

University of Groningen

Het streven naar kwaliteit in scholen voor primair onderwijs

Doolaard, S.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2013

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Doolaard, S. (2013). Het streven naar kwaliteit in scholen voor primair onderwijs. Groningen: s.n.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

G I O N

Gronings Instituut voor Onderzoek van Onderwijs
Grote Rozenstraat 3 9712 TG Groningen



rijksuniversiteit
 groningen / gion

Het streven naar kwaliteit in scholen voor primair onderwijs

Dr. Simone Doolgaard

Het streven naar kwaliteit in scholen voor primair onderwijs

Dr. Simone Doolgaard

Het streven naar kwaliteit in scholen voor primair onderwijs

Integraal eindrapport van onderzoeken in de BOPO-programmalijn
Onderwijskwaliteit PO 2009-2012 (413-09-060)

Dr. Simone Doolaard

GION/RUG

Februari 2013

ISBN 978-90-6690-520-7

2013. GION, Gronings Instituut voor Onderzoek van het Onderwijs

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook zonder voorafgaande schriftelijke toestemming van de directeur van het instituut.

No part of this book may be reproduced in any form, by print, foto print, microfilm or any other means without written permission of the director of the instituut.

Inhoud

Managementsamenvatting	1
1 Inleiding	5
2 Achtergrond.....	7
2.1 Schets van de beleidsachtergrond	7
2.2 Theoretische en empirische achtergrond	14
3 Opzet van Streef	19
3.1 Inleiding	19
3.2 Werving en respons	20
3.3 De deelprojecten	22
3.3.1 Deelproject 1 – Werken met referentieniveaus	22
3.3.2 Deelproject 2 – Gebruik maken van opbrengsten	23
3.3.3 Deelproject 3 – Differentiatie en excellentie.....	24
3.4 Dataverzameling	25
3.4.1 Dataverzameling bij leerkrachten.....	25
3.4.2 Dataverzameling bij leerlingen	26
3.5 Opzet analyses	26
4 Resultaten	29
4.1 Ontwikkeling bij schoolteams.....	29
4.2 Werken met referentieniveaus	31
4.3 Gebruik maken van opbrengsten	32
4.4 Differentiatie en excellentie.....	33
5 Integratie en discussie	37
Literatuur	43
Bijlage.....	45

Managementsamenvatting

Het beleid van de landelijke overheid gericht op het verbeteren van de kwaliteit van het primair onderwijs uitgedrukt in prestaties van leerlingen richt zich met name op drie speerpunten: het invoeren van referentieniveaus zodat leerkrachten een duidelijk doel hebben om naar toe te werken, het stimuleren van opbrengstgericht werken en het stimuleren van onderwijs aan excellente leerlingen. Om te onderzoeken hoe leerkrachten daar mee omgaan en of en hoe zij en hun leerlingen daarvan profiteren heeft het GION/RUG en Cito een samenhangend onderzoeksprogramma opgezet dat op de deelnemende scholen is uitgevoerd als schoolverbeteringsprojecten - onder de naam *Streef* - met daaraan gekoppeld dataverzameling bij leerkrachten en leerlingen om achteraf de werking van de interventies vast te stellen. In de deelprojecten – *Werken met referentieniveaus*, *Gebruik maken van opbrengsten* en *Differentiatie en excellentie* – werden leerkrachten getraind in en begeleid op de genoemde thema's.

Opzet van Streef

De oorspronkelijke opzet waarin scholen (de onder- of de bovenbouw) random aan een deelproject toegewezen zouden worden als zij mee wilden doen aan Streef is tijdens de werving losgelaten omdat scholen zelf wilden kiezen aan welk deelproject ze mee zouden doen en omdat interne motivatie en actieve participatie noodzakelijk waren om effecten vast te kunnen stellen. Wel is grotendeels vastgehouden aan het gekruiste design waarin elke school met één bouw aan een deelproject en het onderzoek meedeed en de andere bouw alleen aan het onderzoek, waardoor deze bouw tot de controlegroep behoorde voor de andere deelprojecten. In het onderzoek zijn via een vragenlijst vooraf en na afloop gegevens verzameld bij leerkrachten en directies, van leerlingen zijn vijfmaal toets- en achtergrondgegevens verzameld. In het totaal hebben aan Streef 120 scholen deelgenomen, 20 scholen aan *Werken met referentieniveaus*, 38 scholen aan *Gebruik maken van opbrengsten* en 9 scholen aan *Differentiatie en excellentie*. Dit laatste deelproject had het karakter van een casestudie.

