgang (wanneer hij of zij bijvoorbeeld van een E-score aan de onderkant, naar een hoge E-score gaat), of juist van achteruitgang dan wel een stagnering in de ontwikkeling van het kind.

Analyse en bespreking van de resultaten

De ib-er maakt, na de invoering van de LVS-resultaten in het digitale systeem, overzichten op groeps- en schoolniveau en signaleert de opvallende ontwikkelingen in relatie tot het landelijk gemiddelde en de prestaties van dezelfde groep in het vorige schooljaar. Het verslag over de hele school heen laat zien op welke gebieden er meer ingezet moet worden. De school- en groepsoverzichten worden, tot nu toe, besproken met het gehele team. Wel is de school op zoek naar een wat efficiëntere aanpak van die besprekingen. Men denkt daarvoor aan een voorselectie door het managementteam van de meest opvallende zaken die aansluiten bij de speerpunten in de schoolontwikkeling op dit moment: taal en rekenen.

De ib-er en de leerkracht samen stellen de groepsstaten op, bekijken de ontwikkelingscurves van zorgleerlingen en werken de handelingsplannen uit. De ib-ers, plus enkele van de leerkrachten hebben bovendien de cursus 'Haal meer uit je rekentoets'¹⁹ gevolgd. Daarmee is het mogelijk de scores op de rekentoetsen wat

¹⁹ De schoolbegeleidingsdiensten in Utrecht en Amsterdam hebben bij elke rekentoets een foutenanalyse ontwikkeld gebaseerd op de vaardigheidsprofielen van het CITO. Met behulp van de foutenanalyse kan de leerkracht een overzicht krijgen van het soort fouten dat een leerling gemaakt heeft. Door het invullen van de bijbehorende groepsstaat, wordt bovendien zichtbaar welke problemen bij meer leerlingen of in de gehele klas spelen. De publicaties zijn te bestellen via www.JIP4all.nl

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

nauwkeuriger te analyseren. Het maakt zichtbaar waar voor individuele of groepjes leerlingen de belangrijkste problemen nog zitten, zodat een handelingsplan kan worden samengesteld. Ook de leerkracht van groep 3 geeft aan dat zij analyses maakt op basis van vragen als: Welke kinderen vallen uit op welk onderdeel en hebben dus individueel of groepsgewijs extra hulp nodig? Zijn er onderdelen die door het merendeel van de kinderen nog niet zijn begrepen? “Als de meeste leerlingen bijvoorbeeld de bussommen nog niet blijken aan te kunnen, dan moet ik als leerkracht daar iets mee.” Een enkele keer levert het bekijken van de ontwikkelingscurves van leerlingen algemenere aandachtspunten op, die worden dan voorgelegd en besproken in de betreffende bouwvergadering. De school heeft geen schoolbrede kindbesprekingen meer.

Terugkoppeling en onderwijsverbeteringen

De interpretaties van de gegevens leiden tot beslissingen over het onderwijs op verschillende niveaus: individueel, groepsgewijs en voor de hele school.

De gegevens worden in de eerste plaats gebruikt om per kind na te gaan of er sprake is van voldoende groei en ontwikkeling. De leerkracht bekijkt welke leerlingen eventueel verlengde instructie nodig hebben in de groep, en wie er gebaat zijn bij extra hulp buiten de groep waar wat meer kan worden geboden dan alleen de groepsinstructie. Met name de leerlingen die op D- en E-niveau presteren krijgen extra aandacht. Vervolgens wordt de ontwikkeling van deze kinderen opnieuw bekeken en getoetst en wordt de vraag gesteld: wat is er nog meer nodig? Eventueel schakelt de leerkracht de ib-er in, of vraagt zij een observatie aan om te zien of het gaat om een ‘zorgkind’. Als in een groep zorgleerlingen (SBO-, SO- of rugzakkinderen) zitten op E-niveau, dan wordt voor die leerlingen een individueel handelingsplan gemaakt. Zolang het mogelijk is, volgen ook de zorgleerlingen het normale programma binnen een ‘minimum-maximum range.’ Pas vanaf groep 5 gaat de school voor deze leerlingen eventueel op een tweede leerlijn zitten, waarin de landelijke normeringen worden losgelaten. Daaraan vooraf gaat een onderzoek naar IQ, leerachterstanden en natuurlijk een oudergesprek. Uiteindelijk streefdoel voor de groep echte zorgleerlingen is: uitstroom op niveau van groep 7, naar het VMBO/LWOO, en een zo hoog mogelijk niveau van sociale redzaamheid. De leerkrachten gebruiken de individuele leerlingresultaten bovendien bij de overdracht naar de volgende groep. Zij informeren elkaar over wat

Opbrengricht werken; over de waarde van meetgestuurd onderwijs

welk kind al kan en wat eventueel nog niet, wat daarvan de achtergrond is en wat de beste benadering is.

De analyses leiden in de tweede plaats tot het ontwikkelen van groepshandelingsplannen door de leerkrachten zelf. De ib-ers geven aan dat de leerkrachten: "over het algemeen heel nieuwsgierig zijn naar de resultaten en meteen zelf al bedenken hoe zij hun onderwijs daarop moeten aanpassen." Voor Veilig Leren lezen bijvoorbeeld worden de groepsgegevens uitgedraaid en vergeleken met de landelijke gemiddelden. De leerkracht van groep 3 benadrukt dat het belangrijk blijft om daarnaast te kijken naar de individuele leerlinggegevens: "Als het groeps-gemiddelde goed genoeg is, is er nog altijd aandacht nodig voor de individuele leerlingen die daarbij achterblijven." Eén van de leerkrachten uit de bovenbouw geeft aan dat zij na de methodegebonden toetsen regelmatig een diagnostisch gesprekje houdt met elke leerling om na te gaan waar precies de problemen zitten. Op basis daarvan vindt dan, waar nodig, remediëring plaats, ofwel klassikaal ofwel individueel.

Deze leerkracht kan dat doen omdat zij dit jaar lesgeeft aan de (groep 6-)schakelklas met 12 leerlingen. De keuze om deze schakelklas te vormen is voortgekomen uit de signalering dat de rekenscores sterk achterbleven sinds er een sterker beroep wordt gedaan op de taalvaardigheid van de leerlingen. De directrice licht de keuze van de school als volgt toe: "Het lijkt erop dat de trend van tegenvallende rekenprestaties vooral is ingezet sinds het Cito een nieuwe toets heeft ontwikkeld waar een sterker accent ligt op het talige rekenen. Realistisch rekenen is voor onze, anderstalige leerlingen, een stuk moeilijker. Het uitvoeren van de bewerkingen lukt hen over het algemeen wel. Bovendien zit het realistisch rekenen nog onvoldoende in de methodes. Begrippen als 'vergelijk', 'minstens', 'via' of 'per' herkennen de leerlingen niet meteen. Via de methodiek van Met woorden in de weer gaan we nu werken aan het goed aanleren van dit soort rekenbegrippen. Daaraan werken we op het moment samen met de reken- en taal-specialist van het ABC die bij de school betrokken zijn." Hier gaat het dus om een beslissing op schoolniveau. De schakelklas is ingericht om de 'lage C-ers' die in potentie met behulp van wat meer gerichte instructie een hoger niveau zouden kunnen bereiken, uit de 'gevarenzone' van uitstroom op LWOO-niveau te houden. Het accent ligt in deze groep dan ook zowel op de taalontwikkeling, als op versterking van het rekenonderwijs. De leerkracht van de schakelklas is heel enthousiast over de mogelijkheden die deze werkwijze biedt om aan de leerlingen

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

extra/voldoende aandacht te geven. Zij verwacht dat het effect daarvan zeker zichtbaar zal zijn straks bij de eindtoets en het uitstroomniveau in groep 8: "Het gaat om gemotiveerde leerlingen, die zelf willen leren en leerbaar zijn, die het wel in zich hebben om verder te komen, maar waar dat in een kleine groep beter uitgemaakt kan worden. Er zitten geen leerlingen met gedragsproblemen in de groep, en de groep is echt een eenheid geworden, gaat heel goed met elkaar (en met hun juf) om: er is rust en er wordt geleerd."

Effecten

(a) Effecten op de taal- en rekenprestaties

Of het meetgestuurd werken bijdraagt aan verbetering van de taal- en/of rekenprestaties is nog moeilijk te zeggen. De school ziet wel duidelijke verbeteringen bij het technisch lezen en hoopt ook op vooruitgang op de woordenschattoetsen wanneer in januari 2009 de nieuwe Cito-woordenschattoetsen gebruikt kunnen worden. De looptijd van de schakelklas is nog te kort om al tot verbetering in de rekenprestaties geleid te kunnen hebben, maar er is zeker meer inzicht ontstaan in de onderdelen die voor de leerlingen problematisch zijn. Op de Cito-toetsen in groep 8 ziet de school verschillen die samenhangen met gehanteerde methodes. De resultaten van de leerlingen die met de methode Wis & Reken hebben gewerkt, lijken aan het einde van hun schoolloopbaan bij te trekken ten opzichte van de normering. Bij de rekenmethode die op de andere locatie werd gebruikt, was dat niet het geval. De keuze voor een overstap naar Wis & Reken voor de hele school was overigens al gemaakt voordat dit effect was geconstateerd.

(b) Overige effecten

In de woorden van de leerkracht van groep 2: "Het succes zit in het leren analyseren van je toetsresultaten, en het meer kritisch kijken naar je onderwijs. Dat maakt het goed mogelijk om je onderwijs aan te passen. En als je dan ziet dat de leerlingen sterk vooruitgaan in hun toetsscores, dan is dat motiverend voor de leerkracht. Het stimuleert je om de lessen nog steeds beter voor te bereiden en allerlei spullen mee te nemen die daarbij ondersteunend of illustratief zijn. Als je ziet wat de kinderen bereiken, dat ze zelf nu een boekje gaan pakken en 'lezen', dan word je ook als leerkracht enthousiast."

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

De rol van het schoolbestuur en bovenschools management

Op het moment heeft de school te maken met een bestuur op afstand, en een bovenschools management met veel oog voor de cijfers en alle details daarvan. De school ziet echter als de meest wenselijk taakverdeling tussen bestuur en bovenschools management dat (a) de school zorgt voor het bijhouden van de resultaten en de afweging over wat er binnen de context van het eigen onderwijs moet gebeuren om die resultaten te verbeteren, terwijl (b) het bestuur en bovenschools management daarvan notie nemen, de verschillende scholen onderling eventueel vergelijken, zorgen voor intervisie en overdracht, en vooral voor het faciliteren van de benodigde schoolontwikkeling.

6.3.2 School 2

School 2 is een rooms-katholieke school in Zeeland. Het schoolteam bestaat, inclusief de directeur en de ib-er uit 10 teamleden. In totaal heeft de school op het moment 71 leerlingen, waarvan 84% allochtone leerlingen en 16% autochtone leerlingen uit sociaal zwakkere en financieel minder welgestelde gezinnen. De hele leerlingenpopulatie is zeer taalzwak. De schoolpopulatie weerspiegelt de samenstelling van de bevolking in de omliggende wijk waarin relatief veel sprake is van achterstanden en problemen. De school is gevestigd in een 52-jaar oud schoolgebouw, maar hoopt per 1 maart 2009 te kunnen verhuizen naar nieuwbouw. De school vestigt zich dan in een wijkgebouw waarin allerlei voorzieningen samengebracht worden, onder andere een peuterspeelzaal. De school is brede school en heeft ondermeer een VVE-aanbod en verlengde schooldagactiviteiten.

De school is momenteel een taalpilotschool, en neemt deel in het project 'Meer taalkansen voor Zeeuwse kinderen.' (ondersteund door het RPCZ). Daarin ligt het accent op begrijpend lezen en woordenschat. Eerder al was de school onderwijskansenschool, maar door veel 'gerommel', wisselingen en onrust is men pas in het laatste jaar van onderwijskansen toegekomen aan het daadwerkelijk werken aan schoolverbetering. De huidige directeur is in 2005 gestart met de opdracht van het bestuur om te zorgen voor 'rust in de tent' en om de school weer op de kaart te zetten. Vanaf dat moment is er sprake van opbouw. Het team staat positief tegenover de ingezette veranderingen. Het nieuwe schoolplan voor de periode 2007-2011 heeft als motto meegekregen: 'Meer dan school alleen'.

Gesproken is met: de directeur, één leerkracht uit de onderbouw (groep 2), één leerkracht uit de middenbouw (groep 5/6).

Leerstandaarden en streefdoelen

De leerstandaarden die de school hanteert zijn in de eerste plaats gebonden aan de doelstellingen uit de methodes, die weer aansluiten op de kerndoelen voor het basisonderwijs (Ministerie van OCW). Voor rekenen wordt gewerkt met de methode 'Alles telt'. Voor het aanvankelijk lezen gebruikt men 'De Leeslijn' en voor de overige vakken (taal, begrijpend lezen en de zaakvakken), sinds het huidige schooljaar de geïntegreerde methode 'Alles in één'.

Daarnaast geven de toetsen uit het Cito-LVS, die afgenomen worden volgens de toetskalender, richtlijnen aan voor de te bereiken doelen bij de leerlingen. Het streven is dat minimaal 60% van de leerlingen op C-niveau of hoger zit bij de toetsen uit het Cito-LVS. Wanneer leerlingen op de Cito-LVS toetsen een D- of E-score halen wordt actie ondernomen. Ook op schoolniveau wordt bekeken: wat kunnen de kinderen nog niet dat ze volgens de normen van de Cito-toetsen al wel zouden moeten beheersen, en waar wordt het gewenste niveau wel gehaald? "De belangrijkste vraag is steeds: bereiken we wat we willen bereiken, en zo niet wat zouden we dan moeten veranderen?" geeft de directeur aan.

Voor de kleuters hanteert men, naast Taal voor Kleuters uit het Cito-LVS, de tussendoelen beginnende geletterdheid als richtlijn. De doelstelling is dat alle leerlingen deze tussendoelen beheersen aan het eind van groep 2. Voor rekenen neemt men, naast de resultaten op de Cito-LVStoets Ordenen, als richtlijn de tussendoelen die men op papier heeft gezet voor het aanvankelijk/voorbereidend rekenen. In die tussendoelen is ook expliciet aandacht voor de begrippen die gebruikt worden in de Cito-toetsen -zoals meer, minder, het grootst- aangezien een groot aantal leerlingen een andere moedertaal dan het Nederlands heeft: "Het is belangrijk dat de kinderen met dit soort taalgebruik vertrouwd raken, zodat de toetsen een reëel beeld geven van hun vaardigheden op het gebied van (aanvankelijk) rekenen," aldus de leerkracht van groep 2.

Gegevens over taal- en rekenprestaties

De school gebruikt om inzicht te krijgen in de taal- en rekenprestaties ten eerste de toetsen uit het Cito-LVS: Ordenen, Taal voor Kleuters, Begrippentoets III, Woordenschattoets, Taalschaal, Rekenen, Lezen met begrip en Begrijpend Lezen. De gegevens daarvan worden ingevoerd in het bijbehorende computerprogramma, en op basis daarvan worden analyses uitgevoerd en grafieken gemaakt. Voor de

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

kleuters wordt aan het eind van groep 1 voor zowel Taal voor Kleuters als Ordenen een foutenanalyse uitgevoerd op basis waarvan de leerkracht in groep 2 heel gericht en sturend kan werken aan die onderdelen die zij nog niet beheersen: elke leerling heeft een eigen werkmapje met opdrachten waaraan zij in elk geval twee keer per week werken. De leerkracht schrijft daar opmerkingen bij over het leerproces en geeft aan of er sprake is van vooruitgang. Halverwege groep 2 worden dezelfde toetsen opnieuw afgenomen en de nieuwe foutenanalyses vormen dan de richtlijn voor de gerichte opdrachten voor het kind. Veel van deze opdrachten zijn gericht op het aanvankelijk rekenen. De taalontwikkeling krijgt op andere momenten, in de kring bijvoorbeeld, al veel extra aandacht via taalspelletjes gericht op auditieve discriminatie, rijmen etc. De leerkracht geeft aan dat de werkwijze zo gegroeid is omdat het noodzaak is, ook al zou zij als kleuterleerkracht liefst wat meer aansluiten bij de eigen keuzes van kinderen: “Voor onze leerlingen is wat meer sturing en stimulans nodig. Uit zichzelf zoeken ze weinig uitdaging.” Naast de methodeonafhankelijke toetsen, neemt de school ook de methodetoetsen af. De resultaten daarvan worden door de leerkracht ingevoerd in de toetsmanager op de computer²⁰, een set in Excel gemaakte toepassingen voor de registratie van de leerlingengegevens met verschillende mogelijkheden om die gegevens te ordenen zodat zichtbaar wordt welke leerlingen extra uitleg of zorg nodig hebben. Voor de methode ‘Alles in één’ maakt men daarbij gebruik van de Excel sheets die een handige leerkracht van een andere school heeft ontwikkeld. De indruk is echter dat het nog efficiënter zou kunnen en daarom wordt momenteel geëxperimenteerd met een eigen registratiesysteem en toetsmanager.

Analyse en bespreking van de resultaten

De directeur is -momenteel nog- degene die de analyses uitvoert op basis van de Cito-toetsen. Hij brengt de groepsoverzichten in kaart en heeft daarmee een goed zicht op de stand van zaken. De belangrijkste informatie koppelt hij via een kort feedbackgesprek terug aan de groepsleerkracht. Dat leidt tot een informele uitwis-

²⁰ Zie bijvoorbeeld: www.allestelt.nl/assets/allestelt/toetsmanager/manager1.html

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

seling over wat goed gaat, wat tegenvalt en wat daarvoor verklaringen zouden kunnen zijn. In de toekomst is het de bedoeling dat de leerkrachten zelf de gegevens gaan invoeren en leren welke de mogelijkheden er zijn om de groepsgegevens in kaart te brengen. De ib-er pakt de overzichten verder op en zorgt voor de diagnosticering van individuele leerlingen, en de bijbehorende handelingsplannen.

Twee keer per jaar worden de LVS-resultaten bekeken met het hele team gezamenlijk. Er worden dan dwarsdoorsnedes en trendanalyses besproken over de hele school heen. Centraal in deze besprekingen staan de groepsgegevens. Ook vergelijkingen met de landelijke gegevens komen aan de orde: de vorderingen en prestaties worden vergeleken met de landelijke normen vanuit het Cito-LVS. De overzichten worden gepresenteerd op een scherm en de teamleden geven elkaar feedback en informatie over de achtergronden van de resultaten. Er wordt gezamenlijk gebrainstormd over verklaringen en verbeteringsmogelijkheden. Eén van de leerkrachten verwoordt het belang van de teambesprekingen als volgt: “Naast het groepsprofiel wordt ook een alternatief groepsprofiel gelegd met de lijn waarop de kinderen met hun prestaties zouden moeten zitten. Dan besef je soms als leerkracht dat je stof aan het behandelen bent die veel te hoog gegrepen is, en dat het nodig is om terug te grijpen op de leerstof van het jaar ervoor om niet de aansluiting te verliezen bij wat de kinderen weten en kunnen. Aan de andere kant moet je soms de lat juist wat hoger leggen om je leerlingen voldoende uitdaging te bieden.”

In de leerlingenbesprekingen van de ib-er en de groepsleerkracht komen met name de zorgleerlingen aan bod: voor de leerlingen die uitvallen worden de gegevens uitgezet in een grafiek over een langere periode. Zo wordt duidelijk of er sprake is van een stijgende lijn, terugval of een op- en neergaande beweging. In de zorgstructuur onderscheidt de school vier niveaus: (a) de geen-zorg-leerlingen, (b) de aandachtsleerlingen, (c) de zorgleerlingen voor wie met een handelingsplan wordt gewerkt, en (d) de rugzakleerlingen, voor wie er naast een handelingsplan, ook ambulante begeleiding is.