Resultaten

Uit analyse van de vragenlijsten blijkt dat deelname aan Streef leidt tot meer gebruik van toetsresultaten en reflectie daarop, en tot een positievere houding ten aanzien van opbrengstgericht werken en de mogelijkheden om prestaties van leerlingen te verbeteren. Leerdoelen blijken nog vrij vaak (door de helft van de leerkrachten) individueel vastgesteld te worden. Samenwerken in het team blijft meestal beperkt tot het uitwisselen van lesideeën. Differentiëren gebeurt nog maar zeer beperkt en in teams waarbij wel gedifferentieerd wordt zijn de opbrengsten over het algemeen hoger.

Het deelproject *Werken met referentieniveaus* laat zien dat leerkrachten slecht op de hoogte zijn van de hoogte en inhoud van referentieniveaus hoewel ze wel bijna allemaal gehoord hebben van de niveaus. Met name het hoogste niveau stellen zij desgevraagd veel hoger dan wettelijk vastgesteld. Deelname aan dit deelproject leidt tot meer begrip van de (implicaties van de) niveaus en meer realistische verwachtingen van leerlingen. De verwachtingen van de niveaus in het algemeen zijn gedaald. Deelname aan dit deelproject had geen effect op leerlingprestaties.

Resultaten van *Gebruik maken van opbrengsten* laten zien dat leerkrachten hun analysevaardigheden behoorlijk kunnen verbeteren en daar ook beter op kunnen reflecteren. Hoewel de prestaties van leerlingen in de breedte een verschuiving in de positieve richting laten zien is het effect (net) niet significant en dooft het in het jaar na deelname weer uit. Scholen geven zelf aan een stok achter de deur nodig te hebben.

Differentiatie en excellentie geeft een gedetailleerd beeld te zien van de manier waarom scholen nu met excellente leerlingen omgaan en de gebieden waarop verbetering mogelijk is. De urgentie van het thema is op alle scholen duidelijk. Allen doen zij op dit gebied al iets, maar de feitelijke situatie is minder positief dan blijkt uit mondelinge of schriftelijke rapportage. Knelpunten zijn met name een aantal vooroordelen (moeilijker werk is een straf, eigen keuze is goed en goede leerlingen kunnen wel zelfstandig werken), de mogelijkheid te differentiëren in klassikale situaties, de eigen beheersing van de lesstof, visie/doelstelling/evaluatie, de benodigde voorbereidingstijd, de interpretatie van de reacties en het gedrag van leerlingen op basis van kennis over de ontwikkeling van leerlingen. In de genoemde knelpunten is in een scholjaar tijd verbetering te boeken bij leerkrachten, met effecten op gedrag en attitude van leerlingen.

Uit de resultaten van de afzonderlijke deelprojecten blijken dus positieve effecten op de vaardigheden, de attitude en het handelen van leerkrachten en een positieve tendens bij leerlingen, hoewel die niet lang aanhoudt. De effecten zijn het grootst op de factoren die het dichtst bij de inhoud van de interventies liggen.

Oorzaken voor het ontbreken van significant positieve effecten op leerlingen liggen waarschijnlijk zowel in de opzet en de intensiteit van de interventie als in de inhoud. De opzet van *Gebruik maken van opbrengsten* voldoet aan de meeste uit de literatuur bekende kenmerken van effectieve interventies. Voor *Werken met referentieniveaus* geldt dat in mindere mate. Minder sterke punten van *Gebruik maken van opbrengsten* waren de mate waarin er aandacht is besteed aan de inbedding in het team, de sturing door de directie en de borging in de toekomst. Wat betreft de inhoud ligt het werken met referentieniveaus verder af van het dagelijks werk in de klas dan opbrengstgericht werken. Uit veel effectiviteitsonderzoek blijkt dat factoren op leerkracht- en klasniveau een grotere rol spelen dan factoren op school of stelselniveau. In *Werken met referentieniveaus* is wel geprobeerd om juist de abstracte

referentieniveaus te ‘vertalen’ naar de klassensituatie, maar het onderzoek laat zien dat dat moeilijk is; leerkrachten hebben weinig idee van de hoogte en de inhoud van de niveaus.

Beleidsrelevantie en evaluatie

Over het concept Opbrengstgericht werken en de aandacht voor excellente, goede leerlingen is men over het algemeen positief. Ook van de referentieniveaus verwachtte men in eerste instantie veel, maar dat lijkt wat te temperen. Er is weinig idee van de inhoud, de hoogte, de doelstellingen en de consequenties van (het vaststellen van) de niveaus. Een (hernieuwde) koppeling aan opbrengstgericht werken (expliciteren referentieniveaus in leerlingvolgsystemen, eindtoets, toezichtskader, methoden) zou meer duidelijkheid kunnen geven over de referentieniveaus. Opvallend is dat leerkrachten met name zeer uiteenlopend denken over wat een goede leerling zou moeten/kunnen behalen. Voor het onderwijs aan deze leerlingen bieden de huidige referentieniveaus weinig houvast.

Gemeenschappelijk in de deelprojecten is de aanbeveling om benodigde kennis voor opbrengstgericht werken in de Pabo's aan te bieden: analyse-vaardigheden en kennis over op onderzoek gebaseerd differentiatiemogelijkheden. Het uitvoeren van verantwoord gedifferentieerd onderwijs kan niet verwacht worden van een pas afgestudeerde leerkrachten. Van alle leerkrachten blijkt volgens inspectie-onderzoek maar de helft over de benodigde vaardigheden te beschikken. Onderzoek naar de ontwikkeling van leerkrachten laat ook zien dat vragen zoals ‘wat heeft een leerling nodig, wat leren leerlingen eigenlijk van mijn les’ pas na verloop van tijd door de (beginnende) leerkracht gesteld worden. Maar wanneer en hoe legt een gevorderde leerkracht daar dan een ‘proeve van bekwaamheid’ voor af?

Zorgwekkend is dat in de deelprojecten bleek dat er nog steeds (ervaren) leerkrachten zijn die tekort schieten op elementaire kennis en vaardigheden die te maken hebben met klassenmanagement, (vak)didactische vaardigheden en ontwikkelingspsychologie. Besturen en schoolleiding zouden daarop nog meer hun verantwoordelijk moeten nemen. Goede leerlingen die een paar jaar achter elkaar niet her/erkend zijn lopen het risico onderpresteerders te worden en daarna nooit meer aangesproken te worden op hun capaciteiten.

Tenslotte blijft (kosten)effectiviteitsonderzoek naar de effecten van aanpak/methoden op het gebied van excellente leerlingen noodzakelijk. Van reguliere methoden kunnen scholen beschikken over gegevens over de mogelijke resultaten en inzichten in welke methode geschikt is voor welke doelgroep. Voor excellente/goede leerlingen zijn deze gegevens zeer schaars, mede omdat onduidelijk is welke doelen de methoden/het materiaal nastreven waardoor het voor scholen heel moeilijk is om voor deze leerlingen een doorgaande leerlijn te realiseren en verantwoorde keuzes te maken.

1 Inleiding

In ‘Scholen voor Morgen’, de kwaliteitsagenda voor het primair onderwijs, werden in 2007 de ambities van het overheidsbeleid ten aanzien van het primair onderwijs onder andere verwoord als: (1) “In 2009 is duidelijk vastgelegd wat leerlingen aan het eind van het primair onderwijs op het terrein van taal en rekenen moeten kennen en kunnen” en (2) ‘In 2011 zijn de gemiddelde leerprestaties voor alle groepen leerlingen op taal en rekenen aantoonbaar gestegen ten opzichte van 2005’ (MOCW, 2007, p.4). Deze laatste doelstelling is later geconcretiseerd met behulp van experts en onderzoekers en wordt vanaf 2008 jaarlijks gevolgd. In 2011 zijn de ambities nog verder gespecificeerd en aangescherpt in ‘Basis voor presteren’ (MOCW, 2011).

Een reeks aan beleidsinitiatieven is in deze jaren genomen om de ambities te bereiken, waaronder het opstellen van referentieniveaus, de taal- en rekenpilots, subsidie voor projecten op het terrein van onderwijs aan excellente leerlingen.

Een van de programmalijnen voor Beleidsgericht Onderzoek Primair Onderwijs is specifiek gericht op deze beleidsinitiatieven gericht op het verhogen van de kwaliteit van het primair onderwijs (NWO, 2009). De onderzoeken binnen deze programmalijn dienen gezamenlijk nader inzicht te verschaffen in hoe deze beleidsinitiatieven in de dagelijkse onderwijspraktijk worden vormgegeven en in hoeverre de verschillende vormgevingsvarianten verschillen in de mate van doeltreffendheid.