Terugkoppeling en onderwijsverbeteringen

De besprekingen van de resultaten leiden tot beslissingen over het onderwijs op de school: de grote ommezwaai van het werken met methodes naar projectonderwijs is daarvan een duidelijk voorbeeld. De keuze voor ‘Alles in één’ komt ener-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

zijds voort uit de zwakke toetsresultaten bij taal, anderzijds uit de ervaringen van de leerkrachten: “We waren steeds bezig de lesjes aan te passen aan de kinderen om te zorgen dat de leerstof voldoende begrijpelijk was. Er was in het team behoefte aan meer interactief onderwijs met name vanwege de woordenschatontwikkeling. Dat is nu veel beter mogelijk doordat taal in het hele thema verweven is: je kunt op die manier lekker uitpakken met de kinderen, de hele klas aankleden in de sfeer van het thema en zo zorgen voor meer verdieping van de woordkennis. Welke resultaten het straks oplevert, moeten we nog afwachten, maar de leerlingen zijn in elk geval zichtbaar meer betrokken bij het onderwijs. Ze nemen zelfs dingen mee van thuis, dat gebeurde eerder helemaal niet. Op die manier bouw je hun kennis op, en krijg je veel meer diepgang in plaats van hapsnap leren.” De overstap naar De Leeslijn is eveneens voortgekomen uit het bestuderen van de toetsresultaten: met de vorige methode (Veilig Leren Lezen) bleken de leerlingen teveel globaal te blijven lezen.

In de functionerings- en beoordelingsgesprekken met de groepsleerkrachten komen de gezamenlijk gemaakte afspraken terug. Dit jaar is bijvoorbeeld gestart met het werken met groepsplannen voor verbetering van het technisch lezen. Er is een overzicht gemaakt van de scores op de DMT en voor elke leerling individueel wordt het beoogde doel voor het volgende toetsmoment vastgesteld. De leerlingenniveaus en bereikte resultaten worden vervolgens meegenomen in de functioneringsgesprekken met de vraag: wat is er bereikt, wat niet, waardoor komt dat, wat is er nodig om dat wel te bereiken? Het uitvoeren van voorafgaande klassenbezoeken door de directeur is daarbij van belang, maar praktisch –nog- niet altijd mogelijk.

Effecten

(a) Effecten op de taal- en rekenprestaties

Of er werkelijke verbetering optreedt in de taal- en rekenprestaties als gevolg van de onderwijsvernieuwing en het meer meetgestuurd werken, zal de toekomst moeten uitwijzen. Daarvoor zijn de vernieuwingen nog te kortgeleden ingezet.

(b) Overige effecten

Zichtbaar is voor de leerkrachten en directie wel dat het themagericht werken in groep 1 t/m 4, en het projectonderwijs in groep 5 t/m 8 leiden tot veel meer overleg door de hele school heen, nu nog vooral over praktische en didactische af-

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

stemming. Bovendien is er een grotere betrokkenheid merkbaar van de leerlingen bij de leerstof. Vanuit de aanname dat een sterkere betrokkenheid gunstig is voor de leerresultaten, zou dat tot een verbetering van de taal- en rekenprestaties moeten leiden. De eveneens kortgeleden ingezette aanpak van het werken met groepsplannen voor de DMT, heeft de leerkrachten meer gespitst gemaakt op de leerlingenresultaten. “Dat gebeurt eigenlijk vanzelf al als je vraagt: wat belooft je?” volgens de directeur.

Het Cito-LVS is een belangrijke leidraad. Dat geeft aan wat de leerlingen bijvoorbeeld aan het eind van groep 4 op het gebied van taal moeten kunnen. De school is overgestapt van het volgen van het kind in zijn/haar ontwikkeling, naar meer resultaatgericht onderwijs gegaan. “De achtergrond van de leerlingen is zo dat van hen niet zomaar vanzelfsprekend verwacht kan worden dat zij de volgende stap in hun ontwikkeling zelf zetten. Het is belangrijk hen expliciet volgende ontwikkelingsstappen aan te bieden” licht één van de leerkrachten toe. De meer sturende aanpak blijkt positieve effecten te hebben op de vorderingen van de leerlingen. Het gestructureerde taalaanbod in groep 2 heeft geleid tot hogere scores op de LVS-toetsen Taal voor Kleuters en Ordenen.

De rol van het schoolbestuur en bovenschools management

Het schoolbestuur van school 2 is een bestuur op vrijwillige basis, de bestuursleden vervullen hun taak meestal naast hun baan. Het is een heel nabij bestuur, maar men houdt zich niet direct bezig met de leerprestaties en schoolvorderingen. Het bestuur ontvangt de rapportages van het CED voor alle scholen, maar de overzichten worden besproken in de werkgroep onderwijskundige zaken van het directieberaad. Die werkgroep wordt gevormd door de directies en een aantal bestuursleden.

De werkgroep onderwijskundige zaken heeft onlangs overleg gehad met de Inspectie over de uitvoering van het nieuwe Toezichtskader. Daarbij is onder meer besproken dat het aansturen van meetgestuurd onderwijs door het schoolbestuur een beroep doet op vaardigheden in het omgaan met data. In het type schoolbestuur waar school 2 onder valt, beschikken de bestuursleden echter niet per definitie over de daarvoor benodigde competenties en tijd.

6.3.3 School 3

School 3 is een samenwerkingsschool in Rotterdam, vallend onder een protestants-christelijk schoolbestuur. De school staat in een buurt waar onlangs veel huizen zijn gesloopt en waar nog meer huizen op de nominatie staan om gesloopt te worden. De beloofde nieuwbouw van woningen laat op zich wachten. Dit heeft tot gevolg dat het leerlingenaantal op de school enorm is teruggelopen én dat er bovendien een toename is van het aantal zorgleerlingen (uit gezinnen die in de sloopwoningen zijn getrokken). Twee jaar geleden had de school 192 leerlingen, nu zijn dat er maar 113. Het percentage allochtone en autochtone leerlingen bedraagt elk zo'n 50%. Ongeveer 30% van de leerlingen heeft een taalarme achtergrond.

De scholen die deelnemen in de Taalpilots in Rotterdam richten zich vooral op verbetering van woordenschat en begrijpend lezen.

Gesproken is met: vier leerkrachten (van groep 1/2, 3, 5 en 7), van wie er twee tevens ib-er zijn voor respectievelijk de kleutergroepen en de groepen 3 en 4, en met de directeur, de senior-ib-er en de coördinator onderwijs.

Leerstandaarden en streefdoelen

Voor alle vakken en leerjaren streeft de school naar een percentage van 80% van de leerlingen op A-, B- en C-niveau (waarvan minstens 50% op A- of B-niveau) voor de toetsen in het leerlingvolgsysteem, zowel voor rekenen als voor taal. Het gaat daarbij om een doelstelling vanuit het schoolbestuur.

Voor het schooljaar 2008-2009 streeft de school ernaar dat 95% van de leerlingen aan het eind van groep 3 AVI-2 niveau haalt, en eind groep 4 AVI-5, conform de streefdoelen van de AVI-leesniveaus. Voor 2009-2010 heeft de school zich ten doel gesteld dat 90% van de leerlingen aan het eind van groep 5 op AVI-9 zit. Nu geldt dat nog voor ongeveer 50 à 60%. Deze doelstellingen komen voort uit de deelname aan de pilot Taalbeleid. Om het gewenste niveau te bereiken wordt gestructureerd gewerkt aan de ontluikende geletterdheid vanaf de kleutergroepen, met daar de doelstelling dat de kinderen aan het eind van groep 2 minimaal 16 letters kennen.

Voor het overige hanteert de school de kerndoelen zoals die in de methodes aan de orde komen als richtlijn.

Gegevens over taal- en rekenprestaties

Twee keer per jaar worden de toetsen uit het leerlingvolgsysteem afgenomen. Kinderen die een D of E scoren, worden in de loop van het jaar een keer extra getoetst. De leerlinggegevens op basis van de LVS-toetsen van het Cito voor rekenen, taal, woordenschat, spelling, begrijpend, lezen, AVI, DMT, tempotoets rekenen en Leestechiek en tempo, worden ingevoerd in het computerprogramma van het Cito, opgeslagen in ESIS en verstuurd naar de toetsservice van het CED²¹. De school ontvangt vervolgens de resultaten van de leerlingen op papier, onder andere in grafieken en analyseert deze zelf verder.

In de kleutergroepen werkt de school met de methode Ik & Ko²² en de aan deze methode verbonden toetsen. Daarnaast legt de leerkracht in een observatielijst vast hoe de kinderen de aangeboden leerstof hebben verwerkt en hoe zij zich ontwikkelen in andere opzichten zoals gedrag en motoriek. In groep 3 maakt men gebruik van de methodegebonden toetsen voor bijvoorbeeld Veilig Leren Lezen: “Er is een nieuwe site van Veilig Leren Lezen, hierin voeren we de gegevens van de leerlingen van groep 3 in. Na het invoeren verwerkt het programma de resultaten en volgen er allerlei suggesties om de prestaties van kinderen te verbeteren. Het is erg handig, het geeft een objectief beeld van de problemen van een kind,” aldus de betreffende groepsleerkracht.

Ook in de bovenbouw werkt men met een combinatie van genormeerde Cito-LVStoetsen volgens de toetskalender, methodegebonden toetsen (uit bijvoorbeeld Rekenrijk en Taalactief) en observaties op individueel en/of groepsniveau voor taal, technisch lezen en rekenen. Elke leerkracht houdt van de resultaten op de methodegebonden toetsen een eigen administratie bij.

Analyse en bespreking van de resultaten

Er vindt evaluatie plaats van de taal- en rekenprestaties op schoolniveau (in team- en bouwbesprekingen), op groepsniveau (via een cyclus van groepsbesprekingen elke zes weken) en op individueel leerlingenniveau (in de vorm van leerlingensprekingen).

²¹ De CED-Groep is een landelijke educatieve dienstverlenende instelling die onder meer een Toetsservice biedt met monitorrapporten op schoolniveau gebaseerd op de aangeleverde toetsresultaten van de leerlingen. Zie: www.cedgroep.nl (doorklikken naar: onderzoek en ontwikkeling, en vandaar naar onderzoek en evaluatie)

²² Zie: www.ikenko.nl

Opbrenstgericht werken; over de waarde van meetgestuurd onderwijs

Bij de analyses die de school maakt, ligt het sterkste accent op het volgen van de ontwikkeling van de individuele leerlingen. De groepsleerkracht en ib-er analyseren samen de resultaten, daarbij maken zij ook een vergelijking met de prestaties van een leerling in eerdere leerjaren. De ib-er kijkt naar de ontwikkelingscurve van de leerlingen door de hele schoolloopbaan heen: “Als een leerling problemen heeft, ga ik na of dat komt door het onderwijsaanbod, of de leerkracht er een rol in speelt, of er op sociaal-emotioneel terrein iets aan de hand is. Ik ga ook na of het nodig is om de ouders in te schakelen.”

In de bouw- en teamvergaderingen worden de ontwikkelingen van alle leerlingen bekeken en besproken. Gekeken wordt waar kinderen op uitvallen, welke prestaties achterblijven, welke leerlingen meer herhaling nodig hebben, welke kinderen uitdagender onderwijs nodig hebben. Op basis van de resultaten van de leerlingen passen de leerkrachten -waar nodig- hun onderwijs aan: voor de leerlingen die beneden het gewenste niveau presteren, maakt de leerkracht een individueel handelingsplan in overleg met de ib-er, wanneer de hele groep of een deel van de klas achterblijft wordt een groepshandelingsplan opgesteld. De in een handelingsplan vastgestelde doelen moeten over het algemeen na zes tot acht weken zijn behaald. Dan volgt een nieuwe toetsronde om te kijken hoe het gaat. Soms worden al na drie weken de eerste gegevens aan de ib-er gepresenteerd, in de vorm van een verslag door de leerkracht zijn of resultaten van methodegebonden toetsen.

Aan het begin van het schooljaar ontvangen de leerkrachten een overzicht van de D- en E-leerlingen in hun (nieuwe) klas, zodat zij meteen rekening kunnen houden met deze leerlingen.

Voor de Cito-toetsen en de Eindtoets in groep 8 worden de prestaties van de leerlingen met het landelijk gemiddelde vergeleken. Ook gaat men na of de resultaten leerkrachtgebonden zijn of dat het om een zwakke groep gaat. Alle gegevens worden teruggekoppeld naar het bestuur dat de prestaties van de leerlingen van de verschillende scholen globaal met elkaar vergelijkt. Daarnaast hebben de ib-ers overzichten gemaakt van de woordenschatontwikkeling van de leerlingen (voor groep 1 tot en met 4, en groep 5 tot en met 8).

Terugkoppeling en onderwijsverbeteringen

De terugkoppeling van de resultaten op groepsniveau gebeurt via de bouw- en teamvergaderingen. In de afgelopen periode ligt daarbij het accent op het lesonderwijs, zo wordt bijvoorbeeld de ontwikkeling van het fonemisch bewustzijn en

de beginnende geletterdheid van de kinderen in groep 1 en 2 nauwkeurig gevolgd en besproken.

De leerkrachten interpreteren voor zichzelf de resultaten van de individuele leerlingen in hun eigen groep en bespreken deze daarna met de ib-er. Daardoor zijn zij bewuster bezig met hun eigen onderwijsaanbod, zoals één van de leerkrachten beschrijft: "We zien nu beter waar de knelpunten zitten. Onlangs bij de herfstsignalering bleek in groep 3 dat de helft van de kinderen moeite heeft met het maken van zinnen. Daarmee ben ik, voor die leerlingen, gericht aan het werk gegaan, terwijl de anderen op hun eigen niveau verder werken. Ik kan nu maatwerk leveren naar de kinderen. Door zo'n specifieke benadering kun je beter aansluiten op het niveau van de leerling en is het onderwijs efficiënter."

Een andere leerkracht benoemt dat zij in haar onderwijsaanbod ook wel rekening houdt met de inhoud van de toetsen: "Wanneer we weten dat klokkijken in een bepaalde toets aan de orde komt, maar in de methode niet of nauwelijks, dan zorgen we ervoor daar wel aandacht aan te besteden. Als we dat niet zouden doen, dan zouden de leerlingen slecht scoren op de Cito-toets terwijl dat niet nodig is." Verder leidt het overzicht over de prestaties van de leerlingen ertoe dat de leerkrachten wat meer durven te selecteren in de leerstof voor de leerlingen die het goed doen, waardoor het onderwijs voor hen uitdagender wordt. Voor de zwakkere leerlingen wordt juist meer herhaling ingebouwd, of soms iets weglaten zodat er meer aandacht kan worden besteed aan de zaken die echt van belang zijn, zoals klokkijken, in plaats van heel veel rijtjes sommen van dezelfde soort.

Effecten

(a) Effecten op taal- en rekenprestaties

Er is in de twee voorgaande schooljaren gericht aandacht besteed aan het verbeteren van de vaardigheid van de leerlingen in begrijpend lezen (via Nieuwsbegrip en Kidsweek). De effecten daarvan ziet het team terug in de uitslagen op de Toets Begrijpend Lezen (TBL) uit het Cito-LVS: in de groepen 5 t/m 8 zijn de scores op de TBL zonder uitzondering rond de 90% op A-, B- of C-niveau. De scores op de DMT zijn duidelijk verbeterd in alle jaargroepen. In maart hadden bijvoorbeeld alle leerlingen in groep 3 een A, B of C-score.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

(b) Overige effecten

Team en schoolleiding staan positief tegenover meetgestuurd onderwijs: "Het is vanzelfsprekend dat dit tot resultaten leidt. Door het meten kom je tot feedback geven: je koppelt terug naar de kinderen en je geeft hen feedback, daardoor help je de kinderen. Je komt achter bepaalde problemen. Je kunt vervolgens andere mensen inschakelen of andere instrumenten gaan gebruiken. Als je niet meet, ontdek je niet waar de problemen zitten." Een belangrijke verbetering vindt men bovendien de toegenomen openheid: ib-ers, maar ook leerkrachten onderling kijken bij elkaar in de les, de deuren staan open, en er worden video-opnames gemaakt door schoolbegeleiders die in het team worden besproken. Men geeft elkaar tips en feedback en wisselt informatie uit. De sfeer is, anders dan een aantal jaren geleden, veilig voor de leerkrachten. Die veiligheid biedt vervolgens ruimte voor kwaliteitsbeleid en meetgestuurd onderwijs.

De rol van het schoolbestuur en bovenschools management

Het stichtingsbestuur van school 3 heeft veel over het beleid van de school te vertellen, maar kan helaas niet voldoende faciliteiten bieden. "Het probleem is, dat wanneer de stichting de ene school iets geeft, het alle andere scholen ook iets moet geven of dat nou nodig is of niet. Dit is een onwenselijke situatie" aldus de schoolleider. De school heeft behoefte aan veel meer ondersteuning dan er nu geboden wordt: de gezinnen die in de sloopwoningen trekken brengen een complexe problematiek met zich mee, allerlei voorzieningen zijn juist uit de buurt vertrokken, waardoor ouders met al hun problemen naar de school komen. De school is daar niet voor toegerust.

6.3.4 School 4

School 4 is een openbare, buurtgerichte school in een Enschedese nieuwbouwwijk. Het leerlingenaantal is de laatste jaren flink gestegen door de snelgroeiende Vinex-locatie waarin de school ligt. Er zijn nu bijna 500 leerlingen, die overwegend afkomstig zijn uit 1.0 gezinnen met tweeverdienende ouders. Enkele leerlingen komen uit Duitsland.

De schoolbesturen en de samenwerkingsverbanden WSNS in Enschede hebben in overleg met de gemeente gekozen voor een integrale aanpak in de vorm van een taaloffensief, uitgevoerd met geormerkte gelden voor de schoolbegeleidingsdienst, in combinatie met deelname aan de taalpilots. Centraal staat het vergroten

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

van de eigen deskundigheid van de school om de taal/leesprestaties van de kinderen te verbeteren. (Projectplan Pilots Taalbeleid).

Gesproken is met: vier leerkrachten (van groep 2, 3, 4 en 7), twee van deze leerkrachten zijn tevens taalcoördinator), en met de ib-er en de schoolleider.

Leerstandaarden en streefdoelen

De school werkt met de kerndoelen uit het Leesverbeterplan-Enschede. Dit plan is mogelijk gemaakt via de Taalpilots van het PO-Platform en staat onder leiding van het CPS. Het CPS bouwt daarvoor voort op de theoretische uitgangspunten van Kees Verwoort. Het accent ligt op het bereiken van betere resultaten voor het technisch lezen. Het doel voor het schooljaar 2008-2009 is een score van 85% voldoende resultaat bij de leerlingen op de DMT en AVI toetsen. In 2007-2008 is de score 80% gehaald, voor 2010 wordt gestreefd naar 95%. Eind groep 3 wordt gestreefd naar AVI 3, eind groep 4 naar AVI 5 en eindgroep 5 naar AVI 9. Het gaat om een 'Enschede-brede-aanpak', en dezelfde streefdoelen worden gehanteerd op de overige 32 scholen in Enschede. Ook voor de overige toetsen uit het Cito-LVS, op het gebied van taal, spelling en rekenen is het streefdoel dat 85% van de leerlingen presteert op A-, B- of C-niveau.

Gegevens over taal- en rekenprestaties

De school werkt met het Cito-leerlingvolgsysteem van groep 1 tot en met groep 8, inclusief AVI-toetsen. De leerkrachten nemen de toetsen af, de administratie voert de gegevens in en maakt een uitdraai, en de ib-er verzamelt de overzichten van de prestaties bij rekenen, taal, lezen en spelling. Woordenschat wordt buiten beschouwing gelaten omdat daar voor de leerlingenpopulatie geen problemen in zitten. Als eindtoets gebruikt de school de NIO-toets²³ omdat deze een goed zicht geeft op de potentie van de leerlingen. Verder vullen de groep 8-leerlingen de NVP in, een vragenlijst die persoonlijke eigenschappen van kinderen meet.

In groep 1 werkt de school met observatielijsten waarin de resultaten van activiteiten worden vermeld en met de observatielijst sociaal-emotionele ontwikkeling (ook wel 'de Praxislijst' ontwikkeld door Luc Koning). In groep 2 nemen de leerkrachten, naast de LVS-toetsen, de toetsen fonemisch bewustzijn en letterkennis

²³ NIO= Nederlandse Intelligentietest voor Onderwijsniveau, bedoeld voor advisering van de schoolkeuze aan het eind van groep 8 van het basisonderwijs tot en met klas 3 in het voortgezet onderwijs.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

af. Op basis van de prestaties op deze toetsen beslist men welke kinderen in groep 2 al met lezen kunnen beginnen. In de groepen 3 t/m 8 worden naast de LVS-toetsen ook methodegebonden toetsen afgenomen voor lezen, taal en rekenen. De resultaten van de methodegebonden toetsen administreren de leerkrachten zelf.

Analyse en bespreking van de resultaten

Twee á drie keer per jaar bekijkt men in de teamvergadering de trendanalyses op basis van de resultaten van de midden- en eindtoetsen uit het leerlingvolgsysteem. Daarbij worden de resultaten van de parallelgroepen vergeleken en groepsgemiddelden uit de vorige toetsronde gelegd naast de nieuwe resultaten voor rekenen, spelling en begrijpend lezen. Zo wordt de groei van de leerlingen door de jaren heen in kaart gebracht. Ook een korte vergelijking met de landelijke gemiddelden komt aan de orde, maar er ligt een sterker accent op de vergelijking met de andere scholen in Enschede.

De leerkrachten analyseren zelf de resultaten op leerlingniveau. De leerlingen met een D- of E-score voor één of meer vakken worden nader besproken met de ib-er. Ook leerlingen die uitvallen naar boven (A+), of leerlingen met sociaal-emotionele problemen worden besproken. Voor al deze leerlingen maakt de leerkracht een handelingsplan, waarbij de ib-er eventueel nog pedagogische of didactische suggesties geeft. Zo nodig worden ouders ingeschakeld en volgt er een gesprek. De school werkt zowel met persoonlijke handelingsplannen voor individuele leerlingen, als met groepshandelingsplannen voor een groepje leerlingen of de hele groep. Bij een handelingsplan voor de hele groep wordt een programma ontwikkeld voor 2 à 3 maanden en als het nodig is, wordt er tussentijds getoetst.

De toetsresultaten worden bovendien gebruikt voor het indelen van de leerlingen in drie niveaugroepen voor technisch lezen, en in de groepen 5 t/m 8 voor rekenen.

Terugkoppeling en onderwijsverbeteringen

De school heeft een Leesverbeterplan 2008-2011, in samenwerking met het CPS. Daarin wordt gewerkt met het IGDI-model (IGDI = Interactieve Gedifferentieerde Directe Instructie) voor het technisch lezen. Op grond van hun toetsresultaten worden de leerlingen ingedeeld in drie niveaugroepen: er is een A + B groep, een C + D groep en een D + E groep per leerjaar. Voor alle leerlingen wordt het lesdoel concreet benoemd en is er een korte, interactieve groepsinstructie. Voor een

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

beperkt groepje leerlingen geeft de leerkracht een verlengde instructie. Er wordt steeds gezorgd voor voldoende effectieve leestijd. Hetzelfde model gebruikt de school ook voor het rekenonderwijs in de groepen 5 t/m 8. Een belangrijke reden voor het hanteren van het IGDI-model is om de leerlingen die naar boven toe uitvallen (de A+ leerlingen) te kunnen voorzien van uitdagender onderwijs. De invoering van het model op school wordt gedaan door de schoolbegeleidingsdienst, de ib-er en de directie. Na verloop van tijd zullen klassenconsultaties worden uitgevoerd door de schoolbegeleiders, en er vindt intervisie plaats met de andere scholen die met het leesverbeterplan zijn gestart.

De resultaten van de leerlingen worden bijgehouden door de leerkrachten zelf in informatiemappen, die bij de overgang naar de volgende groep worden overgedragen aan de nieuwe leerkracht. De ib-er beheert de mappen van de zorgleerlingen. Tijdens de bouwvergaderingen presenteren de leerkrachten de groepsprestaties. Kern is dat de leerkracht met name de prestaties volgt en de ib-er het proces. Wanneer de ontwikkeling van een leerling stagneert, gaat de leerkracht eerst te rade bij de ib-er en andere collega's op de school, daarna eventueel bij collega's van andere scholen, en als het nodig is wordt ook de schoolbegeleidingsdienst ingeschakeld. Op leerlingniveau vindt afstemming van het onderwijs plaats door de leerkracht zelf, die bijvoorbeeld verdiepingsstof inzet wanneer de rekenprestaties daartoe aanleiding geven, die materiaal zoeken voor de zwakke leerlingen of daarover advies vragen aan de ib-er.

In de functioneringsgesprekken komt het werken met het IGDI-model aan de orde en zijn de resultaten besproken van de leerlingen die achterblijven. Uiteindelijk is het de bedoeling dat deze informatie ook in de beoordelingsgesprekken meegenomen wordt, maar zover is het nu nog niet.

Effecten

(a) Effecten op taal- en rekenprestaties

De leesresultaten op de school zijn verbeterd. In het schooljaar 2007-2008 haalde in alle groepen –met uitzondering van groep 5- zo'n 80% van de leerlingen de gestelde normen voor de AVI en de DMT-toets. Bij de meting in oktober van het huidige schooljaar 2008-2009, was dat percentage voor de groepen 5 t/m 8 al 85% of hoger.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

(b) Overige effecten

De leerkrachten merken bovendien dat de leerlingen meer plezier hebben gekregen in lezen doordat zij meer op maat worden bediend en meer worden uitgedaagd. Het Leesverbeteringsplan heeft opgeleverd dat het team nu veel bewuster bezig is met het technisch lezen, meer inzicht heeft in wat effectief is, meer onderling uitwisselt en beter inziet dat de leerkracht ertoe doet.

“Men ziet nu veel beter wat er gebeurt met de leerlingen. De resultaten zijn helder. In groep 6 bleven bijvoorbeeld de leesresultaten achter. De leerkracht wilde zich daar niet op laten afrekenen. Er volgde een open discussie waarin we zijn nagegaan waardoor het tegenvallende resultaat veroorzaakt kon zijn. Vervolgens is er een discussie geweest met leerkrachten van andere scholen. Daar bleek zich hetzelfde probleem voor te doen. Dat bracht ons tot de conclusie dat de stap van AVI-8 naar AVI-9 erg groot en veel meer tijd kost dan de andere AVI-niveaus. Door een nieuw plan van aanpak zitten de prestaties nu op een score van 81%. De netwerkbijskomsten met andere scholen, die horen bij het leesverbeterplan zijn heel nuttig.”

Het effect van de vernieuwing van het onderwijs voor de leerlingen is dat er nu een vaste structuur is, waardoor zij goed weten wat er gaat gebeuren en wat van hen wordt verwacht. Zij krijgen meer eigen verantwoordelijkheid en moeilijkere teksten. Ze mogen de teksten vaker oefenen en merken zelf dat zij door de herhaling beter gaan presteren. De nieuwe aanpak heeft een omslag gevraagd van de leerkrachten, maar: “Er worden op alle niveaus resultaten geboekt, zowel op kindniveau, als op groepsniveau als op schoolniveau,” aldus de schoolleider.

De rol van het bestuur en het bovenschools management

Het schoolbestuur van school 4 houdt goed bij wat er op de scholen gebeurt, en speelt een belangrijke rol in de faciliteiten om het verbeteringstraject mogelijk te maken. Doel is dat 80% van de tijd van de schoolleider besteed wordt aan onderwijskundig leiderschap, en om die reden hebben alle Enschedese scholen twee dagen in de week administratieve ondersteuning gekregen. Verder is er overleg over WSNS, en zijn studiedagen voor schoolleiders, Ib-ers en leerkrachten. “Als het bestuur niet voor al deze faciliteiten had gezorgd, waren de prestaties op de school nooit zover gekomen.”

6.3.5 School 5

School 5 is een openbare SBO-school in Enschede. De school heeft ongeveer 200 leerlingen. Op deze school zitten alleen moeilijke lerende kinderen met een IQ tussen de 60 en 80. Er zijn verspreid over het schooljaar vier instroomdata. De instroom is het grootst in de groepen 3, 4 en 5, vooral in januari. In groep 3 zitten nu 12 kinderen, het maximum is 15 á 16 leerlingen, daarna wordt een nieuwe groep geformeerd. Er is een JORIK-groep ((J)Ongere RIsico Kinderen 4 tot 6 – jarigen) van ongeveer tien leerlingen. Aan het einde van groep 2 worden deze leerlingen gescreend. De moeilijk lerende kinderen blijven op de school, de LOM-kinderen gaan naar de andere SBO-school in Enschede.

Leerstandaarden en streefdoelen

Het streven is een gemiddeld leerrendement van 50% te bereiken voor alle leerlingen in alle vakken. Het doel is dat de leerlingen, wanneer zij de school verlaten, minimaal het niveau van eind groep 5 van het regulier basisonderwijs hebben bereikt. De school telt zowel (enkele) rugzakleerlingen met een IQ van 60 of lager, als leerlingen met een IQ van 80 of hoger. De kinderen worden per vak geplaatst in niveaugroepen: onder, op of boven niveau. De school probeert hen op zo hoog mogelijk niveau te brengen. Voor de kinderen met een IQ van 75 of hoger, streeft de school naar een leerrendement van 75%.

De doelen voor het taalonderwijs in de kleutergroepen zijn gebaseerd op het ontwikkelen van de ontluikende geletterdheid. Daarin spelen fonemisch bewustzijn, pictolezen en het uitbreiden van de woordenschat een belangrijke rol. De leerkrachten werken met de tussendoelen beginnende geletterdheid, waarvan steeds tussentijds wordt nagegaan of de vooraf bepaalde leerstandaard is behaald. Voor het rekenonderwijs aan de kleuters werkt men aan tussendoelen voor ontluikende gecijferdheid.

Voor technisch lezen gebruikt de school de normen uit het SBO-protocol Leesproblemen en Dyslexie, met als uitgangspunt dat elk kind per schooljaar twee AVI-niveaus zou moeten (kunnen) stijgen, ongeacht het startniveau bij binnenkomst: “Als een kind daarvoor meer tijd nodig heeft dan gepland, dan wordt er extra tijd ingestopt.” In groep 3 wordt gestart met Veilig Leren Lezen. Waar in het verleden drie kernen per jaar werden behandeld, bleken er vorig jaar 7 á 8 haalbaar. Het streven is dit jaar om er zelfs negen te behandelen. Het streefdoel voor de leerlingen aan het einde van groep drie is AVI-1.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Gegevens over taal- en rekenprestaties

De school werkt met de toetsen uit het Cito-leerlingvolgsysteem en met Eduscoop. De Cito-toetsen worden afgenomen in januari en juni. Daarnaast worden ook de DMT en de AVI-toetsen gebruikt, en worden er methodegebonden toetsen afgenomen: de school werkt met themaleren en na een blok van vijf weken moeten alle kinderen de vragen van het thema af hebben. De leerkrachten observeren de kinderen regelmatig. Bovendien is men bezig met het aanleggen van een portfolio voor de kennisgebieden waarin verschillende werkjes van de kinderen verzameld worden.

Kinderen die voor technisch lezen onder de norm van het SBO presteren, gaan werken met de Ralfi-methode. In groep 3 en 4 krijgen deze leerlingen dubbele leestijd waardoor zij sneller vooruitgaan in hun technische leesvaardigheid, en leerlingen uit groep 8 gaan met hen tutor-lezen. Als dat lukt, heeft het meteen effect op de prestaties in de andere vakken zoals rekenen en spelling. Bij deze methode worden de leerlingen om de tien weken getoetst, dus vier keer per schooljaar getoetst.

In groep 8 doen de leerlingen een schoolverlaterstoets de NIO of de NDT²⁴, omdat voor sommige leerlingen de NIO-toets te moeilijk is.

Analyse en bespreking van de resultaten

Het afnemen van de toetsen gebeurt door de leerkrachten, de ib-ers verzamelen de gegevens en plaatsen ze in een 'datamuur' per jaargroep. De analyse gebeurt met name door de ib-ers en de schoolleider, die voor het vergelijken met de landelijke gemiddelden kijken naar de SBO-normeringen.

In de groepsbesprekingen probeert men bij de leerkracht na te gaan waarom bepaalde resultaten niet zijn gehaald, of waarom sprake is van stagnatie van de ontwikkeling van bepaalde kinderen. De leerkrachten zijn vervolgens degenen die het groepshandelsplan opstellen voor rekenen, technisch lezen en spellen. In dat plan staat welke kennis de leerlingen moeten verwerven. Ook is de leerstofplanning uitgewerkt in een schema waarin de tijd, wat gedaan moet worden (bijvoorbeeld de te halen kernen) en de toetsmomenten zijn opgenomen.

²⁴ NDT= Nederlandse Differentiatie Testserie. De NDT is een intelligentietest voor leerlingen van 11 tot 15 jaar die in aanmerking komen voor plaatsing binnen het praktijkonderwijs of vmbo al dan niet met leerwegondersteuning.

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

De ib-er bekijkt de scores van de individuele leerlingen. Die toetscore bepaalt de instructiegroep waarin de leerling komt: op, boven of onder niveau. De IQ-verschillen tussen leerlingen kunnen, met name in de bovenbouw, erg groot zijn. Voor rekenen wordt daarom groepsdoorbrekend, per leerjaar, op niveau gewerkt, dat gebeurt ook bij het AVI-lezen en lezen met de Ralfi-methode. De ib-er gaat bovendien na hoe de ontwikkeling van de individuele leerlingen verloopt en geeft aan wie een langzame, snelle of gemiddelde 'groei' doormaakt. Voor kinderen van wie de groei achterblijft, voeren de ib-er en de groepsleerkracht een foutenanalyse uit, zodat duidelijk wordt waar precies de problemen zitten. De rugzakleerlingen zijn de echte zorgkinderen. De resultaten van deze kinderen worden apart besproken. Veelal wordt voor deze kinderen extra begeleiding geregeld en is er ambulante hulp.

Terugkoppeling en onderwijsverbeteringen

Twee keer per jaar wordt op schoolniveau met het hele team gekeken naar de prestaties van alle kinderen in dat schooljaar. De ib-er en de schoolleider presenteren de gegevens en bespreken per vakgebied welke conclusies er te trekken zijn uit de resultaten. Het team zoekt samen naar oplossingen en verbeterpunten voor resultaten die zijn tegengevallen. Er worden geen vergelijkingen gemaakt met landelijk gemiddelden, omdat alle leerlingen op de school moeilijk lerend zijn. Alle leerlingen op de school zijn zorgleerlingen. De vorderingen en prestaties worden vergeleken op leerlingenniveau, groepsniveau en schooljaarniveau. De ene keer ligt het accent bij de leerlingen die onder het streefniveau presteren, de andere keer bij de leerlingen die juist boven het niveau presteren.

De leerkrachten zelf interpreteren de gegevens van de leerlingen in de eigen groep en overleggen, als dat nodig is, met de ib-er en met leerkrachten van vergelijkbare groepen. Daarbij gaat het over de vraag hoe zij het lesprogramma kunnen aanpassen, waar de mogelijke problemen zitten (op school, thuis, de puberteit), of er sprake is van gedragsproblemen. Vervolgens worden de doelen bijgesteld, en wordt voor de leerling eventueel extra leerstof aangeboden en tijd ingeruimd voor bijvoorbeeld het lezen.

Aan de ouders doet de school verslag tijdens de ouderavonden. Ook gaan de leerkrachten minimaal een keer per jaar op huisbezoek om met de ouders te praten over het functioneren van de kinderen. Verder komen de uitkomsten van de metingen aan de orde tijdens studiedagen, in de jaarverslagen en de schoolbrochure.

Opbrengricht werken; over de waarde van meetgestuurd onderwijs

De school heeft drie speerpunten geformuleerd waarop men de prestaties van de leerlingen wil verbeteren: technisch lezen, begrijpend lezen en woordenschat. Ook over spelling bestaan zorgen, dat wordt een volgend speerpunt. De leerkracht staat daarin centraal. Om die verbeteringen te bereiken is er tijdens de teamvergaderingen is er veel aandacht voor effectieve instructie. Alle teamleden volgen nascholing op dat gebied. Wanneer leerkrachten aangeven dat zij een probleem hebben, gaan de ib-ers kijken in de klas. Zij maken soms ook video-opnames: “Dat is heel leerzaam voor de leerkrachten, alleen al het zien van de opnames is soms voldoende” vindt een van de ib-ers. De schoolleider voert ook klassenconsultaties in controlerende zin uit. Hij beoordeelt de leerkrachten op hun competenties en op basis van aandachtspunten uit hun eigen ontwikkelplan. Na verloop van tijd volgt dan een beoordelingsgesprek, maar dat is tot nu toe nog niet gebeurd.

Effecten

(a) Effecten op taal- en rekenprestaties

“De resultaten bij het technisch lezen zijn opvallend positief en ook de andere vakken liften mee op de vooruitgang in de technische leesvaardigheid”, aldus de directeur. In het schooljaar 2007-2008 deden zo’n 50 van de 150 leerlingen uit groep 5 tot en met 8 mee met het RALFI-lezen. Bij 86% van de leerlingen in deze groep werd de doelstelling ‘minimaal twee AVI-instructieniveaus hoger’ bereikt. In de periode augustus tot oktober 2008 deden 38 leerlingen mee, en werd voor 44% van hen het doel ‘minimaal een AVI instructieniveau hoger’ ook daadwerkelijk bereikt. Er is tussen januari 2006 en juni 2008 sprake van een groeiend percentage leerlingen dat de norm haalt voor technisch lezen (van 51% naar 79%), voor begrijpend lezen (van 26% naar 55%) en voor rekenen (van 32% naar 54%), zoals een staafdiagram in het jaarverslag van de school laat zien.

(b) Overige effecten

“Voorheen werkten we kindvolgend, nu werken we kinduitdagend. Sinds deze aanpak zijn de kinderen vooral met lezen enorm vooruitgegaan. De resultaten geven aan dat we de norm nog kunnen verhogen. We zijn bij Veilig Leren Lezen van drie naar negen kernen per jaar gegaan. Als ons vier jaar geleden was gezegd dat dit mogelijk was, hadden we het niet geloofd,” geven de teamleden aan. “Door de taalpilots blijkt dat we in het verleden, zelfs bij de kleuters, de kinderen vaak hebben onderschat. We namen ze te veel in bescherming. We weten nu dat

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

de kinderen meer aankunnen dan we vaak denken. In groep 2 leerden we in het verleden vijf letters per jaar aan, nu streven we naar 20 letters. We hebben ons aanbod aan de kinderen vergroot en verruimd. Uiteraard kunnen niet alle kinderen het aan, maar we bieden in eerste instantie alles aan de hele groep aan. Voor wie het te hoog gegrepen is wordt (meer) herhaling ingebouwd. We gaan er van uit dat elk kind moet kunnen leren lezen en het blijkt haalbaar.”

Het meetgestuurd onderwijs in combinatie met het effectieve instructiemodel biedt de leerkrachten meer duidelijkheid: “Je weet beter welke instructie je moet geven aan welk kind. Je differentieert meer. Je weet beter van elkaar als leerkrachten waar je mee bezig bent. Wel is het nodig goed in de gaten te houden dat kinderen zich niet overvraagd gaan voelen. Het gevoel van veiligheid is van groot belang.” De nieuwe manier van werken heeft een omschakeling betekend voor het onderwijs, maar de leerkrachten geven aan dat zij bewuster bezig zijn met hun onderwijs: “Het plannen van de leerstof, het werken met een handelingsplan geeft houvast. Je weet nu veel beter waar je aan werkt. Je plant, het werkt en dat geeft voldoening. We kijken nu in het SBO veel meer naar de leerresultaten van de kinderen dan voorheen. We zien dat het dus anders kan. Ook de kinderen komen veel beter uit de verf: ze kunnen vertellen wat ze van school vinden, wat ze er leren.” De school gebruikt voor het volgen van de sociaal-emotionele ontwikkeling de SCOL, ook daarop wordt zichtbaar dat de leerlingen zich gunstiger ontwikkelen van voorheen. Met name op aspecten als aardig doen, opkomen voor jezelf en omgaan met ruzie wordt hoger gescoord.