Het GION (Rijksuniversiteit Groningen) en Cito hebben een samenhangend onderzoeksprogramma opgezet, in samenwerking met twee regionale onderwijsadviesorganisaties (Cedin en Marant) waarin drie interventiestudies zijn uitgevoerd waarvan de effecten vastgesteld zijn op basis van gegevens van twee dataverzamelingsprojecten, een bij leerlingen en een bij leerkrachten en schoolleiders. Waarbij de laatst genoemde ook een aantal op zichzelf staande onderzoeksvragen heeft. De drie interventiestudies betreffen trainings- en begeleidingstrajecten voor teams van leerkrachten op het terrein van de drie beleidsspeerpunten; het werken met referentieniveaus, opbrengstgericht werken en excellentie. Er is gebruik gemaakt van een groep van 120 scholen die met behulp van een gekruist design zowel als (quasi-) experimentele als controleschool dienst deden. Werving en communicatie naar de scholen heeft plaatsgevonden onder de naam ‘Streef’. Behalve de drie interventiestudies die in het kader van het BOPO-programma zijn uitgevoerd bevatte Streef voor de scholen ook nog 2 projecten die in het kader van TIER door het GION zijn uitgevoerd en die inhoudelijk een sterke samenhang vertonen met de BOPO-onderzoeken. Deze deelprojecten worden in deze rapportage verder buiten beschouwing gelaten.

Voor het samenhangende onderzoeksprogramma zijn de volgende hoofdvragen geformuleerd:

1. Hoe kijken leerkrachten aan tegen de referentieniveaus? Kunnen ze ermee werken?
2. Welke professionaliseringsactiviteiten ondernemen (teams van) leerkrachten in het kader van een kwaliteitsverhoging? En wat zijn de effecten op de gecijferd- en geletterdheid van de leerlingen?
3. Hoe gaan leerkrachten om met de gevraagde extra nadruk op de basisvaardigheden taal/lezen en rekenen? Wat zijn de effecten hiervan op de gecijferd- en geletterdheid van de leerlingen? En wat zijn de eventuele neveneffecten voor andere domeinen?
4. Welke activiteiten ondernemen scholen om excellentie (niet alleen cognitief opgevat) bij leerlingen te bevorderen, en wat zijn de effecten hiervan op de ontwikkeling bij de betreffende leerlingen?

Deze hoofdvragen die enerzijds inventariserend, descriptief van aard zijn en anderzijds gericht zijn op effecten, zijn per deelproject verder uitgewerkt. De uitgewerkte vragen komen in hoofdstuk 3 aan de orde.

Van elk van de interventiestudies en het onderzoek onder leerkrachten en schoolleiders is een apart eindrapport verschenen. In dit integrale eindrapport zal de opzet van het gehele project Streef toegelicht worden en zullen de resultaten van de deelprojecten in samenhang worden beschreven, gevolgd door een aantal aanbevelingen. Van het programma Streef als geheel is tevens een praktijkboek verschenen waarin de opbrengsten van de interventiestudies dusdanig zijn beschreven dat ook andere scholen daarvan gebruik kunnen maken.

Hoofdstuk 2 geeft een beknopt overzicht van de theoretische en beleidsachtergrond van Streef. In hoofdstuk 3 wordt ingegaan op de opzet van Streef en in hoofdstuk 4 komen de resultaten van de afzonderlijke interventiestudies en het onderzoek onder leerkrachten en schoolleiders aan de orde. Hoofdstuk 5, ten slotte, bevat de conclusies, de integratie van de resultaten en de aanbevelingen.