De rol van het schoolbestuur en bovenschools management

Het bestuur ondersteunt met name het onderwijskundig leiderschap van de schoolleider. De schoolleider stuurt het team aan in de gezamenlijke visie op onderwijs, en er is -veel meer dan vroeger- sprake van een gezamenlijke verantwoordelijkheid voor de inhoud van het onderwijs en het verbeteren van het primaire proces.

6.4 Knelpunten en succesfactoren

In de interviews op de verschillende scholen komt een aantal knelpunten en risico's naar voren die telkens lijken terug te keren, in deze paragraaf zijn de belang-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

rijkste daarvan bij elkaar genomen. Ook de ervaren winstpunten, en de factoren die daarin een rol hebben gespeeld worden door de respondenten 'gedeeld', en komen hieronder aan de orde.

6.4.1 Knelpunten en risico's

In de 14 gesprekken die wij op de vijf scholen hebben gehouden, brachten de leerkrachten, schoolleiders, ib-ers en taalcoördinatoren enkele knelpunten en risico's van meetgestuurd onderwijs naar voren. Voor al die punten geldt dat ze door de betrokkenen over de verschillende scholen heen worden gedeeld.

Knelpunten

Bij de knelpunten gaat het om problemen die zich daadwerkelijk op de scholen hebben voorgedaan, of die nog steeds spelen in de ervaring van de respondenten. Het eerste, en meest genoemde, knelpunt (7g/4s)²⁵ is dat het afnemen en goed registreren van alle toetsen, en het vervolgens werken op verschillende niveaus vraagt om meer handen in de klas dan er nu beschikbaar zijn. Voor het speciaal basisonderwijs komt daarbij dat er sprake is van verdichting van de problematiek: er komen meer kinderen met een IQ onder de 60 naar de scholen, de kinderen stromen vaak te laat in (pas in groep 7 of 8), en de groepen worden groter waardoor de kwaliteit onder druk komt te staan.

Het tweede knelpunt (5g/3s) dat de respondenten naar voren brengen is de neiging om teveel te gaan toetsen, registreren en administreren, zeker bij deelname aan projecten van instanties die elk hun eigen gegevensoverzichten vragen.

Het derde knelpunt (3g/2s) is het gebrek aan aansluiting tussen enerzijds de toetsen uit het Cito-leerlingvolgsysteem en anderzijds de leerstof zoals die in de methodes aan de orde komt. Wanneer bepaalde onderdelen in de methode nog niet zijn behandeld of op een andere manier worden bevraagd dan in de toets, geven de toetsscores geen adequaat beeld van het niveau van de leerlingen. Er is dan het risico dat je als leerkracht de leerlingen gaat trainen voor de toets.

Risico's

Behalve deze knelpunten signaleren de geïnterviewden ook twee risico's van meetgestuurd onderwijs. Risico's zijn –anders dan knelpunten– geen actuele pro-

²⁵ Het knelpunt is genoemd in zeven gesprekken (7g) verspreid over vier scholen (4s).

blemen die op dit moment spelen op de scholen, maar problemen waarvan de respondenten onderkennen dat ze in de toekomst, of in ongunstige omstandigheden zouden kunnen optreden.

Ten eerste is er een risico dat er binnen het team, of bij de directie een afrekencultuur ontstaat (4g/4s) waarin leerkrachten onzeker worden als hun leerlingen slecht presteren. De vier citaten geven aan hoe men over dit risico denkt, maar ook hoe de scholen het ondervangen:

- "Meetgestuurd werken vraagt een open teamcultuur, en ruimte voor reflectie op je eigen handelen. De leerprestaties moeten zonder 'veroordeeling' besproken (kunnen) worden, zonder vingerwijzingen naar collega's of directie."
- "Je moet realistische doelstellingen voor ogen blijven houden en een veilig klimaat hebben waarin je met elkaar bespreekt wat er wel of niet goed is gegaan, en wat beter kan."
- "Negatief voor leerkrachten kan zijn dat wanneer je een zware groep hebt, waar je veel energie in moet stoppen en je heel erg je best doet en vervolgens niet de gestelde 85% score haalt. Het kan heel teleurstellend werken als je ziet dat je in de toekomst met die groep ook nooit de 95% gaat halen. In dat soort situaties is het van belang dat er klassenconsultaties worden gedaan om te kijken op welke wijze de leerkracht geholpen kan worden. Niet om de leerkracht er op af te rekenen, maar om steun te bieden."
- "Er bestaat het risico dat leerkrachten het gevoel krijgen zich te moeten verantwoorden. Het slagen van meetgestuurd werken valt en staat met het vertrouwen dat leerkrachten hebben in de begeleiders en de schoolleider. De leerkracht is de belangrijkste persoon in de groep. De schoolleider moet zorgen voor de faciliteiten, de juiste ondersteuning, de goede materialen, het helpen zoeken naar oplossingen bij problemen, het bespreekbaar maken van resultaten en personeelsbeleid voeren etc."

Ten tweede bestaat de kans dat de school of de leerkracht teveel gaat leunen op de toetsuitslagen (4g/3s), en dat daarmee de eigen professionele blik op de ontwikkeling van de leerling of op het eigen handelen van de leerkracht in de groep naar de achtergrond verdwijnt. Eén van de respondenten verwoordt dat als volgt: "Het is belangrijk als leerkracht ook te (kunnen) blijven vertrouwen op de eigen observaties in je groep, en je eigen beoordelingen. Je moet durven han-

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

delen, ook zonder de toetsresultaten ben je professioneel genoeg om te signaleren wat nodig en goed is voor je leerlingen." Een andere leerkracht geeft aan: "Ontwikkeling gaat soms met sprongen, soms is een kind opeens ergens wel aan toe. Dat betekent dat het niet altijd nodig is heel sturend en gericht te werken, maar dat je zoekt naar evenwicht tussen 'niet frustreren', maar ook niet te lang wachten met het aanbieden van bepaalde leerstof."

6.4.2 Winstpunten en succesfactoren

Net als de knelpunten en risicofactoren, worden ook de winstpunten en succesfactoren breed gedeeld over de scholen heen. Hieronder eerst een overzicht en vervolgens een bespreking.

Winstpunten

Onder de winstpunten worden de diverse opbrengsten van meetgestuurd onderwijs beschreven, die breder zijn dan de beoogde verbetering van de taal- en rekenprestaties bij de leerlingen. Bijzonder is dat eigenlijk geen van de geïnterviewden die verbetering van de prestaties benoemt als opbrengst of winstpunt. Wellicht komt dat doordat daarover nog niet zo heel veel feitelijke informatie beschikbaar is op de scholen, anderzijds is het een verschijnsel dat zich vaker voordoet bij het bevragen van leerkrachten en schoolleiders over onderwijsvernieuwingen: zij beschrijven wat er in de praktijk van het eigen onderwijs en het handelen van de leerkracht is veranderd, en wat zij (subjectief) observeren bij hun leerlingen, en zijn veel minder gericht op de 'harde' gegevens over prestaties van die leerlingen.

Het belangrijkste winstpunt (8g/5s) van meetgestuurd onderwijs is dat het leidt tot gericht werken aan de leerstof, waarbij de leerkracht de methode als leidraad neemt, maar zelf keuzes maakt en bewust handelt op basis van inzicht in het niveau en de ontwikkelingslijnen van de leerlingen. Eén van de leerkrachten van de SBO-school omschrijft deze winst als volgt: "We weten nu beter wat we binnen een week aan de kinderen kunnen leren. Vroeger deden we wel twee weken over het aanleren van één letter, nu doen we dat in één dag. Als leerkracht kun je nu sneller je strategie aanpassen, dat is een voordeel. We hebben nu veel meer overzicht. Onze leerstofplanning is duidelijk en we weten in welk tempo we moeten werken. We overwegen onze keuzes wat betreft het aanbod aan de kinderen. Onze kinderen doen lang over werkjes met knippen en plakken. Die oefeningen vergen

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

te veel tijd en zijn niet direct van belang voor de taak, daarom laten we ze weg. Daardoor werken we efficiënter.” Bijkomend voordeel is dat de resultaten helderder opleveren in de contacten met ouders: “Je kunt aan de hand van de resultaten voor de ouders duidelijker, en al vroeger in de schoolloopbaan, aangeven wat goed gaat en welke dingen problemen geven. Daardoor kun je de informatie beter overbrengen.” Enkele leerkrachten maken wel kanttekeningen bij dit belangrijke winstpunt. De eerste geeft aan dat het bereiken van doelstellingen, niet de aantrekkelijkheid van het onderwijs mag gaan overvleugelen: “Het is belangrijk dat de kost die je hen voorschotelt natuurlijk wel zo plezierig en leuk mogelijk is. Daarvoor is het nodig je in te leven in de kinderen en hen niet het idee te geven dat ze bezig zijn met een leerproces en met verplichte kennis.” Twee andere leerkrachten geven aan dat zij graag meer zouden willen leren over het analyseren van de toetsresultaten en nuancering daarvan. “Ik zou graag beter inzicht willen krijgen, in waar nu precies de problemen met de leerstof zitten, hoeveel stappen je terug moet doen, en met welke materialen je dan het best aan de slag kunt gaan.” Het tweede winstpunt (6g/4s) dat uit de gesprekken naar voren komt is de ervaring van de teamleden dat zij beter kunnen terugvallen op hun collega's, en dat er binnen het team meer sprake is van het geven van feedback aan elkaar, en van het uitwisselen over het onderwijs en het kritisch durven kijken naar de eigen rol. Voorwaarde daarvoor is natuurlijk wel dat er sprake is van een veilige sfeer en goede verhoudingen binnen het team en tussen team en directie. Ten derde (3g/3s) benoemen de teamleden als winstpunt van meetgestuurd onderwijs dat het motiverend is voor zowel leerkrachten als leerlingen wanneer de resultaten goed zijn. Eén van de SBO-leerkrachten geeft daarbij het volgende voorbeeld: “Sommige kinderen zijn erg faalangstig. Zij hebben op voorgaande scholen alleen maar teleurstellingen te verwerken gekregen omdat ze dingen niet konden. Ze zijn heel trots als ze bijvoorbeeld heel lang in AVI-1 zijn blijven hangen en nu vooruitgaan.”

Succesfactoren

Tot slot zijn er twee factoren die bepalend zijn voor het succes van meetgestuurd onderwijs. Deze succesfactoren kunnen opgevat worden als de belangrijkste invoeringscondities, volgens de respondenten.

Ten eerste is dat een goede directie (3g/2s) die niet van bovenaf bepaalt wat de leerkrachten moeten doen, maar die blijft luisteren naar ‘de werkvloer’ en zorgt

Opbrengricht werken; over de waarde van meetgestuurd onderwijs

voor borging van goede ontwikkelingen. De kern van een goed functionerende school is dat alle teamleden belangstelling hebben voor elkaar en voor het eigen onderwijs en dat zij meedenken over mogelijke verbetering daarvan.

De tweede succesfactor (2g/2s) vloeit daaruit voort, en houdt in dat bij leerkrachten het gevoel van verantwoordelijkheid voor hun eigen onderwijs en voor de resultaten van hun leerlingen wordt ontwikkeld. “Toetsresultaten is geen ‘vies woord’ meer in school. De leerkracht is eigenaar van die resultaten en maakt op basis daarvan de eigen plannen. Dat is inmiddels niet meer alleen een zaak van de ib-er. In de school is het gelukt om leerkrachten in het ontwikkelen van groeps-handelingsplannen ook zelf handelingsbekwaam te maken. Dat betekent dat een omslag is gemaakt in de functie van de ib-ers, die vroeger juist degenen waren die de toetsen afnamen, de analyses uitvoerden en de handelingsplannen ontwikkelden. Er is geen Remedial Teaching meer, de leerkracht is de verantwoordelijke.”

6.5 Tussentijdse conclusies

Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?

De scholen maken alle vijf gebruik van het Cito-LVS, en nemen alle toetsen die daarin zijn opgenomen af volgens de toetskalender. Daarnaast wordt gewerkt met de resultaten op methodegebonden toetsen voor taal en rekenen. Een enkele keer vinden er diagnostische gesprekken plaats of observaties door de leerkracht, maar het zwaartepunt ligt op de toetsresultaten.

In hoeverre en hoe worden deze gegevens gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?

De toetsgegevens worden altijd gebruikt voor diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen, zeker voor de leerlingen van wie de prestaties op D- of E-niveau liggen. Voor die leerlingen wordt over het algemeen een handelingsplan ontwikkeld, en na enige tijd opnieuw gekeken naar hun prestaties. Ook vindt evaluatie plaats van het onderwijs voor de hele groep of subgroepen:

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

soms worden groepshandelingsplannen opgesteld wanneer een deel van de leerlingen bepaalde onderdelen van de leerstof nog niet blijkt te beheersen. In uiterste consequentie worden bij achterblijvende leerresultaten op schoolniveau ook verbeteringstrajecten ingezet: ingrijpende vernieuwing van het lees- of rekenonderwijs, invoering van een methodiek voor woordenschatontwikkeling, integratie van het taal- en zaakvakonderwijs etc. Alle vijf de taalpilotscholen hanteren daarbij vooraf gestelde leerdoelen, uitgangspunt zijn de kerndoelen zoals die in de methodes aan de orde komen, daarnaast wordt meestal vastgelegd voor welk percentage van de leerlingen minimaal een voldoende niveau (een A-, B- of C-score) op de taal- en/of rekentoetsen gehaald moet worden. Ongeveer 10 tot 15% van de leerlingen op D- of E-niveau vinden de scholen meestal acceptabel, alleen voor technisch lezen worden strengere normen gesteld (95 to 100%).

Opvallend is overigens dat geen van de –drie- scholen die bij de leerlingen de Cito-Eindtoets afneemt, de gemiddelde schoolscore op deze toets noemt als instrument op basis waarvan zij hun leerstandaard bepalen. Het lijkt erop dat de scholen de resultaten op deze toets alleen benutten als eindmeting ten behoeve van het schooladvies voor de leerling, niet als longitudinale meting van het eigen onderwijsniveau.

Welke handelingen/stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom?

Er worden doelen bepaald, zowel op schoolniveau, als op groepsniveau, als voor individuele leerlingen. Via bestudering van de toetsresultaten, eveneens op alle drie de niveaus, wordt geëvalueerd in hoeverre de gestelde doelen zijn bereikt. Wanneer dat niet het geval is, overlegt men over de achtergronden daarvan, en over mogelijkheden ter verbetering van het onderwijs. Wanneer het om individuele leerlingen gaat, is er meestal sprake van overleg tussen de groepsleerkracht en de leraar over toetsresultaten, ontwikkelingsverloop en het handelingsplan. De groepsresultaten op schoolniveau worden over het algemeen besproken in een teamoverleg. Een enkele keer vormen ze ook (al) een aandachtspunt in het functioneringsoverleg. Vaste aandachtspunten zijn: wat gaat goed, wat nog niet, hoe komt dat en hoe kan het beter? Zowel op individueel niveau, als op groeps- en schoolniveau worden vervolgens nieuwe doelen gesteld, en nieuwe evaluatiemomenten bepaald.

Opbrengstgericht werken; over de waarde van meetgestuurd onderwijs

Bestaat er draagvlak bij leerkrachten en schoolleiders voor meetgestuurd onderwijs?

Uit de interviews blijkt dat er zowel bij de schoolleiders als bij de leerkrachten van de vijf Taalpilotscholen zeker draagvlak bestaat voor meetgestuurd onderwijs. Men is met name enthousiast over het feit dat meetgestuurd werken leidt tot gericht en efficiënter onderwijs, meer bewuste keuzes en meer greep op de leerstof. Er vindt meer uitwisseling plaats binnen het team, leerkrachten hebben meer zicht op de ontwikkelingscurves van hun leerlingen, en dat betekent dat er in grotere mate sprake is van doorgaande leerlijnen.

Wat zijn risico- en succesfactoren bij meetgestuurd onderwijs? Zijn er eventueel onbedoelde effecten?

De belangrijkste risico's van meetgestuurd onderwijs vinden de scholen dat er een te sterke nadruk op de toetsresultaten zou kunnen ontstaan. Wanneer alleen de resultaten nog tellen, kan dat uiteindelijk ook ten koste gaan van de eigen professionaliteit en inzichten van de leerkracht. Een tweede risico is het ontstaan van een 'afrekencultuur' waarin de leerkracht verantwoordelijk wordt gehouden voor achterblijvende prestaties van de leerlingen, en de oorzaken te weinig aandacht krijgen.

De succesfactoren zijn eigenlijk de tegenpolen van de genoemde risico's: de gerichte aandacht voor prestaties, het nauwgezet volgen van de ontwikkeling van de leerlingen en het daarop afstemmen van het onderwijs, maken dat de leerkrachten meer greep hebben op hun onderwijs en bewustere keuzes maken in de leerstof en methodelessen. Een veilige, open, collegiale sfeer in het team en goede relaties tussen team en directie vormen de tweede belangrijke succesfactor. In die situatie voelen de leerkrachten zich inderdaad verantwoordelijk voor de leerprestaties, en eigenaar van het onderwijs in hun groep, maar voelen zij zich ook in voldoende mate gesteund om die verantwoordelijkheid, gezamenlijk, te dragen. De constatering dat de leerprestaties inderdaad vooruitgaan werkt daarbij voor zowel de leerkrachten als de leerlingen zelf extra motiverend.

Meetgestuurd onderwijs op vijf Taalbeleid-pilotscholen

Wat hebben schoolleiders en leerkrachten nodig om meetgestuurd onderwijs te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

Misschien is in de gesprekken onvoldoende expliciet gevraagd wat leerkrachten of schoolleiders nodig hebben voor de uitvoering van meetgestuurd onderwijs, toch is het eerste dat opvalt dat slechts een enkele respondent opmerkt dat zij behoefte heeft aan scholing in het leren omgaan met data. (Met name om meer genuanceerde, diagnostische gegevens te verkrijgen dan alleen een indicatie voor het niveau waarop een leerling presteert). Voor het overige bieden de knelpunten en succesfactoren die genoemd zijn, zeker een aantal aanknopingspunten voor wat nodig is om meetgestuurd onderwijs met succes te kunnen uitvoeren.

De eerste conditie is dan de kwaliteit van de schoolleiding én de mogelijkheid voor die schoolleiding om zich daadwerkelijk bezig te houden met de onderwijskwaliteit. De administratieve ondersteuning voor de Enschedese scholen, gefaciliteerd door het schoolbestuur, is daarvan een goed voorbeeld. De tweede conditie is dat er sprake moet zijn van een goede onderlinge sfeer in het team, waarin men zich gezamenlijk verantwoordelijkheid voelt voor het onderwijs en voor de resultaten die bij de leerlingen worden bereikt. Ook in die zin is voor de invoering van meetgestuurd onderwijs ondersteuning van het onderwijskundig leiderschap van belang: de schoolleider moet in staat zijn gezamenlijke verantwoordelijkheid en eigenaarschap van de visie bij de teamleden tot stand te brengen, ofwel 'sturing te geven aan een lerende organisatie' waar in een open en veilig klimaat feedback kan worden gegeven aan collega's.

Verder komt uit de interviews naar voren dat het afnemen van de toetsen én het werken op verschillende niveaus veel tijd kost. Om die reden is ook 'meer handen in de klas' een aanbeveling voor het uitvoeren van meetgestuurd onderwijs. En tot slot vindt men een goede afstemming tussen de methodes enerzijds en de inhoud van de toetsen anderzijds van belang, evenals stroomlijning in de af te nemen toetsen en de registratie daarvan.

7 Conclusies en discussie

7.1 Doel van het onderzoek

Het uitgangspunt voor dit onderzoek was het begrip data driven teaching. De opdracht was om na te gaan wat dit begrip inhoudt, welke stappen in data driven teaching kunnen worden onderscheiden, of en hoe scholen voor primair onderwijs deze manier van onderwijzen toepassen, hoe het staat met het draagvlak daarvoor en wat de condities zijn voor de uitvoering ervan. Om op deze vragen antwoord te geven hebben we een aantal deelonderzoeken uitgevoerd:

- een literatuurstudie, tevens begripsverkenning
- interviews met praktijkdeskundigen: inspectie, schoolbegeleiders, medewerkers van het Cito
- een korte inventarisatie van de aandacht voor het onderwerp op de pabo's
- een verkenning van de praktijk in het primair onderwijs, via een telefonische enquête
- een verdere verkenning van die praktijk via casestudies op enkele scholen die deelnemen aan de pilots taalbeleid
- analyse van de ervaringen van leerkrachten met proefreferentieniveaus, opgesteld door het Cito

In dit hoofdstuk brengen we de resultaten van die deelonderzoeken bij elkaar en beantwoorden we de onderzoeksvragen. Per onderzoeksvraag vatten we samen wat het beeld is dat uit de deelonderzoeken naar voren komt. Daarbij grijpen we terug op het kader voor het onderzoek, zoals geschetst in hoofdstuk 1. Aan het eind van dit hoofdstuk evalueren we de uitkomsten in een discussieparagraaf, die we afsluiten met enkele aanbevelingen.

7.2 De begripsanalyse

De term *data driven teaching* is in de onderwijskundige literatuur niet erg gangbaar, zo bleek uit het literatuuronderzoek. Wel zijn er andere termen in omloop

Opbrengstgericht werken

voor de praktijk die er mee bedoeld wordt: diagnostic teaching, adaptief onderwijs, data-based decision making. Die termen verwijzen meestal naar het onderwijs in de klas ('teaching'). In de Nederlandse context wordt het begrip echter ruimer opgevat: het gaat om het verzamelen en gebruiken van gegevens binnen de hele school, om een 'onderzoekende' werkwijze waarbij uitkomsten van evaluaties of metingen worden ingezet. In het landelijk beleid is 'data driven teaching', in overeenstemming met deze bredere bedoeling, inmiddels vertaald als 'opbrengstgericht onderwijs'. In dit onderzoek hebben we echter, op basis van de begripsanalyse, de voorkeur gegeven aan een andere aanduiding, namelijk *meetgestuurd onderwijs*. Dat hebben we omschreven als onderwijs waarbij betrokkenen (scholen, leraren, leerlingen) zich in hun taakuitvoering laten sturen door de uitkomsten van metingen. Meetgestuurd onderwijs ligt qua betekenis dichter bij data driven teaching dan opbrengstgericht onderwijs. Beide termen hebben namelijk gemeen dat ze de nadruk leggen op het gebruik van data of de uitkomsten van metingen. Opbrengstgericht werken heeft die connotatie niet, deze term mist een verwijzing naar systematische gegevensterugkoppeling of feedback.

Een belangrijk onderscheid is op welk niveau en voor welk doel meetgestuurd onderwijs wordt ingezet. Leerkrachten gebruiken allerlei bronnen van feedback over de prestaties van hun leerlingen: observaties van het leerproces in de klas, schriftelijk werk van de leerlingen, methodegebonden toetsen. Die kunnen op klasniveau allemaal worden ingezet voor meetgestuurd onderwijs. Methode-onafhankelijke, gestandaardiseerde toetsen worden ook door leerkrachten gebruikt, maar nemen in het geheel van data maar een bescheiden plaats in, omdat ze niet frequent (kunnen) worden afgenomen. Op schoolniveau zijn echter juist de resultaten van deze toetsen interessant, omdat daarmee vergelijkingen kunnen worden gemaakt: tussen klassen, in de tijd, met andere scholen. Dit wordt ook wel aangeduid met het onderscheid tussen formatieve evaluatie (gegevens gebruiken om voortgang te bewaken en eventueel de werkwijze bij te stellen, evaluatie lopende het proces) versus summatieve evaluatie (gegevens gebruiken om een stand van zaken op te maken en/of te bepalen of gestelde doelen zijn gehaald, evaluatie achteraf).

Verder is het zo dat meetgestuurd onderwijs niet alleen betrekking hoeft te hebben op het gebruik van gegevens over leervorderingen. Het kan ook gaan om me-

tingen van andere zaken (zoals tevredenheid, welbevinden, pedagogische en didactische kwaliteit van het onderwijs). In dit rapport ligt het accent echter wel op het gebruik van gegevens over leervorderingen, omdat de context voor het onderzoek het streven naar verbetering van de taal- en rekenvaardigheden van leerlingen in het primair onderwijs is, zoals verwoord in de Kwaliteitsagenda.

7.3 Stand van de kennis in de literatuur

De eerste vraag in dit onderzoek luidde:

Uit welke handelingen/stappen bestaat meetgestuurd onderwijs? Wat is volgens bestaand onderzoek de beste werkwijze voor meetgestuurd onderwijs?

De literatuurstudie heeft hierover het volgende beeld opgeleverd.

Uit welke stappen bestaat meetgestuurd onderwijs?

Het gebruik van vorderingengegevens is onlosmakelijk verbonden met de evaluatieve cyclus. Deze cyclus bestaat uit vijf stappen: het vastleggen van doelen en standaarden, het verzamelen van informatie, het registreren ervan, het interpreteren en het nemen van beslissingen. Het gaat om een eenvoudige en voor zichzelf sprekende indeling in stappen, al langer bekend in de onderwijswereld, die ook van toepassing is voor meetgestuurd onderwijs. Idealiter begint de cyclus met het stellen van doelen en criteria (standaarden). Het is echter ook mogelijk te beginnen met het verzamelen van informatie nog zonder een vooraf geformuleerd doel. Bijvoorbeeld om een stand van zaken op te maken of een diagnose te stellen.

Voor elke stap zijn specifieke aandachtspunten te formuleren:

- Voor het *vastleggen van doelen en standaarden* geldt dat dit bij voorkeur moet plaatsvinden in meetbare termen en voldoende specifiek. Hierbij kan differentiatie plaatsvinden, bijvoorbeeld in minimumdoelen en streefdoelen. De referentieniveaus taal en rekenen die onlangs zijn voorgesteld door de Expertgroep Doorlopende leerlijnen Taal en Rekenen kennen een dergelijke differentiatie.
- Het *verzamelen van informatie* kan op verschillende manieren plaatsvinden en met verschillende instrumenten. Een combinatie van instrumenten verdient vaak de voorkeur omdat daarmee de beperkingen van afzonderlijke instrumenten verminderd kunnen worden.

Opbrengstgericht werken

- *Registreren* kan eveneens op verschillende manieren, maar invoer van de gegevens in de computer is het meest gewenst, omdat dat betere analysemogelijkheden biedt. Een goed systeem op papier kan echter ook functioneel zijn, als dat beter past bij voorkeuren van leerkrachten, het soort gegevens of mogelijkheden van de school. Belangrijk is dat er goede toegang tot de gegevens is voor zowel leerkrachten als schoolleiding.
- Voor de stap *interpreteren* kan een onderscheid gemaakt worden tussen leerstofgerichte interpretatie (welk deel van de leerstof wordt beheerst?) en normgerichte interpretatie (wat is het prestatieniveau vergeleken met een norm?). Interpretatie wordt eenvoudiger als er vooraf gestelde doelen zijn (zie stap 1).
- Voor het *nemen van beslissingen* geldt dat dit niet altijd een beslissing in de richting van verbetering of verandering hoeft te zijn. Een beslissing kan ook zijn dat de geconstateerde situatie goed is zoals het is.

Is er een 'beste werkwijze'?

Meetgestuurd onderwijs is een wijze van feedback geven aan leerkrachten en/of leerlingen. Er is empirisch bewijs dat feedback geven prestaties verhoogt en dus gaat het in beginsel om een effectieve werkwijze. Voorwaarden zijn wel dat het om valide metingen en om realistische doelen gaat en dat er bereidheid is om conclusies te verbinden aan de uitkomst van de metingen.

De onderzoeksliteratuur geeft weinig rechtstreekse aanwijzingen voor een 'beste werkwijze'. Wel zijn er verschillende valkuilen of risico's uit af te leiden:

1° Gebruik van toetsresultaten moet in functie staan van het geven van aanwijzingen (aan leerkrachten, leerlingen) over hoe iets beter kan. Het moet dus taakgericht zijn. In de praktijk is het echter nogal eens persoonsgericht en worden er uitspraken gedaan in termen van goed/slecht, voldoende/onvoldoende. Dit kan competitiegedrag in de hand werken en/of leiden tot een overmatige nadruk op het voldoen aan verantwoordingseisen.

2° De neiging bestaat om bij analyse van toetsresultaten vooral te kijken naar leerlingen die beneden een norm presteren. Zij krijgen vervolgens extra aandacht en stimulans. Dit kan leiden tot onderstimulering van de betere leerlingen, die de doelen/de norm wel halen. Zij worden niet uitgedaagd tot verbetering van hun prestaties.

3° Invoeren van goede toets- en feedbackprocedures kost tijd, vooral als er oude ineffectieve patronen moeten worden doorbroken.

4° Summatief toetsen op schoolniveau (om te controleren of scholen voldoende kwaliteit leveren) kan leiden tot strategisch gedrag, zoals zwakke leerlingen buiten de toetsing houden en *teaching to the test*. Dat laatste kan een verarming van het curriculum in de hand werken (leerdoelen die niet getoetst worden krijgen dan minder aandacht).

7.4 Beschikbare gegevens over leervorderingen

De tweede vraag in dit onderzoek luidde:

Welke objectieve gegevens over de ontwikkeling van taal- en rekenvaardigheden van leerlingen hebben scholen tot hun beschikking en van welke niet-objectieve gegevens maken leerkrachten verder gebruik?

Ook deze vraag kon al vanuit de literatuurstudie beantwoord worden. Daaruit blijkt dat vrijwel alle scholen/leraren in het basisonderwijs verschillende soorten gegevens verzamelen en gebruiken over leervorderingen. Naast observaties, mondelinge beurten, schriftelijk werk en methodegebonden toetsen gebruiken ze ook vrijwel allemaal een leerlingvolgsysteem met methodeonafhankelijke toetsen. Daardoor beschikken ze over regelmatige objectieve gegevens over het prestatieniveau van de leerlingen, in ieder geval voor de basisvaardigheden (lezen, rekenen, spelling). Het meest gebruikte leerlingvolgsysteem is het Cito-leerlingvolgsysteem. De verspreidingsgraad hiervan bedraagt rond de 90 procent. Daaruit mag worden afgeleid dat de overgrote meerderheid van leraren gebruik maakt van de daarin opgenomen toetsen voor technisch en begrijpend lezen, spelling en rekenen/wiskunde en dat de meeste scholen zich houden aan het door het Cito aanbevolen afnameschema, dat grosso modo neerkomt op een halfjaarlijkse afname. De scores op methodeonafhankelijke toetsen worden in de regel centraal opgeslagen (op schoolniveau), de overige gegevens meestal niet.

De overige deelonderzoeken hebben dit beeld bevestigd en op onderdelen aangevuld. Zo meldt het Cito dat

meer dan 60 procent van de scholen ook het bij het leerlingvolgsysteem behorende computerprogramma in huis heeft, waarmee toetsgegevens goed geanalyseerd

Opbrengstgericht werken

kunnen worden. ‘In huis’ betekent niet automatisch ‘gebruiken’, maar het instrument is in ieder geval op een meerderheid van de scholen aanwezig. Ook traditionele vernieuwingscholen zoals montessorischolen gebruiken in meerderheid het leerlingvolgsysteem van het Cito. Maar deze scholen werken wel relatief vaak met andere systemen. De montessorivereniging werkt momenteel aan een eigen systeem om kinderen te volgen, gebaseerd op observaties. Dit is nog in ontwikkeling, maar leent zich in de nabije toekomst ook voor allerlei analyses op zowel individueel als groepsniveau.

Naast de methodeonafhankelijke toetsen uit de leerlingvolgsystemen worden ook veelvuldig methodegebonden toetsen gebruikt. Er zijn dus in het primair onderwijs veel data voorhanden, dit geldt zowel voor het reguliere basisonderwijs als voor het speciaal basisonderwijs. Er is meestal ook een vorm van centrale invoer; in de regel heeft de interne begeleider hierover de regie. Soms zijn er externe bureaus die toetsdata bij de scholen opvragen en op basis daarvan schoolrapportages maken.

7.5 Gebruik van de gegevens over leervorderingen

De derde vraag in dit onderzoek luidde:

In hoeverre en hoe worden deze gegevens gebruikt voor (a) diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen en (b) evaluatie van het onderwijs voor de hele groep of specifieke subgroepen? Wordt daarbij gebruik gemaakt van vooraf gestelde leerdoelen?

Uit enkele eerder verrichte Nederlandse empirische studies is gebleken dat de gegevens gebruikt worden om het onderwijs voor individuele leerlingen te plannen/bij te stellen, om het onderwijsaanbod te evalueren, om het effect van handelingsplannen vast te stellen, voor rapportage aan ouders en voor het onderbouwen van schoolkeuzeadviezen. Daarbij staat diagnostiek en aanpassing van het onderwijs aan individuele leerlingen voorop, evaluatie van het onderwijs aan de hele groep krijgt minder aandacht. Verder blijkt uit deze studies dat slechts een minderheid van de leraren daarbij vooraf doelen stelt. Een andere bevinding is dat de gegevens relatief weinig gebruikt worden voor feedback aan leerlingen (leer-

lingen inzicht geven in eigen vorderingen); er vindt vooral terugkoppeling plaats naar leerkrachten en ouders.

Deze studies hebben enkele jaren geleden plaatsgevonden. Is dit beeld nog actueel, als we kijken naar wat het onderzoek naar meetgestuurd onderwijs heeft opgeleverd? We vinden op sommige punten bevestiging, op andere nuanceringen en aanvullingen.

Bevestigd wordt dat analyse van gegevens wat vaker plaatsvindt op individueel niveau dan op groepsniveau. De telefonische enquête heeft dit laten zien en de interviews met de praktijkdeskundigen eveneens. Het meest kijkt men naar ontwikkelingscurves van zorgleerlingen (75% van de basisscholen doet dit en 64% van de scholen voor speciaal basisonderwijs). Daarbij bekijkt men ook of individuele (zorg)leerlingen gestelde doelen halen, aan de hand van de doelen in de handlingsplannen. De meeste basisscholen zeggen ook analyses op groepsniveau uit te voeren. Dit gebeurt op scholen die volgens de inspectie zwakkere leerresultaten hebben (voor taal) wat vaker dan op scholen die volgens de inspectie voldoende resultaten behalen: 69% van de basisscholen zegt te analyseren op groepsniveau en 74% van de taalzwakke scholen. In het speciaal basisonderwijs vindt analyse op groepsniveau echter heel weinig plaats (20% van de scholen). Vooral analyses die vergelijking met landelijke gemiddelden impliceren zijn in het speciaal basisonderwijs nagenoeg afwezig.

De vervol ginterviews met een klein deel van deze scholen hebben informatie opgeleverd over de achtergronden van deze bevindingen. Gebleken is dat:

- analyseren op groepsniveau vaker gebeurt op scholen met zwakkere leeropbrengsten dan op scholen met voldoende prestaties, dit komt doordat men bij voldoende resultaten niet meer de behoefte heeft om verder te analyseren;
- de aanleiding om analyses op groeps- en schoolniveau te gaan doen nogal eens de wens van inspectie of bestuur is;
- men in het speciaal basisonderwijs weinig voelt voor groepsanalyses omdat men vindt dat er te veel verschillen zijn tussen de leerlingen en groepsresultaten daarom weinig zeggen;

Opbrengstgericht werken

- zowel voor analyses op groepsniveau als op individueel niveau de normen uit het Cito leerlingvolgsysteem leidend zijn; het stellen van eigen, schoolspecifieke doelen komt minder voor;
- men in de analyses vaker gebruik maakt van de grove indeling in de vijf niveaugroepen A tot en met E uit het Cito systeem dan van de meer precieze informatie biedende vaardigheidsscores;
- de meeste scholen bij de groepsanalyses alleen globaal in de gaten houden of ze niet te veel afwijken van de landelijke verdeling wat betreft de vijf niveaugroepen.

Op de vijf scholen uit de casestudies vinden zowel individuele analyses als groepsanalyses plaats. De toetsgegevens worden altijd gebruikt voor diagnostiek en aanpassingen in het onderwijs voor individuele leerlingen, zeker voor de leerlingen van wie de prestaties op D- of E-niveau liggen (de twee laagste niveaus). Voor die leerlingen wordt over het algemeen een handelingsplan ontwikkeld, en na enige tijd opnieuw gekeken naar hun prestaties. Ook vindt evaluatie plaats van het onderwijs voor de hele groep of subgroepen: soms worden groepshandelingsplannen opgesteld wanneer een deel van de leerlingen bepaalde onderdelen van de leerstof nog niet blijkt te beheersen. In uiterste consequentie worden bij achterblijvende leerresultaten op schoolniveau ook verbeteringstrajecten ingezet: ingrijpende vernieuwing van het lees- of rekenonderwijs, invoering van een methodiek voor woordenschatontwikkeling, integratie van het taal- en zaakvakonderwijs etc. Alle scholen hanteren daarbij vooraf gestelde leerdoelen, uitgangspunt zijn de kerndoelen zoals die in de methodes aan de orde komen, daarnaast wordt meestal vastgelegd voor welk percentage van de leerlingen minimaal een voldoende niveau (een A-, B- of C-score) op de taal- en/of rekentoetsen gehaald moet worden. Ongeveer 10 tot 15% van de leerlingen op D- of E-niveau vinden de scholen meestal acceptabel, alleen voor technisch lezen worden strengere normen gesteld. De Cito-Eindtoets die wordt afgenomen in groep 8 wordt door de scholen niet gebruikt om leerstandaarden mee te bepalen.

De praktijkdeskundigen oordeelden verschillend over de huidige praktijk ten aanzien van meetgestuurd onderwijs. Enkele schoolbegeleiders, die ervaring hebben met het begeleiden van scholen in schoolverbeteringsprojecten, zijn overwegend positief: ze rapporteren dat leerkrachten goed gebruik maken van toetsresultaten

om het onderwijs af te stemmen op de behoeften van leerlingen en dat scholen toetsresultaten analyseren om te beoordelen of men gestelde doelen behaalt. Het gaat hier echter niet om 'doorsnee' scholen, maar om scholen met veel leerlingen uit achterstandsgroepen voor wie het bewaken en verbeteren van het onderwijsniveau een nadrukkelijke opgave is.

De inspectie toont zich veel kritischer en wijst er op dat meer dan de helft van de scholen nog niet voldoet aan normen ten aanzien van evalueren van leren en onderwijzen. Ook vindt de inspectie dat weinig scholen goede analyses maken, deels vanwege ontbrekende deskundigheid en deels vanwege onvoldoende intrinsieke motivatie (scholen zien er volgens de inspectie niet zelf het nut van in). Ook maken scholen volgens de inspectie weinig gebruik van leerstandaarden op schoolniveau, zij het dat zwakke scholen op dit gebied wel iets actiever zijn. Toetsen uit het leerlingvolgsysteem worden volgens de inspectie vooral benut voor de individuele leerlingenzorg, veel minder voor het uitvoeren van analyses op groepsniveau en nog minder voor trendanalyses.

Ook de Cito-medewerkers vinden dat er nog veel verbeterd kan worden. Ze zijn van mening dat scholen nog lang niet voldoende de mogelijkheden benutten die het Cito leerlingvolgsysteem biedt voor analyse van gegevens ter evaluatie van het onderwijs.

We zien dus dat er een zekere discrepantie is tussen dat wat scholen zelf melden over wat ze doen aan meetgestuurd onderwijs en dat wat inspectie en Cito-medewerkers waarnemen. De oorzaak daarvan ligt in een verschil in standaarden. De Cito-medewerkers merken in hun contacten met scholen (bijvoorbeeld tijdens de cursussen die ze geven) dat er nog veel deskundigheid ontbreekt en dat schoolleiders en interne begeleiders lang niet altijd voldoende doorzien wat ze uit hun gegevens kunnen halen. De inspectie verwacht van scholen dat ze de leeropbrengsten analyseren en daarbij per vak naar ontwikkelingen in de tijd kijken. De scholen weten niet altijd wat er mogelijk is, respectievelijk verwacht wordt en vinden het in veel gevallen voldoende als zorgleerlingen worden gesignaleerd en globaal wordt bijgehouden of de gemiddelde leerprestaties ongeveer op het verwachte niveau liggen.