2 Achtergrond

2.1 Schets van de beleidsachtergrond

In het Nederlandse onderwijsbeleid heeft lange tijd, met een korte onderbreking in de jaren 70, de autonomie van de school voorop gestaan. De overheid zorgde voor de randvoorwaarden en de scholen moesten zich achteraf verantwoorden. Lange tijd werden er echter weinig consequenties verbonden aan onvoldoende resultaten en ook het beleid werd weinig systematisch geëvalueerd. Eind jaren 90 begint er een kentering op te treden onder andere ingegeven door het feit dat er verschillen bestaan tussen scholen in de mate waarin zij de kerndoelen behandelden en daarmee in het eindniveau dat leerlingen behaalden, waardoor er voor die leerlingen geen ononderbroken ontwikkelingsgang door het onderwijs mogelijk was. De Onderwijsraad publiceerde in 1999 een advies waarin een pleidooi werd gehouden om leerstandaarden te formuleren voor wat leerlingen halverwege en aan het eind van het primair onderwijs zouden moeten kennen en kunnen. Daarbij werden leerstandaarden gedefinieerd als ‘een geoperationaliseerde en genormeerde omschrijving (...) van cruciale doelen die aan het eind van bepaalde onderwijsfasen beheerst moeten zijn. Cruciale doelen zijn (...) de doelen die voorwaardelijk zijn voor het verdere leer- en ontwikkelingsproces. Operationalisering van die doelen betekent dat wordt aangegeven hoe beheersing van die doelen kan blijken, en normering houdt in dat ook het gewenste beheersingsniveau gespecificeerd wordt - bijvoorbeeld in de vorm van een beheersingsscore die bij een toets gehaald moet worden, of van bepaalde criteria waaraan bij een opdracht moet worden voldaan’ (Onderwijsraad, 1999, p. 54). Er werd een minimum niveau geformuleerd, te bereiken door 90% van de leerlingen, en een voldoende niveau, te bereiken door driekwart van de leerlingen. In een later verschenen advies introduceerde de Onderwijsraad (2007) ook nog een niveau ‘gevorderd’ of ‘excellent’, bedoeld om de meer begaafde leerlingen extra te stimuleren.

Het formuleren van standaarden past goed binnen de aangepaste koers van de landelijke overheid waarin er meer aandacht is voor sturing, ook op processen, voor bepaalde kernvakken en voor evidence-based werken. Het aanpassen van de koers wordt ook gevoed door het aanhoudend debat over het niveau van het primair onderwijs, waarover de meningen overigens uiteenlopen.

De aangepaste koers wordt verwoord in de Kwaliteitsagenda's voor het onderwijs uit 2007, waarin voor het primair onderwijs doelen op procesniveau zijn geformuleerd en doelen op het niveau van leerlingprestaties.

In dat jaar wordt ook de Expertgroep Doorlopende Leerlijnen, in de wandelgangen de commissie Meijerink genaamd, geïnstalleerd. Deze kreeg als opdracht mee om te adviseren over de vraag wat leerlingen moeten kennen en kunnen op het gebied van taal en rekenen bij belangrijke overgangen in het onderwijs. Met andere woorden; om het voorstel van de

Onderwijsraad om leerstandaarden te formuleren nader in te vullen. Deze commissie duidde de betreffende beheersingsniveaus als referentieniveaus aan, en beperkte zich bij het primair onderwijs - conform de opdracht – tot het einde van deze periode, als leerlingen ongeveer 12 jaar oud zijn. De referentieniveaus werden daarbij omschreven als ‘beschrijvingen van kennis en vaardigheden die leerkrachten een houvast bieden voor het bepalen, volgen en stimuleren van de ontwikkeling van leerlingen’ (Expertgroep, 2008, p. 7). Daarbij werd een onderscheid gemaakt tussen een fundamenteel niveau (dat 75% van de leerlingen zou moeten beheersen) en een (hoger) streefniveau. Voor het streef niveau geldt dat bij taal de cesuur op basis van de gegevens die gebruik zijn voor de nulmeting door het Cito gebruik makend van het PPON-onderzoek in het schooljaar 2008-2009 is gelegd bij de score die 25% van de leerlingen haalt en voor rekenen voor de score die 50% van de leerlingen haalt.

Alhoewel het idee van de referentieniveaus op zich gedragen lijkt te worden in het onderwijsveld, is er ook enige scepsis die te maken heeft met het feit dat invoering van een dergelijk instrument ook tot een hardere afrekencultuur ten aanzien van de prestaties van de school als geheel kan leiden.

Het formuleren van het hogere streefniveau had enerzijds te maken met het extra stimuleren van de meer begaafde leerlingen. Want hoewel de meningen verschillen over de vraag of de prestaties van leerlingen gemiddeld dalen, is men het er wel over eens dat er in Nederland te weinig leerlingen excellent presteren. Internationale peilingen wijzen uit dat Nederlandse leerlingen in het topsegment relatief ondervertegenwoordigd zijn. Anderzijds vloeide dit voort uit de aandacht om ‘andere’ intelligenties bij leerlingen te stimuleren. Daarbij werd onder meer gedacht aan het – niet onomstreden werk van Gardner die een achttal intelligenties onderscheidde (bijvoorbeeld intra- en interpersoonlijke intelligentie, muzikaliteit). De kritiek op het werk van Gardner heeft onder meer betrekking op de gebrekkige empirische onderbouwing (Waterhouse, 2006).

Voor de volgende kabinetsperiode (kabinet Rutte I) wordt een nieuw beleidsdocument opgesteld, genaamd Actieplan Basis voor Presteren. Hierin wordt in grote lijnen het beleid voortgezet en wordt veelvuldig de term opbrengstgericht werken genoemd. Hoewel deze manier van werken al wel eerder gestimuleerd werd, werd die ten tijde van de Kwaliteitsagenda nog niet zo expliciet gebruikt. Opbrengstgericht werken is een voorbeeld van een onderwijsproceskenmerk. Dus naast de uitwerking van “wat” er op school geleerd moest worden, bepleitte de overheid ook steeds meer het “hoe” via cyclische evaluaties, collegiale intervisie en verdere professionalisering (Onderwijsraad, 2008a). In de Kwaliteitsagenda wordt het als volgt beschreven: ‘Om de taal- en rekenprestaties te verbeteren is het nodig dat leraren en schoolleiders de aanwezige informatie over leerprestaties in de school benutten en leren van en met collega’s en experts, zowel in als buiten de school, over de meest effectieve onderwijsmethoden’ (p. 5). Verwijzend naar eerder

onderzoek – onder andere van de Inspectie van het Onderwijs - wordt erop gewezen dat op basisscholen wel al een schat aan informatie aanwezig is maar dat dat nog onvoldoende gebruikt wordt. Men zou gestimuleerd moeten worden om vaardigheden te ontwikkelen om de uitkomsten te analyseren, te begrijpen en te vertalen naar concrete voorstellen voor verbetering. Deze vaardigheden worden dan nog niet ondergebracht onder de term ‘opbrengstgericht werken’ maar onder ‘kwaliteitszorg’. De doelstelling luidt dan ook: ‘In 2011 voldoet minimaal 80 procent van de scholen aan de standaarden van kwaliteitszorg zoals omschreven in het toezichtskader van de inspectie. Dat betekent dat de informatie over onderwijskwaliteit die daaruit voortvloeit, beter wordt benut door leraren. Dit houdt in het systematisch gebruiken van toetsen, de toetsdata interpreteren en gebruiken voor het concrete onderwijsaanbod aan alle leerlingen en als maatstaf: waar sta ik ten opzichte van mijzelf (vorig jaar, parallelklas, etc.), andere, vergelijkbare scholen en de landelijke referentieniveaus’ (p.11).

In de jaren daarna raakt de term ‘opbrengstgericht werken’ meer ingeburgerd. In de eerste voortgangsrapportage van 7 juli 2008 wordt het analyseren en evalueren van toetsen en het op basis daarvan aanpassen van het onderwijs ‘een belangrijk onderdeel’ van kwaliteitszorg genoemd. De ambitie die daarbij aansluit luidt dat in 2011 minimaal 70 procent van de scholen aan de indicatoren van goede kwaliteitszorg zal voldoen. In een voetnoot wordt gemeld dat overleg met de inspectie heeft geleid tot precisering en bijstelling van de eerder geformuleerde ambitie. Van de indicatoren van het bredere aspect kwaliteitszorg gaat het hierbij om twee specifieke indicatoren: het vergroten van de opbrengstgerichtheid en het evalueren van het onderwijsproces en de opbrengsten. In 2007 voldeed 39% van alle scholen aan beide indicatoren.

In de tweede voortgangsrapportage van 29 juni 2009 worden twee onderdelen genoemd die extra aandacht behoeven. Een daarvan is het percentage scholen dat aan beide specifieke kwaliteitszorgindicatoren voldoet; in het schooljaar 2007/2008 was dat weliswaar gestegen naar 52% maar daarmee lag het nog ver van de beoogde 70% af. Daarom is de inzet op dit terrein in de twee schooljaren daarna geïntensiveerd. Dit betekent dat er nog meer scholen betrokken worden in de lerende netwerken op het terrein van taal en rekenen (de taal- en rekenpilots), er komen financiële middelen voor verbetertrajecten op 50 nieuwe scholen, er komt meer aandacht voor het delen van effectief gebleken praktijkvoorbeelden van opbrengstgericht werken op een aantal regionale bijeenkomsten en meer experts om de scholen te ondersteunen. Ook zullen vanaf augustus 2010 de referentieniveaus wettelijk vastgesteld zijn en kunnen dienen als instrument om opbrengstgericht werken te realiseren.