7.6 De stappen in de evaluatieve cyclus

De volgende vraag in het onderzoek luidde:

Welke handelingen/stappen van meetgestuurd onderwijs worden op scholen uitgevoerd, welke niet, en waarom?

We hebben dit ingevuld als: in hoeverre nemen leraren en schoolleiders de evaluatieve cyclus als leidraad voor hun evaluaties?

De stappen van de evaluatieve cyclus: zijn het vastleggen van doelen en standaarden, het verzamelen van informatie, het registreren ervan, het interpreteren en het nemen van beslissingen.

Volgens de praktijkdeskundigen worden de stappen ‘verzamelen’ en ‘registreren’ door vrijwel alle scholen gezet, maar is dat veel minder het geval voor de stappen doelen vastleggen, gegevens interpreteren en beslissingen nemen. Op scholen die dat laatste wel doen ontstaan positieve effecten, volgens de geïnterviewde schoolbegeleiders die betrokken zijn bij projecten waarin dataverzameling en –analyse van leerprestaties plaatsvindt:

- men wordt zich sterker bewust van de eigen onderwijskwaliteit;
- men gaat nauwkeuriger signaleren waar in het onderwijs op de school nog zwakke plekken zitten;
- het voorkomt ongemerkt afzakken van het niveau, omdat de resultaten worden vergeleken met landelijke gegevens, en/of worden afgezet tegen een referentieniveau
- scholen die zwak presteren gaan gerichte maatregelen nemen om hun onderwijskwaliteit te verbeteren.

De scholen zelf zijn ook hier weer wat positiever over de eigen praktijk dan de deskundigen rapporteren. Uit de telefonische enquête bij de scholen is naar voren gekomen dat een meerderheid van zichzelf vindt dat ze alle stappen van de evaluatieve cyclus uitvoeren. Per stap gaat het om de volgende resultaten:

Stap 1, doelen stellen

In het reguliere basisonderwijs hebben de meeste scholen (ca 80 procent) vastgelegde leerstandaarden, behalve in de groepen 1 en 2. Dit geldt voor zowel lezen

als rekenen. Wel is er enig verschil tussen lezen en rekenen: leerstandaarden voor lezen heeft men in de onderbouw (vooral groep 3 en 4) iets vaker dan voor rekenen, in de bovenbouw is dat juist omgekeerd. Op taalzwakke scholen zijn er wat minder vaak leerstandaarden dan op niet taalzwakke scholen. Van de scholen voor speciaal basisonderwijs heeft maar ongeveer een kwart vastgestelde leerstandaarden, dit geldt voor zowel lezen als rekenen en voor alle groepen.

Stap 2 en 3, verzamelen en registeren

Vrijwel alle scholen beschikken over een leerlingvolgsysteem met opslag van gegevens over leervorderingen op verschillende vakken. Nagenoeg altijd voor lezen en rekenen, zeer vaak voor overige onderdelen van het taalonderwijs (zoals spelling, woordenschat) en nog relatief weinig voor andere leerdomeinen. Dit geldt voor het regulier en het speciaal basisonderwijs.

Stap 4, interpreteren

Een groot deel van de scholen geeft aan de verzamelde gegevens ook te analyseren en te interpreteren, namelijk ruim 80 procent van de basisscholen en 74 procent van de scholen voor speciaal basisonderwijs.

Stap 5, beslissingen nemen

Beslissingen over te ondernemen stappen worden voornamelijk genomen voor individuele leerlingen, het gaat hier om het opstellen en bijstellen van handelingsplannen voor zorgleerlingen. Indien nodig worden ook beslissingen genomen over het onderwijsaanbod voor de hele groep, dan gaat het bijvoorbeeld om de keuze van een nieuwe methode of het inzetten van extra lesmateriaal.

Een discrepantie tussen norm en eigen resultaten is niet altijd aanleiding tot actie. Soms vindt men de norm voor de eigen school te hoog gegrepen of heeft men het vertrouwen/de ervaring dat een geconstateerde achterstand nog wel bijtrekt. Slechts een enkele (taalzwakke) school maakt melding van acties ter verbetering van de kwaliteit van de leerkrachten (coaching, scholing).

Zo samengevat, ontstaat gemakkelijk de indruk dat het er vrij goed voorstaat met het toepassen van de evaluatieve cyclus. Maar uit de vervol ginterviews met de scholen kwamen enkele zaken naar voren die dit beeld nuanceren.

Opbrengstgericht werken

Zo viel het op dat men soms niet goed raad wist met de stap ‘interpreteren’. Hieronder bleek men niet altijd het interpreteren van de gegevens zelf te verstaan, maar eerder het uitzoeken waar het aan ligt dat prestaties mee- of tegenvallen. We vermeldde al dat analyse van gegevens op sommige scholen niet verder gaat dan ‘globaal in de gaten houden’. Op groepsniveau kijkt men dan vooral naar de verdeling van prestaties op een bepaald moment, ten opzichte van de normindeling van het Cito, en soms naar de ontwikkeling van de groep leerlingen in de tijd. Wat veel minder plaatsvindt is prestaties vergelijken met die van parallelgroepen, groepen uit een vorig leerjaar of andere scholen.

Verder worden leerkrachten niet altijd bij de interpretatiestap betrokken. Resultaten worden dan alleen met hen besproken als er ‘onder de maat’ wordt gepresteerd. Ook komt het maar weinig voor dat resultaten van de (groeps)analyses worden besproken met bestuur of medezeggenschapsraad.

Op de scholen uit de casestudies (voorbeeldscholen) is men zich er goed van bewust dat het juist leerkrachten zijn die zich moeten buigen over de leerresultaten. Wanneer het om individuele leerlingen gaat, is er meestal sprake van overleg tussen de groepsleerkracht en de interne begeleider over toetsresultaten, ontwikkelingsverloop en het handelingsplan. De groepsresultaten op schoolniveau worden over het algemeen besproken in een teamoverleg. Een enkele keer vormen ze ook (al) een aandachtspunt in het functioneringsoverleg. Vaste aandachtspunten zijn: wat gaat goed, wat nog niet, hoe komt dat en hoe kan het beter? Zowel op individueel niveau, als op groeps- en schoolniveau worden vervolgens nieuwe doelen gesteld, en nieuwe evaluatiemomenten bepaald.

7.7 Draagvlak voor meetgestuurd onderwijs

De vijfde vraag in het onderzoek was:

Bestaat er draagvlak bij leerkrachten en schoolleiders voor meetgestuurd onderwijs?

Uit het feit dat scholen naar hun eigen opinie in meerderheid (elementen van) meetgestuurd onderwijs al toepassen kan afgeleid worden dat ze er ook positief tegenover staan, het belangrijk vinden. Dat blijkt ook uit de uitkomsten van enkele

le stellingen die we voorlegden in de telefonische enquête. De respondenten zijn het in meerderheid eens met de volgende uitspraken:

- verbetering van het onderwijs in taal en rekenen kan niet zonder regelmatig meten van je onderwijsopbrengsten
- op onze school is er voldoende openheid in het team om met elkaar de leeropbrengsten in ieders groep te bespreken
- op onze school is er voldoende deskundigheid om analyses te plegen op de gegevens in het leerlingvolgsysteem
- jezelf vergelijken met andere scholen met hetzelfde type leerlingenpubliek is een goede manier om te evalueren of de leerlingen presteren wat verwacht mag worden.

Uit de vervol ginterviews komen enkele aanvullingen op dit beeld. Zo blijkt dat op scholen die net begonnen zijn met meetgestuurd onderwijs nog wel wat weerstand van (een deel van) de leerkrachten moet worden overwonnen. Het gaat dan niet zozeer om weerstand tegen het zichtbaar maken van de resultaten, maar wel tegen het werk dat daarvoor nodig is (regelmatig toetsen, toetskalender aanhouden, toetsscores registreren en invoeren). Ook komt het voor dat leerkrachten het bespreken van de resultaten bedreigend vinden. Weerstand kan ook optreden wanneer de uitkomsten aanleiding geven tot substantiële veranderingen in het onderwijs.

De reden dat het draagvlak tamelijk groot lijkt, kan ook samenhangen met het feit dat analyses die scholen plegen in een aantal gevallen bescheiden of oppervlakkig zijn. Men erkent wel dat het soms beter of preciezer zou kunnen, maar dat kost meer tijd. Scholen die voldoende of goede resultaten boeken, geven daaraan geen prioriteit. *'We beschikken over heel veel gegevens en doen er niet zoveel mee. Dat roept wel eens vragen op. We kiezen er voor om wel de vinger aan de pols te houden, maar niet door te analyseren zolang de resultaten goed zijn'. 'Trendanalyses zouden kunnen, maar de directe aanleiding ontbreekt en je kunt je tijd maar een keer uitgeven'*, zijn twee uitspraken die in dit verband door schoolleiders zijn gedaan.

Voor het speciaal basisonderwijs geldt dat daar weinig draagvlak bestaat voor analyses met instrumenten en procedures die afkomstig zijn uit het reguliere basisonderwijs. De sbo-scholen hebben behoefte aan een leerlingvolgsysteem dat meer is toegesneden op het eigen schooltype en waarmee leerlingen op maat kunnen worden getoetst.

Opbrengstgericht werken

Op de scholen waar de casestudies zijn gehouden is het draagvlak groot, zowel bij de schoolleiders als bij de leerkrachten. Men is met name enthousiast over het feit dat meetgestuurd werken leidt tot gericht en efficiënter onderwijs, meer bewuste keuzes en meer greep op de leerstof. Er vindt meer uitwisseling plaats binnen het team, leerkrachten hebben meer zicht op de ontwikkelingscurves van hun leerlingen, en dat betekent dat er in grotere mate sprake is van doorgaande leerlijnen.

Volgens de praktijkdeskundigen is over het geheel genomen het draagvlak voor meetgestuurd onderwijs groeiende, maar nog te gering. De scholen zijn nog lang geen 'lerende organisaties', ze zijn nog te weinig nieuwsgierig naar de eigen kwaliteit en te weinig gericht op het bereiken van een zo hoog mogelijk niveau, zo is meermalen opgemerkt. Vooral scholen met weinig kinderen uit achterstandsgroepen hebben volgens de experts geringe belangstelling voor meetgestuurd onderwijs. Als de resultaten toch wel goed zijn, ziet men geen aanleiding om veel aandacht te besteden aan analyse van resultaten, is de ervaring. Op zwakke scholen of scholen met veel kansarme leerlingen is volgens de experts het draagvlak groter: daar is de noodzaak om prestaties te bewaken en waar nodig te verhogen groter.

Een middel om draagvlak te vergroten is het ervaren van succes, volgens de schoolbegeleiders. Leerkrachten die merken dat beter inzicht in resultaten leidt tot aanpassing van het onderwijs en dat die aanpassing vervolgens weer tot betere resultaten leidt, worden enthousiast.

Referentieniveaus

In het kader van de verkenning van draagvlak is in de vervolginterviews met schoolleiders/interne begeleiders ook gevraagd naar hun mening over de invoering van landelijke referentieniveaus. Daarover blijkt wisselend te worden gedacht. Een minderheid vindt de plannen voor landelijke referentieniveaus een goede ontwikkeling. Eveneens een minderheid is er beslist tegen. De meesten vinden het of overbodig ('*We hebben de normering van de Cito LVS toetsen toch al?*') of zijn het wel met invoering eens, maar onder de voorwaarde dat er bij het opstellen van referentieniveaus rekening wordt gehouden met verschillen tussen scholen in leerlingenpubliek.

We stelden deze vraag ook aan de leerkrachten die hebben meegedaan aan een pilotproject van het Cito voor de ontwikkeling van referentieniveaus voor groep 5. Bij hen troffen we het volgende beeld aan:

- De betrokken leraren zijn overwegend positief over het gebruik van landelijke referentieniveaus. Iets minder dan de helft van hen kiest voor het hanteren van twee niveaus, een minimum- en een voldoende niveau zoals in het pilotproject van het Cito. Eén op de drie leraren hanteert echter liever meer dan twee niveaus om beter recht te kunnen doen aan verschillen tussen leerlingen.
- Ondanks deze positieve houding zien de leerkrachten toch ook bezwaren. Een op de drie leraren vreest dat de niveaus nadelig uitpakken voor scholen met veel achterstandsleerlingen. Zulke scholen hebben het moeilijker de afgesproken niveaus te halen. Aan de andere kant noemen leraren ook voordelen. Ongeveer de helft van hen ziet als voordeel dat landelijke referentieniveaus klaarheid scheppen die er nu nog niet is.
- Ongeveer de helft van de leraren onderschrijft de verwachting dat referentieniveaus zullen bijdragen aan betere resultaten. Leraren beseffen wel dat alleen het vastleggen van landelijke referentieniveaus daarvoor niet genoeg is. Het gaat om de follow-up: wat doet de school als resultaten tegenvallen? Een deel van de leraren is sceptisch, omdat men meent dat de marges voor onderwijsverbetering smal zijn. De kwaliteit is al heel hoog.

De onderzochte scholen voor speciaal basisonderwijs staan niet afwijzend ten opzichte van een landelijk referentiekader, maar dan wel in aangepaste vorm. *‘Als er een landelijke meetlat voor sbo-leerlingen zou zijn, zou dat wel een spiegel voor de scholen kunnen vormen’*. De scholen zijn beducht voor ‘afrekening’ op hun eindresultaten.

7.8 Risicofactoren, succesfactoren, onbedoelde effecten

De zesde vraag in het onderzoek was:

Wat zijn risico- en succesfactoren bij meetgestuurd onderwijs? Zijn er eventueel onbedoelde effecten?

Opbrengstgericht werken

Risicofactoren, belemmerende factoren, negatieve bijeffecten

Uit de literatuurstudie zijn enkele valkuilen naar voren gekomen die zich kunnen voordoen bij vormen van meetgestuurd onderwijs (zie 7.2). We herhalen ze hier nog even kort: bevorderen van competitiegedrag, te veel nadruk op verantwoordingsseisen, te sterke focus op zwakke leerlingen en daardoor onderstimulering van de betere leerlingen en uitlokken van strategisch gedrag van scholen, zoals zwakke leerlingen buiten de toetsing houden en teaching to the test.

Zijn er in dit onderzoek aanwijzingen gevonden dat dergelijke valkuilen zich daadwerkelijk voordoen?

De praktijkdeskundigen noemen de volgende risico's, waarbij overigens moet worden aangetekend dat niet altijd duidelijk was of dit risico's zijn die zij vrezen of die zij al daadwerkelijk zijn tegengekomen:

- Scholen kunnen verkeerde conclusies trekken uit de verzamelde cijfers of blijven steken in een te mechanistische aanpak
- Teaching to the test kan zich voordoen
- Een overwaardering van de toetsresultaten en een te sterke prestatiegerichtheid kan ten koste gaan van de aandacht voor het leerproces zelf.
- Teveel toetsen kan ten koste gaan van de onderwijstijd. Het risico van teveel toetsen ontstaat wanneer scholen oude systemen niet voldoende durven loslaten wanneer er nieuwe voor in de plaats komen, wanneer het alleen gebeurt 'omdat het moet van de inspectie', of wanneer er vanwege afspraken tussen schoolbesturen (bijvoorbeeld per gemeente) op alle scholen dezelfde toetsen moeten worden afgenomen. In die situaties zijn de gegevens namelijk niet voor alle scholen ook echt informatief, en zijn er dus geen effecten op de onderwijsopbrengst.
- Meetgestuurd onderwijs kan leiden tot een afrekencultuur: ofwel op bestuurlijk niveau van scholen die zwakker presteren, ofwel op schoolniveau van leerkrachten die minder goede resultaten weten te behalen. In dat geval raakt het doel, verhogen van de kwaliteit van het onderwijs, uit beeld.
- Een (te) sterke nadruk op toetsen en prestaties kan leiden tot pressie van ouders op leerkrachten om het kind naar een volgend niveau te laten doorgaan, en tot prestatiedruk bij kinderen, met als gevolg wedijver, faalangst, het verlies van plezier in leren of welbevinden op school.
- Voor het montessori-onderwijs botst een werkwijze waarin op vaste momenten in het schooljaar toetsen moeten worden afgenomen (zoals met het

Cito-LVS) met de onderwijsvisie. Het risico is dat scholen of leerkrachten de neiging krijgen om bepaalde leerstof dan toch maar alvast te gaan behandelen voordat de toetsperiode aanbreekt, ook als dat niet aansluit bij het verloop van de ontwikkeling van het kind op dat moment.

De scholen zelf noemen niet veel risico's. Wat de scholen het meest naar voren brengen is dat *gebrek aan tijd* een bedreiging vormt voor het toepassen van meetgestuurd onderwijs. Het kan dan gaan om gebrek aan tijd om de metingen uit te voeren en om (goed) te analyseren, maar vaker doelt men op de tijd die nodig is om de uit de meetresultaten voortvloeiende beslissingen uit te voeren, namelijk veranderingen doorvoeren in de werkwijzen. Scholen zoeken voor dit laatste verschillende oplossingen: extra formatie inzetten, invoeren van zelfstandig werken om extra instructietijd vrij te maken voor de leerkrachten; vergroting van de groepen zodat er meer hulp buiten de les kan worden georganiseerd, slim benutten van het rooster, inzetten van zorgmiddelen (rugzakjes).

Een risico dat scholen wel onderschrijven is dat er een kans is op 'stigmatiseren' van leerlingen of leerkrachten. In dit verband wordt er op gewezen dat meten eventuele tekorten zichtbaar maakt, maar dat dat nog niet betekent dat die tekorten ook altijd opgeheven kunnen worden.

Dat de toetsen het onderwijs teveel gaan bepalen, zien de onderzochte scholen niet als reëel gevaar. '*Geen gevaar voor teaching to the test, we schieten heus niet door*'.

De scholen in de casestudies vinden het belangrijkste risico van meetgestuurd onderwijs dat er in de oordeelsvorming over de prestaties een te sterke nadruk op de toetsresultaten zou kunnen ontstaan. Wanneer alleen de toetsresultaten nog tellen, kan dat ten koste gaan van de eigen professionaliteit en inzichten van de leerkracht. Een tweede risico is het ontstaan van een 'afrekencultuur' waarin de leerkracht verantwoordelijk wordt gehouden voor achterblijvende prestaties van de leerlingen, en de oorzaken te weinig aandacht krijgen.

Succesfactoren, bevorderende factoren, positieve bijeffecten

Het zijn vooral de scholen die, meer dan de praktijkdeskundigen, de successen van meetgestuurd onderwijs noemen: je kunt je onderwijs gericht bijstellen; planmatig werken verbeterd; er komt meer aandacht voor differentiatie, voor af-

Opbrengstgericht werken

stemming tussen de groepen, voor klassenmanagement; het maakt van de school een lerende organisatie; er ontstaat een beter inzicht in zwakke plekken.

Er zijn enkele scholen die ervaring hebben met een sterkte/zwakte analyse van de school waarbij ook analyses van leerprestaties zijn meegenomen. Dit leidt in de regel tot een vrij ingrijpende aanpassing van het onderwijs. De aanleiding om dit te doen is de komst van een nieuwe directeur en/of kritiek van de inspectie.

In het speciaal basisonderwijs is een nieuwe ontwikkeling dat men ontwikkelingscurves van alle leerlingen analyseert met behulp van uitstroomprofielen. Een uitstroomprofiel is een prognose voor wat een kind aan niveau kan halen bij de overgang naar het voortgezet onderwijs en wordt gebruikt om tussentijds te evalueren of men ten opzichte van die prognose op de goede weg zit (er voldoende progressie is). Indien nodig leidt dit tot bijstelling van het onderwijsaanbod. Dit gebeurt aan de hand van gegevens als IQ-tests, leerlingkenmerken (werkhouding etc.), diagnostisch onderzoek, thuissituatie, onderwijsleeftijd (DLE), Cito-uitkomsten (rekenen, spelling) en soms AVI-uitkomsten.