Op 28 juni 2010 verschijnt de derde voortgangsrapportage. Opbrengstgericht werken wordt daarin ‘de rode draad’ door de kwaliteitsagenda genoemd. Ten aanzien van de doelstelling op het terrein van kwaliteitszorg c.q. opbrengstgericht werken is een lichte verbetering te zien.

Het percentage scholen dat voldoende scoort op de indicatoren voor kwaliteitszorg is 54% (Onderwijsverslag 2008/2009). Op elk van de 5 individuele indicatoren is een stijgende trend zichtbaar. De inspectie concludeert dat de voorwaarden aanwezig zijn maar dat het evalueren van de resultaten en het daarop aanpassen van het onderwijs een knelpunt blijft. Eventuele effecten van de geïntensiverde inzet waarover in de 2^e voortgangsrapportage wordt gesproken zijn in deze cijfers nog niet zichtbaar omdat het cijfers uit het schooljaar daarvoor betreft. In de opiniepeiling van Researchnet in 2010 en in een meta-analyse van het IVA uit 2010 zeggen respondenten wel vaker activiteiten te ondernemen die onder opbrengstgericht werken vallen.

Overigens wordt hier gesproken over kwaliteitszorg als belangrijk onderdeel van opbrengstgericht werken, waar opbrengstgericht werken in de eerste voortgangsrapportage in 2008 een belangrijk onderdeel van kwaliteitszorg was.

In 2011 wordt in het Actieplan Basis voor Presteren als streefdoel geformuleerd dat in het primair onderwijs in 2015 60% van de scholen opbrengstgericht werkt en in 2018 90%, volgens het oordeel van de inspectie waarbij er gerekend wordt met een nulmeting in 2010 waarbij 30% van de scholen opbrengstgericht werkte. In de begeleidende brief wordt aangegeven dat opbrengstgericht werken geen doel is maar een effectief middel om de prestaties naar een hoger plan te brengen.

In het voorjaar van 2012 blijkt uit een peiling van Regioplan (Dekker, Krooneman, Brekelmans & Groenwoud, 2012) naar de manieren waarop scholen inspelen op de in de referentieniveaus beschreven eisen voor taal en rekenen opnieuw een versterking van het opbrengstgerichte werken. Volgens eigen zeggen van de respondenten is er meer aandacht voor taal- en rekenzwakke leerlingen en wordt er meer in niveaugroepen gewerkt. Daarnaast is er draagvlak voor de referentieniveaus, het vaststellen van een minimumniveau en extra aandacht voor de basisvaardigheden.

Het laatste onderwijsverslag van de inspectie (2012a) is minder positief: sinds het vorige onderwijsverslag heeft er geen verbetering plaats gevonden in het percentage scholen dat voldoet aan de indicatoren voor opbrengstgericht werken. Het percentage scholen dat systematisch de resultaten van leerlingen evalueert is iets hoger geworden, maar in algemene zin is er geen verbetering zichtbaar.

In 2011 en 2012 verschijnen tenslotte ook de eindevaluaties van de taal- en rekenverbetertrajecten en de Op Maat-trajecten waarin opbrengstgericht werken een belangrijke rol speelden (Inspectie van het onderwijs, 2012b; van Schilt-Mol & van Vijfeijken, 2011; PO-Raad, 2012). Uit deze rapportages blijkt dat de resultaten van reken- en begrijpend lezen wel, maar van technisch lezen niet vooruit zijn gegaan. Wat betreft de onderwijsprocessen signaleren van Schilt-Mol en van Vijfeijken een cultuuromslag; er wordt meer belang gehecht aan het zelf uitvoeren van data-analyses en er is een dialoog op gang gekomen over de relatie tussen leerling-resultaten en leerkracht-handelen. Leerkrachten zijn

steeds meer bereid kritisch te kijken naar en te reflecteren op eigen handelen en op het handelen van collega's.