De ervaringen hiermee zijn nog wisselend. Niet alle leerkrachten zijn al goed in staat om het onderwijs daadwerkelijk aan te passen als dat geïndiceerd is, men is beducht voor eventuele onderstimulering en het blijkt vooral voor jonge kinderen moeilijk om een helder uitstroomprofiel te maken.

De scholen in de casestudies noemen als succesfactoren de gerichte aandacht voor prestaties en het nauwgezet volgen van de ontwikkeling van de leerlingen en het daarop afstemmen van het onderwijs. Dit maakt dat de leerkrachten meer greep hebben op hun onderwijs en bewustere keuzes maken bij de selectie van leerstof. Een veilige, open, collegiale sfeer in het team en goede relaties tussen team en directie vormen de tweede belangrijke succesfactor. In die situatie voelen de leerkrachten zich inderdaad verantwoordelijk voor de leerprestaties, en eigenaar van het onderwijs in hun groep, maar voelen zij zich ook in voldoende mate gesteund om die verantwoordelijkheid, gezamenlijk, te dragen. De constatering dat de leerprestaties inderdaad vooruitgaan werkt daarbij voor zowel de leerkrachten als de leerlingen zelf extra motiverend.

7.9 Conditie

De zevende en laatste vraag in het onderzoek was:

Wat hebben schoolleiders en leerkrachten nodig om meetgestuurd onderwijs te kunnen uitvoeren? Tot welke aanbevelingen leidt dit voor het beleid van de overheid en andere actoren?

In de beleidsnota *Scholen voor morgen*, waarin het belang van meetgestuurd onderwijs wordt uiteengezet, is ook aandacht voor de condities die hiervoor nodig zijn. Conditie die in de nota genoemd worden zijn

- schoolleiders die in staat zijn om kwaliteitsbeleid te voeren en een ontwikkel- en verbetercultuur op de school te creëren
- de inspectie moet stimuleren en intensief toezicht houden op scholen die zwak of risicovol zijn
- leren van elkaar: door uitwisseling van kennis tussen scholen zou meetgestuurd onderwijs bevorderd kunnen worden.

Uit de onderzoeksliteratuur (zie hoofdstuk 1) komen ook condities ook naar voren, zoals:

- een ‘onderzoekende cultuur’ in scholen: schoolleiders en leerkrachten moeten een attitude hebben van ‘willen weten’ en bereid zijn tot reflectie op resultaten van metingen
- schoolleiders en leerkrachten moeten in staat zijn tot kritisch zelfonderzoek
- er moet een schoolklimaat zijn waarin op een open, veilige manier over leeropbrengsten en werkwijzen gesproken kan worden; scholen moeten in staat zijn ‘leerconversaties’ te houden
- schoolleiders moeten in staat zijn leerkrachten te steunen en te stimuleren en een eenduidige visie uitdragen
- schoolleiders en leerkrachten moeten ‘data onderlegd’ zijn, dat wil zeggen data kunnen analyseren, verwerken, interpreteren en benutten; dit vergt onder meer dat ze data kunnen omzetten in betekenisvolle kennis voor de school
- externe ondersteuning kan hierbij bevorderend werken
- schoolbesturen/bovenschools managers en eventueel inspectie kunnen stimuleren in initiëren

Opbrengstgericht werken

De verschillende betrokkenen in dit onderzoek hebben bij de vraag naar de condities verschillende accenten gelegd. De praktijkdeskundigen benadrukken vooral het belang van voldoende deskundigheid bij schoolleiders en interne begeleiders, in de zin van ‘data onderlegd zijn’ (zie boven). Zij constateren dat er op dit gebied nog veel te wensen is en dat de meeste schoolleiders en interne begeleiders nog niet voldoende weten van de mogelijkheden tot data-analyse en nog onvoldoende in staat zijn om vanuit data-analyse te komen tot zinvolle interpretaties en keuze van acties voor verbetering.

De ronde die we hebben gehouden langs de pabo’s heeft laten zien dat studenten op de pabo’s het voor de ontwikkeling van hun kennis over toetsing, interpretaties van resultaten en afstemming van het onderwijsaanbod op die resultaten vooral van hun praktijkstages moeten hebben. Op de opleiding wordt daar weinig systematische aandacht aan besteed, met als uitzondering het oefenen met het opstellen van een handelingsplan. Voor het leren kennen en leren toepassen van meetgestuurd onderwijs zijn de studenten dus afhankelijk van de kwaliteit van hun stage-school. Bovendien is de leer- en oefenstof voor studenten gedeeltelijk afhankelijk van de specialisatie die zij kiezen vanaf het derde jaar. Dat betekent dat niet alle toekomstige leerkrachten in dezelfde mate leren om toetsresultaten op groepsniveau te interpreteren, hun eigen onderwijs te evalueren en/of een handelingsplan te maken op basis van toetsresultaten. En omdat ook de huidige leerkrachten dat niet in hun opleiding hebben geleerd, zijn er dus goede gronden om, zoals de praktijkdeskundigen ook aangeven, te constateren dat er nog sprake is van onvoldoende deskundigheid in scholen.

De scholen zelf hebben echter niet of nauwelijks gewag gemaakt van een eventueel gebrek aan deskundigheid. In de casestudies heeft slechts een enkele leerkracht opgemerkt dat zij behoefte heeft aan scholing in het leren omgaan met data. De respondenten in de telefonische interviews vonden, desgevraagd, dat er geen specifieke deskundigheid voor analyseren van toetsresultaten nodig is. Maar in ongeveer de helft van de gevallen bleek de interne begeleider van deze scholen wel een cursus gevolgd te hebben, vaak bij het Cito.

De tweede conditie die praktijkdeskundigen hebben genoemd is een op onderwijsresultaten gericht schoolbestuur. Een bestuur dat inzicht wil in de resultaten van de leerlingen draagt er toe bij dat scholen data daarover gaan verzamelen en ook zelf willen analyseren. De cases hebben laten zien dat een actieve rol van het

schoolbestuur inderdaad van belang is. Op één van de scholen uit de casestudies voorziet het schoolbestuur in administratieve ondersteuning.

De derde conditie die de praktijkdeskundigen hebben genoemd is het hebben van een duidelijk schoolbeleid: een visie op onderwijs en een meerjaren beleidsplan waarin is vastgelegd waar men aan wil werken, waarom en met welke doelstellingen.

De scholen hebben vooral de nadruk gelegd op het belang van een goede onderlinge sfeer in het team en van gezamenlijke verantwoordelijkheid voor het onderwijs en voor de resultaten die bij de leerlingen worden bereikt. De schoolleider vervult daarin een belangrijke rol. Verder komt uit de casestudies naar voren dat het afnemen van de toetsen én het werken op verschillende niveaus veel tijd kost. En tot slot vindt men een goede afstemming tussen de methodes enerzijds en de inhoud van de toetsen anderzijds van belang, evenals stroomlijning in de af te nemen toetsen en de registratie daarvan.

7.10 Discussie en aanbevelingen

In deze paragraaf dragen we tot slot enkele discussiepunten en aanbevelingen aan, naar aanleiding van de onderzoeksresultaten.

De relatie tussen meetgestuurd onderwijs en prestaties verhogen

De ambitie van het huidige kabinet in het onderwijsbeleid voor het primair onderwijs is het verhogen van de prestaties voor taal en rekenen. Het bevorderen van meetgestuurd onderwijs is één van de instrumenten die worden ingezet om deze ambitie waar te maken. Is het nu, in het licht van de bevindingen van dit onderzoek, plausibel dat bevorderen van meetgestuurd onderwijs er toe zal leiden dat reken- en taalprestaties omhoog gaan? Op deze vraag past een genuanceerd antwoord. Aan de ene kant is het inderdaad zo dat deze praktijk bijdraagt aan een verhoogd inzicht van leerkrachten en schoolleiders in de eigen onderwijsresultaten en daarmee leidt tot het nemen van beslissingen om het onderwijs meer af te stemmen op onderwijsbehoeften van individuele leerlingen of geconstateerde tekorten in het onderwijs te gaan bestrijden. Dit beeld hebben we gezien op de

Opbrengstgericht werken

scholen uit de casestudies. Maar het gaat hier wel om voorlopers in meetgestuurd onderwijs. We hebben aan de andere kant ook geconstateerd dat er scholen zijn die alleen oppervlakkig naar hun toetsresultaten kijken en daar verder niet veel mee doen. Toch vinden ook die scholen dat ze bezig zijn met meetgestuurd onderwijs. Ten onrechte, want het is natuurlijk zo dat alleen het meten en het bekijken van meetresultaten nog niet tot betere leerresultaten leidt. Voor betere leerresultaten zijn veranderingen nodig in de klas, in het handelen van leerkrachten. Blijkens de onderzoeksliteratuur liggen effectieve maatregelen op het terrein van meer leertijd, betere instructie, meer en betere feedback aan leerlingen, hoge eisen stellen, aandacht voor leerstrategieën. Vooral van dergelijke maatregelen valt een direct effect op de prestaties van de leerlingen te verwachten. Alle activiteiten op andere niveaus (schoolleiding, bestuur, ondersteuningsinstellingen, overheid) zijn van voorwaardelijke aard en niet doorslaggevend in effect op leerresultaten (Scheerens, 2008).

Het is dus van belang de maatregelen om meetgestuurd onderwijs te bevorderen zo in te richten dat daarmee daadwerkelijk de leerkrachten worden bereikt. Schoolleiders en interne begeleiders spelen daarbij een cruciale rol. Zij zijn degenen die het voortouw moeten nemen bij het vertalen van uitkomsten van metingen naar schoolbeleid en –activiteiten en leerkrachten moeten activeren om daaraan hun eigen bijdrage te leveren. Uit eerder onderzoek is gebleken dat voor het op gang brengen van veranderingen in de school het verzamelen en analyseren van data een bruikbare methodiek is, maar dat die methodiek impliceert dat (a) er in de school als geheel moet worden nagedacht over de vraag wat de relevantie is van de verzamelde data en (b) de data bij voorkeur worden ingezet in processen van formatieve evaluatie, passend bij zelf geformuleerde doelen. Schoolontwikkeling komt niet tot stand door alleen data aan te bieden (Hofman, Vandenberghé & Dijkstra, 2008). Het gaat juist om het toepassen van de hele evaluatieve cyclus.

Deskundigheidsbevordering

De experts die wij in dit onderzoek raadpleegden vinden vrijwel zonder uitzondering dat er nog te weinig deskundigheid is bij schoolleiders en leerkrachten op het gebied van meetgestuurd onderwijs. Dan gaat het vooral om de stappen interpreteren van de gegevens en nemen van daaruit voortvloeiende beslissingen. Onderwijsgevendenden zijn onvoldoende thuis in verschillende mogelijkheden om toetsgegevens te analyseren en hebben nog te weinig een onderzoekende houding,

te weinig ‘wil om te weten’. Voor het bevorderen van meetgestuurd onderwijs is dus deskundigheidsvergroting van belang. Het gaat daarbij niet alleen om kennis van analysemethoden en regels van verantwoord onderzoek, maar ook om het verhogen van het besef dat het vastleggen en analyseren van leerresultaten tot de kern van het beroep horen en dus een normaal deel van het werk zijn. Dit lijkt een overbodige opmerking, gezien het feit dat in de omschrijving van de beroepsbekwaamheid voor leraren van de SBL aangegeven wordt dat leraren in staat moeten zijn tot ‘waarnemen, diagnosticeren, plannen, uitvoeren, volgen, evalueren, bijstellen en ontwikkelen’. Toch is dat kennelijk niet altijd vanzelfsprekend. Uit onderzoek naar de bureaucratisering in het onderwijs (Visscher, Witziers & Scherrens, 2007) is bijvoorbeeld gebleken dat sommige leerkrachten taken die te maken hebben met monitoren van leerresultaten (toetsen, gegevens invoeren) en overleg met collega’s in het kader van leerlingenzorg beschouwen als secundaire taken, die tot verzwaring van hun beroep leiden. Ook in ons onderzoek hebben we van schoolleiders gehoord dat er leerkrachten zijn die deze taken als ‘werkdruk’ ervaren. Er is dus bij een deel van de leerkrachten ook een omslag in denken nodig. Het besef zal moeten groeien dat registreren en analyseren van leerresultaten niet primair in functie staan van controle en verantwoording, maar instrumenten zijn om de eigen onderwijspraktijk te verbeteren. Dat zal niet altijd eenvoudig zijn. Earl & Katz (2006) stellen dat het een ‘mindshift’ vergt om in scholen meer onderzoekend te gaan werken: het betekent dat leerkrachten moeten gaan werken op basis van *evidence based knowledge* in plaats van de meer impliciete *tacit knowledge*.

Deskundigheidvergroting is niet alleen nodig voor leerkrachten, maar ook voor schoolleiders en interne begeleiders. Voor de schoolleiders geldt dat in veel schoolleidersopleidingen tegenwoordig aandacht wordt besteed aan onderzoeksvaardigheden en aan het leiding geven aan een school met een onderzoekende cultuur (Onderwijsraad, 2009). Voor deze groep is het dus vooral van belang dat zoveel mogelijk (zittende) schoolleiders alsnog dergelijke opleidingen volgen. Voor interne begeleiders bestaat echter niet een dergelijke opleiding. Omdat interne begeleider geen gereguleerd beroep is, bestaat er voor hen ook geen formele omschrijving van bekwaamheden. Op de meeste scholen hebben echter juist de interne begeleiders de taak om te bewaken dat leerkrachten toetsen en met hen de uitkomsten daarvan te bespreken. Zij zijn dus bij uitstek de eerste doelgroep voor deskundigheidsbevordering.

Opbrengstgericht werken

Ook op het niveau van schoolleiders en interne begeleiders gaat het niet alleen om het toevoegen van technische kennis over data-analyse. Wat ook geleerd moet worden, zijn strategieën om in de scholen ‘leerconversaties’ op gang te brengen. In algemenere termen gaat het dan om het ontwikkelen van het vermogen om te leren van data over de eigen school met het oog op het nemen van veranderingsbeslissingen (‘capacity building’, zie Fullan, 2007; Stoll et al, 2002, Wohlstetter et al, 2008; Van Petegem & Vanhoof, 2004).

Een andere doelgroep voor deskundigheidsbevordering zijn docenten van de lerarenopleidingen. In de opleidingen wordt nog maar weinig aandacht besteed aan dataverzameling en –analyse in functie van het eigen onderwijs. Dat lijkt niet alleen een kwestie van te weinig tijd/ruimte in het programma, maar ook van te weinig deskundigheid bij lerarenopleiders. Een suggestie is om te bezien of het mogelijk is toetspecialisten aan te stellen bij de opleidingen, dan wel zittende docenten daarvoor op te leiden. Verder zou gestimuleerd kunnen worden dat lerarenopleidingen werken aan onderzoeksvaardigheden en een onderzoekende houding bij hun studenten en hen leren de evaluatieve cyclus toe te passen.

Innovatieprincipes

In het verlengde van het voorgaande is het nuttig om te wijzen op enkele inzichten uit de innovatieliteratuur die van belang zijn voor de bevordering van/invoering van meetgestuurd onderwijs op scholen. We verwijzen hiervoor in onder meer naar de reviewstudie van Hofman, Vandenberghe & Dijkstra (2008).

- 1) Bij meetgestuurd onderwijs komen de opvattingen en inzichten van verschillende groepen betrokkenen bij elkaar: inspectie, externe begeleiders, onderzoekers, schoolleiders, interne begeleiders, leerkrachten. Al deze actoren hebben een eigen perspectief, een eigen praktijktheorie. Onvoldoende aandacht hiervoor is niet zelden de oorzaak van het mislukken van een vernieuwing. Er is dus veel communicatie nodig over doel en vorm van meetgestuurd onderwijs, om te zorgen dat die verschillende perspectieven op elkaar afgestemd raken.
- 2) De ene school is de andere niet. Een eenheidsaanpak is daarom ongewenst. In ons onderzoek is dat ook gebleken. We selecteerden in de telefonische ronde voor de vervolginterviews alleen scholen die eerder hadden aangegeven op veel verschillende manieren toetsresultaten te analyseren, maar in de vervolginterviews bleek dat er tussen die scholen

toch weer behoorlijke verschillen waren in de mate waarin ze dat serieus en diepgaand doen.

- 3) Hoewel veel scholen in ons onderzoek menen dat ze voldoende open en deskundig zijn in het analyseren en bespreken van toetsresultaten, laten onderzoeksresultaten (zie ook hoofdstuk 1) iets anders zien. Hulp van externen is gewenst, niet alleen in de vorm van scholing maar ook in de vorm van begeleiding op de werkvloer. Daar moet bovendien de tijd voor worden genomen: te snel vernieuwen maakt leerkrachten afhankelijk van de externe krachten en leidt niet tot professionele autonomie van leraren en tot draagvlak.
- 4) Een succesvolle leerstrategie lijkt leren van elkaar, in teams en netwerken. Voor het goed benutten van gegevens (over leervorderingen of anderszins) is het van belang dat er samen wordt geïnterpreteerd, in dialoog tussen schoolleiding en leerkrachten. Ook uitwisseling met andere scholen kan hier worden ingezet.

Aandacht voor de risico's

In het onderzoek zijn, zowel uit de literatuur als uit de interviews, diverse risico's naar voren gekomen bij invoering van meetgestuurd onderwijs. Ze worden door de scholen niet allemaal gedeeld. Zo is men niet bang voor *teaching to the test* en daaruit voortvloeiende verarming van het curriculum. Toch zijn er wel aanwijzingen dat iets dergelijks kan optreden (Townsend, 2007). Het is gewenst daar verder onderzoek naar te doen. Een ander risico dat daadwerkelijk zou kunnen optreden is *window dressing*: alleen voldoen aan de eisen van de buitenwereld (inspectie, bestuur, overheid, ouders), maar geen feitelijke verandering van de onderwijspraktijk. We hebben in het onderzoek gezien dat inspectie en bestuur de scholen kunnen aanzetten tot (elementen van) meetgestuurd onderwijs en dus kunnen bevorderen dat scholen zich daarmee gaan bezighouden. Maar een vaak gesignaleerd en besproken risico hierbij is dat scholen het dan alleen gaan doen omdat het van ze verwacht wordt en ze zich moeten verantwoorden (accountability) en niet omdat ze er zelf het nut van inzien en het willen gebruiken (zie bijvoorbeeld Onderwijsraad, 2009). Dit geldt in het bijzonder voor de scholen die van de inspectie geen onvoldoendes krijgen voor de leeropbrengsten; zij hebben nu al de neiging om minder aandacht te besteden aan analyse van leeropbrengsten. Ook dit is een aandachtspunt voor verder onderzoek.

Opbrengstgericht werken

Terzijde merken we hier op dat ouders over het algemeen niet een prikkel geven aan scholen om meer opbrengstgericht te worden. Ouders stellen zelden vragen over de opbrengsten van de school als geheel (alleen over het eigen kind), zoeken ook weinig informatie daarover en laten het niet of nauwelijks meewegen in hun schoolkeuze (Karsten et al, 2006).

Andere reële risico's zijn:

- Overmatige nadruk op de prestaties van de zwakke leerlingen, ten koste van de beter presterende leerlingen. We hebben gezien dat scholen niet vaak de gewoonte hebben om leerresultaten van *alle* leerlingen te analyseren.
- Onterechte conclusies uit analyses trekken. Het is opvallend dat de meeste scholen bij hun analyses gebruik maken van de vrij grove indeling in kwartielen (niveaugroepen A t/m D/E) uit het Cito leerlingvolgsysteem. Voor evaluatie van vorderingen van leerlingen is dat eigenlijk te beperkt; analyses met vaardigheidsscores zijn geschikter voor het volgen van ontwikkelingscurves, zowel op individueel niveau als op groepsniveau. Het werken met de kwartielen suggereert soms ontwikkelingssprongen die er niet zijn en veronachtzaamt soms vorderingen die er wel zijn. Deskundigheidsbevordering op dit onderwerp is zeer gewenst.

Over doelen stellen, referentieniveaus en leerstandaarden

Het beleidsvoornemen is om binnen afzienbare termijn referentieniveaus vast te stellen voor het primair onderwijs, voortgezet onderwijs en tertiair onderwijs. De verwachting is dat dit scholen duidelijkheid biedt over wat ze moeten bereiken en dat het leidt tot prestatiebewaking en, waar nodig, prestatieverbetering. De wens om dit te doen is niet nieuw: de Onderwijsraad pleit er al geruime tijd voor (Onderwijsraad, 1999, 2009). De Onderwijsraad kiest daarbij voor een indeling in drie niveaus: minimum, voldoende en gevorderd. De Expertgroep Doorlopende leerlijnen Taal en Rekenen heeft in haar onlangs uitgebrachte nota *Over de drempels met taal en rekenen* (2008) gekozen voor twee niveaus: fundamenteel niveau en streefniveau. De PO-raad heeft zich hierbij aangesloten, maar heeft wel voorgesteld om eerst nog nader onderzoek te doen naar de consequenties van een dergelijke indeling voor het onderwijs aan zorgleerlingen (PO-Raad, 2009).

De scholen en leerkrachten die wij in dit onderzoek raadpleegden over de eventuele invoering van referentieniveaus staan daar niet negatief, maar nog wel gereserveerd tegenover. Men pleit in ieder geval voor voldoende differentiatie, om recht te doen aan het feit dat scholen verschillen in leerlingenpubliek (aandeel kansarme en kansrijke leerlingen) en het dus voor de ene school moeilijker zal zijn dan voor de andere om referentieniveaus te halen. In het licht van de vraag of referentieniveaus er toe zullen bijdragen dat leerprestaties verbeteren is het goed om op te merken dat bij onvoldoende differentiatie alleen scholen met veel achterstandsleerlingen voor de opgave zullen komen te staan om de prestaties te verhogen, omdat zij de gestelde standaarden nog niet halen. Scholen met een kansrijker leerlingenpubliek zullen niet geprikkeld worden tot prestatieverbetering, omdat zij die standaarden al wel halen. Dat betekent niet alleen dat er sprake is van een eenzijdig appel op een bepaalde groep scholen, maar ook dat er weinig kans is dat er een algehele prestatieverbetering optreedt: een groot deel van de scholen zal immers niet geneigd zijn om aan prestatieverbetering te werken, omdat men de standaard al haalt en geen tekorten in de prestaties ziet. Als prestatieverbetering in de breedte het doel is, is het verstandig om ook voor de scholen met een kansrijk leerlingenpubliek de lat hoog te leggen en voor hen passende referentieniveaus te formuleren.

Het invoeren van referentieniveaus zou wel kunnen bevorderen dat scholen zich meer bewust worden van het belang van doelen stellen. We hebben in het onderzoek gezien dat scholen wel vrijwel allemaal naar hun leerresultaten kijken, maar niet allemaal even doelgericht. Het evalueren op basis van vooraf zelfgestelde doelen (zoals verondersteld in de evaluatieve cyclus) lijkt nog weinig voor te komen. Dit geldt in het bijzonder voor het speciaal basisonderwijs, dat in dit onderzoek als weinig opbrengstgericht naar voren komt. De casestudies, waarin één school voor speciaal basisonderwijs is opgenomen, hebben laten zien dat op die school een vrij spectaculaire vooruitgang mogelijk was op het gebied van technisch lezen, nadat op dat domein concrete doelen waren geformuleerd. In deze sector is dus waarschijnlijk nog veel prestatiewinst te behalen.

Literatuur

- Amrein-Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, 37(2), 65-75.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, 86(1), 9-21.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5, 7-75.
- Black, P. & Wiliam, D. (2003). 'In praise of educational research': formative assessment. *British Educational research Journal*, 29, 623-637.
- Blok, H., & Breetvelt, I. (2002). *Adaptief onderwijs: betekenis en effectiviteit*. Amsterdam: SCO-Kohnstamm Instituut.
- Blok, H., Otter, M. E., & Roeleveld, J. (2001). *Het gebruik van leerlingvolgsystemen anno 2000*. Amsterdam: SCO-Kohnstamm Instituut.
- Blok, H., Otter, M. E., & Roeleveld, J. (2002). *Leerlingvolgsystemen: praktische lessen uit onderzoek*. Amsterdam: SCO-Kohnstamm Instituut.
- Blok, H., Slegers, P., & Karsten, S. (2006). *Schoolzelfevaluatie in het basisonderwijs; een samenvatting van het onderzoek rond Ziezo*. Amsterdam: SCO-Kohnstamm Instituut.
- Blok, H., Slegers, P., & Karsten, S. (2006) *Schoolzelfevaluatie in het basisonderwijs; een terugblik op de collegiale-visitatiefase van Ziezo*. Den Haag: Q Primair
- Coleman, A. (2007). Leaders as researchers. Supporting practitioners enquiry through the NCSL Research Associate Programme. *Educational Management Administration & Leadership*, 33 (4), 479-498
- Darling-Hammond, L. (2004). Standards, accountability, and school reform. *Teachers College Record*, 106, 1047-1085.
- De Groot, A. D. & Medendorp, F. L. (1986). *Term, begrip, theorie; inleiding tot de signifiante begripsanalyse*. Meppel/Amsterdam: Boom.

Opbrengstgericht werken

- De Wolf, I. F., & Janssens, F. J. G. (2007). Effects and side effects of inspections and accountability in education: an overview of empirical studies. *Oxford Review of Education*, 33, 379-396.
- Earl, L. & Fullan, M. (2003). Using data in leadership for learning. *Cambridge Journal of Education*, 33 (3), 383-394.
- Earl, L. & Katz, S. (2006). *Leading schools in a data-rich world. Harnessing data for school improvement*. Thousand Oaks, California: Corwin Press.
- Earl, L. & Katz, S. (2007). Leadership in networked learning communities: defining the terrain. *School Leadership and Management* 27 (3), 239-258.
- Emmelot, Y., Koopman, P., & Ledoux, G. *Voorstudie taalmonitor* Amsterdam. Amsterdam: SCO-Kohnstamm Instituut.
- Expertgroep Doorlopende leerlijnen Taal en Rekenen (2008). *Over de drempels met taal en rekenen*. Hoofdrapport. Enschede: Expertgroep Doorlopende leerlijnen Taal en Rekenen.
- Geijsel F.P. & Krüger, M.L. (2005). Leren van Onderzoek: de benutting van data-feedback ten behoeve van schoolontwikkeling. *Pedagogische Studiën* 82 (4), 327-342.
- Geijsel, F.P., Slegers, P.J.C., Stoel, R.D., Krüger, M.L. (2009). The Effect of Teacher Psychological, School Organizational and Leadership Factors on Teachers' Professional Learning in Dutch Schools. *The Elementary School Journal* 109 (4).
- Fullan, M. (2002). The change leader. *Educational Leadership*, 59 (8), 16-20.
- Fullan, M. (2007). *The new meaning of educational change*. New York and London: Teacher College Press.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77, 81-112.
- Hoban, G. (2002). *Teacher learning for educational change*. Buckingham: Open University Press.
- Hofman, R.H., Dijkstra, N.J., Hofman, W.H.A., & Boom, J. de (2004). *Kwaliteitszorg in het primair onderwijs*. Deelstudie 1. Groningen: GION.
- Hofman, R.H., Dijkstra, N.J., Hofman, W.H.A., & Boom, J. de (2005). *Kwaliteitszorg in het primair onderwijs*. Eindrapport. Groningen: GION.
- Hofman, R.H., Vandenberghe, R., & Dijkstra, B.J. (2008). *Kwaliteitszorg, innovatie en schoolontwikkeling*. BOPO Review. Groningen: GION.

- Hoffmanns, C. (2007). Kwaliteitszorg: sturen met verstand. In: *Laat getallen spreken. Kengetallen over continuïteit, kwaliteit en schoolontwikkeling*. Utrecht: Projectbureau lumpsum primair onderwijs.
- Ingram, D., Louis, K. S., & Schroeder, R. G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, 106, 1258-1287.
- Ingram, D., Louis, K. S. & Schroeder, R.G. (2004). Accountability policies and teacher decision making: Barriers to the use of data to improve practice. *Teachers College Record*, 106, 1258-1287.
- Inspectie van het Onderwijs (2003). *De kern van de zaak. Werken aan kwaliteitszorg in het primair onderwijs*. Utrecht: Inspectie van het Onderwijs
- Inspectie van het Onderwijs (2006). *De staat van het onderwijs. Onderwijsverslag 2004-2005*. Utrecht: Inspectie van het Onderwijs
- Inspectie van het Onderwijs (2006). *Eindtoets in het basisonderwijs; een onderzoek naar leerlingen die niet meedoen en/of niet mee tellen*. Utrecht: Inspectie van het onderwijs.
- Inspectie van het Onderwijs (2007). *De staat van het onderwijs. Onderwijsverslag 2005-2006*. Utrecht: Inspectie van het Onderwijs
- Inspectie van het Onderwijs (2008). *Basisvaardigheden taal in het basisonderwijs; een onderzoek naar het niveau van de taalvaardigheden in het basisonderwijs en naar verschillen tussen scholen met lage, gemiddelde en goede taalresultaten*. Utrecht: Inspectie van het onderwijs.
- Janssens, F. J. G. (2008). *Een andere kijk op toetsen* (bijlage bij brief PO/Kwaliteit/3518 van Staatssecretaris mevr. Dijkzema aan de Tweede Kamer, d.d. 24 april 2008). [<http://www.minocw.nl/documenten/3518c.pdf>]
- Karsten, S., De Jong, U., Ledoux, G., & Sligte, H. (2006). *De positie van ouders en leerlingen in het governancebeleid*. Amsterdam: SCO-Kohnstamm Instituut.
- Kornhaber, M. L. (2004). Appropriate and inappropriate forms of testing, assessment, and accountability. *Educational Policy*, 18, 45-70.
- Ledoux, G., Overmaat, M., Boogaard, M., Felix, C., & Triesscheijn, B. (2005a). *Onderwijskansen bekeken. De stand van zaken in het onderwijskansenbeleid*. Amsterdam: SCO-Kohnstamm Instituut.

Opbrenstgericht werken

- Ledoux, G., Boogaard, M., & Overmaat, M., m.m.v. Derriks, M. (2005b). *Meer dan een schip met duiten. Eindevaluatie Project Onderwijskansen Rotterdam*. Amsterdam: SCO-Kohnstamm Instituut.
- Ledoux, G., Boogaard, M., Veen, A., & Bremer, F. (2005c). *Evalueren van onderwijsverbetering op scholen. Handreikingen bij het kiezen van instrumenten en het inrichten van een schoolspecifieke monitor*. Rotterdam/Amsterdam: Dienst Stedelijk Onderwijs Gemeente Rotterdam/ SCO-Kohnstamm Instituut.
- Leithwood, K., Aitken, R., & Jantzi, D. (2001). *Making schools smarter: A system for monitoring school and district progress. Second Edition*. Thousand Oaks: Corwin Press.
- McGill-Frantzen, A., & Allington, R. (2006). Contamination of current accountability systems. *Phi Delta Kappan*, 87(10), 762-766.
- Mehrens, W. A. (2002). Consequences of assessment: What is the evidence? In G. Tindal (Ed.), *Large-scale assessment programs for all students: validity, technical adequacy, and implementation* (p. 149-177). Mahwah, N.J.: Erlbaum.
- Meijer, A., Hasselaar, E., & Snepvangers, C. A. M. (2007). *Literatuurstudie scholen en kindercentra: binnenmilieu, gezondheid en leerprestaties*. Delft: Technische Universiteit Delft.
- Ministerie van OCW (2007). *Scholen voor morgen. Samen op weg naar duurzame kwaliteit in het primair onderwijs* Den Haag: Ministerie van OCW.
- Ministerie van OCW en PO Platform Kwaliteit en Innovatie (2006). *Innovatie-agenda Primair Onderwijs*. www.minocw.nl.
- Mulder, L., Roeleveld, J., Veen, I. van der & Vierke, H. (2005). *Onderwijsachterstanden tussen 1988 en 2002: ontwikkelingen in basis- en voortgezet onderwijs*. Nijmegen/Amsterdam: ITS/SCO-Kohnstamm Instituut.
- Onderwijsraad (1999). *Zeker Weten. Leerstandaarden als basis voor toegankelijkheid*. Den Haag: Onderwijsraad
- Onderwijsraad (2006). *Naar meer evidence based onderwijs*. Den Haag: Onderwijsraad
- Onderwijsraad (2007). *Presteren naar vermogen*. Den Haag: Onderwijsraad
- Onderwijsraad (2009). *Partners in onderwijsopbrengst*. Den Haag: Onderwijsraad.

- Ponte (2002). *Actie-onderzoek door docenten: uitvoering en begeleiding in theorie en praktijk*. Leuven / Apeldoorn, Garant.
- PO-Raad (2009). *Advies aan staatssecretaris Dijkma over referentieniveaus taal en rekenen*. www.poraad.nl.
- Putnam, R.T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning. *Educational Researcher*, 29, 4-15.
- Rennie Center (2006). *Data-driven teaching: Tools and trends* (Policy brief). Cambridge, MA: Rennie Center for Education Research and Policy.
- Scheerens, J. (2008). *Een overzichtsstudie naar school- en instructie-effectiviteit*. Enschede: Universiteit Twente.
- Schoot, F. van der (2008). *Onderwijs op peil? Een samenvattend overzicht van 20 jaar PPO*. Arnhem: Cito.
- Schoot, F. van der (2008). Referentieniveaus medio jaargroep 5; een pilotstudie in opdracht van het Ministerie van OCenW onder basisscholen in Arnhem en Den Haag. Arnhem: Cito.
- Shepard, L. A. (2002). The hazards of high-stakes testing. *Issues in Science and Technology*, 19(2), 53-58.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78, 153-189.
- Sleegers, P., & Ledoux, G. (2006). *Innovatie in het primair onderwijs: strategieën, ervaringen en aanbevelingen*. Een literatuurstudie naar werkzame principes. Amsterdam: SCO-Kohnstamm Instituut.
- Smylie, M.A., & Hart, A.W. (1999). School leadership for teacher learning and change: A human and social capital development perspective. In J. Murphy, and K.S. Louis (Eds.), *Handbook of research on educational administration*, 2nd ed. (pp. 297-322). San Francisco: Jossey-Bass.
- Stoll, L., Bolam, R., & Collarbone, P. (2002). Leading for change: Building capacity for learning. In K. Leithwood and Ph. Hallinger (Eds.), *Second international handbook of educational leadership and administration, part I* (pp. 41-73). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Townsend, T. (ed) (2007). *The international handbook of school effectiveness and school improvement. Review, reflection and reframing*. Dordrecht: Kluwer Academic Publishers.

Opbrengstgericht werken

- Van Eck, E., & Boogaard, M. (red.) (2007). *Interne en externe kwaliteitszorg in het basisonderwijs en de rol van het bovenschools management*. Amsterdam: SCO-Kohnstamm Instituut.
- Van Petegem, P., & Vanhoof, J. (2004). Feedback over schoolprestatie-indicatoren als strategisch instrument voor schoolontwikkeling? Lessen uit twee Vlaamse cases. *Pedagogische Studiën*, 81 (5), 338-353
- Van der Veen, I., Van der Meijden, A., & Ledoux, G. (2006). *School- en klaskenmerken Basisonderwijs: basisrapportage Prima-cohortonderzoek, zesde meting 2004-2005*. Amsterdam: SCO-Kohnstamm Instituut.
- Visscher, A. (2001). Publieke schoolprestatie-indicatoren: de problemen op een rij. In A. B. Dijkstra, S. Karsten, R. Veenstra & A. J. Visscher (Red.), *Het oog der natie: scholen op rapport; standaarden voor de publicatie van schoolprestaties* (p. 54-61). Assen: Koninklijke Van Gorcum.
- Visscher, A.J., Witziers, B. & Scheerens, J. (2007). *Omvang en oorzaken van bureaucrativering in het onderwijs. Case studies in het basisonderwijs en indicatoren voor OECD landen*. Twente: Faculteit Gedragwetenschappen, Afdeling onderwijsorganisatie en –management.
- Wang, L., Beckett, G. H., & Brown, L. (2006). Controversies of standardized assessment in school accountability reform: A critical synthesis of multidisciplinary research evidence. *Applies Measurement in Education*, 19, 305-328.
- William, D. (2008). Changing classroom practice. *Educational Leadership*, 65(4), 36-42.
- Verbiest, E. (2003). De rol van de schoolleider in een professionele leergemeenschap. In J.W. Heijmans & B.J.R. Redder (red.), *De schoolleider 'Meesterlijk beschreven'* (pp.64-68). NSA themareeks Leidinggeven in het primair onderwijs nummer 1.

Bijlage

Geïnterviewde deskundigen Meetgestuurd onderwijs

- Channa Woudenberg, Beleidsmedewerker monitoring en evaluatie, Stichting PAS, Arnhem
- Ruud Janssen, Schoolbegeleider CED (onder andere in het Delta+ project)
- Els Kolvoort, Schoolbegeleider CED (onder andere in het Delta+ project)
- Gea Spaans, coördinator pilots Taalbeleid WSNS
- Spier ten Doesschate, Adviseur.consultant APS
- Emile Eshuis + Marcel Boogaerts, RPCZ/Bazalt, coördinatoren Onderwijskansen Zeeland
- Mirjam Stefels, coördinator Montessori-opleiding EHvA, directeur Montessorivereniging

Geïnterviewde PABO's

- IN-Holland Alkmaar
- Groenewoud Arnhem (2x)
- Iselinge Hogeschool Doetinchem
- Driestar Educatief Gouda
- Hogeschool Domstad Utrecht (3x)
- Fontys-PABO Eindhoven
- Hogeschool Zuyd Heerlen/Maastricht.

Casestudiescholen

- SWS Klaver-Heijplaat, Rotterdam
- Obs de Esmarke
- Sbo de Tender in Enschede
- Kruispunt Santenkraam, Amsterdam
- Willibrordschool, Terneuzen

Recent uitgegeven SCO rapporten

- 814 Boogaard, M., Derriks, M.
Waarden en normen in interculturele context
- 813 Schooten, E. van, Slegers, P.
Onderzoek naar de effectiviteit van VVE- en peuterspeelzalen in Oosterhout en Den Bosch
- 810 Derriks, M., Kat, E. de
Maatjesaanpak in een Jenaplanschool
- 809 Jong, U. de, Ledoux, G., Emmelot, Y., Roeleveld, J.
Opleiden in de school
- 808 Daalen, M. van, Vrieze, G., Karsten, S., Kessel, N. van
Onbevoegd lesgeven: een noodoplossing.
- 807 Emmelot, Y., Ledoux, G., Veen, I. van der. m.m.v. Breetvelt, I.
Innovatiemonitor Primair Onderwijs
- 805 Fukkink, R.
Evaluatie van het SHare In Trust-project.
- 804 Peetsma, T., Veen, I. van der
Een tweede onderzoek naar de beïnvloeding van motivatie bij vmbo-leerlingen.
- 803 Veen, A., Roeleveld, J., Daalen, M. van
Implementatie en effecten van Voor- en Vroegschoolse Educatie in Rotterdam. Samenvatting.
- 802 Blok, H., Karsten, S.
Vervangend onderwijs aan kinderen met een richtingbezwaaar.
- 801 Karstanje, P.N., Glaude, M.T., Ledoux, G., Verbeek, A.E.
Beleidsvoerend vermogen van basisscholen: stand van zaken en verklarende factoren.
- 799 Derriks, M., Kat, E. de, Voncken, E.
Iedereen bij de les!
- 797 Roede, E., Derriks, M., Boogaard, M.
Sociale Competentie in het Rotterdamse Onderwijs. Eindrapport.
- 796 Veen, I. van der,
Amsterdamse Schakelklassen in het schooljaar 2006/2007.
- 795 Voncken, E., Derriks, M., Ledoux, G.
Vernieuwen zonder blauwdruk, de plussen en minnen.
- 794 Voncken, E., Derriks, M., Ledoux, G., Slegers, P., Kock, J. de
Een hele toer. Ervaringen van schoolleiders en docenten met de vernieuwing van de onderbouw VO.

Deze rapporten zijn te bestellen via:

<http://www.sco-kohnstaminstituut.uva.nl/webwinkel/bestellen.htm>