Uit dit chronologische overzicht is duidelijk op te maken dat er in de looptijd van de kwaliteitsagenda 'Scholen voor morgen' tot en met de start van het actieplan 'Basis voor presteren' een verschuiving is opgetreden in de definitie van opbrengstgericht werken en de manier waarop dit gemeten wordt. Niet alleen gebruiken verschillende organisaties en instellingen verschillende definities en operationalisaties, ook binnen dezelfde organisatie – zoals bijvoorbeeld de inspectie – zien we verschuivingen. Dat maakt het moeilijk om een duidelijk antwoord te krijgen op de vraag in welke mate er in het basisonderwijs opbrengstgericht gewerkt wordt en of er sprake is van vooruitgang. Duidelijk is wel dat opbrengstgericht werken gezien wordt als middel om de ambities op het gebied van leerlingprestaties te verbeteren en wordt er in het algemeen een manier van werken bedoeld waarin heldere doelen gesteld zijn en hoge verwachtingen, waarin vorderingen regelmatig in kaart worden gebracht en geanalyseerd en waarin vervolgens het onderwijs aangepast wordt aan de behoeften van leerlingen. Het betreft een cyclisch proces dat schoolbreed, door het hele team, wordt doorlopen en dat gericht is op het realiseren van zo hoog mogelijke prestaties door alle leerlingen. Opbrengstgericht werken is dus geen doel maar een middel, een manier van werken, volgens sommigen ook een attitude van leerkrachten en schoolleiders. Verschillende onderzoeken (ref) hebben aangetoond dat individuele leerkrachten lang niet allemaal voldoende vaardig zijn in het analyseren van toetsresultaten en daarvan dus ook geen gebruik maken bij het vormgeven van hun onderwijs, ook op veel scholen ontbreekt het aan een systematiek om gezamenlijk, in teamverband, resultaten te analyseren en te bespreken.

Weerstand en keerzijde

Opbrengstgericht werken wordt vaak gepresenteerd als de panacee, een wondermiddel tegen 'alle kwalen'. Het roept echter ook weerstand op. De weerstand concentreert zich op drie terreinen:

- Opbrengstgericht werken wordt geassocieerd met de basisvaardigheden. Aan taal en rekenen wordt zeer veel aandacht besteed, waardoor andere vakgebieden en andere vaardigheden, zoals sociaal-emotionele en burgerschapsvaardigheden te weinig aandacht krijgen.
- Opbrengstgericht werken wordt geassocieerd met (veel) toetsen. Met name voor jonge kinderen wordt dit niet altijd een geschikte werkwijze gevonden om oog te houden op de ontwikkeling van kinderen. Toetsen voor deze kinderen zouden te weinig zeggen omdat de ontwikkeling niet lineair is. Meer in het algemeen ageren betrokkenen tegen de monopoliepositie van het Cito, de instelling die het leeuwendeel van de toetsen levert en in de ogen van sommige betrokkenen daardoor te bepalend is in wat er getoetst en dus geleerd wordt.

- Opbrengricht werken wordt geassocieerd met het controleren van scholen. Scholen moeten de toetsgegevens aanleveren aan hun bestuur, de inspectie en mogelijk nog andere instanties en daarmee hun onderwijs verantwoorden. Het betreft hier toetsgegevens die ook gebruikt worden om het onderwijs te verbeteren. Dit dubbele perspectief wordt vaak als problematisch gezien.

Naast de formulering van referentieniveaus, het stimuleren van opbrengricht werken en de aandacht voor excellente leerlingen heeft de overheid nog meer aan de verbetering van de onderwijskwaliteit gerelateerde maatregelen genomen zoals het beperken van het aantal taal- en rekenzwakke scholen, het definiëren van bekwaamheidseisen voor schoolleiders en het verhogen van de effectieve leertijd en het financieren van reken- en taalpilots. Daarnaast heeft men het voornemen een leerlingvolgsysteem verplicht te stellen.

De doeltreffendheid van overheidsbeleid

Dit overzicht van het onderwijsbeleid op het gebied van de kwaliteit van het primair onderwijs is begonnen met de constatering dat na jaren waarin de overheid op afstand stond en veel invloed toebedacht was aan het onderwijsveld zelf, er nu (weer) een tendens is om meer te sturen op de inhoud en het onderwijsproces (hoewel de autonomie van de school nog wel hoog in het vaandel staat evenals de verantwoordingsplicht achteraf). Dit hernieuwt ook de vraag naar wat een effectieve beleidsinstrumenten zijn.

De (landelijke) overheid heeft verschillende soorten beleidsinstrumenten tot haar beschikking: juridische, (financieel-)economische en communicatieve instrumenten (Leune, 2001). Voorbeelden van juridisch instrumenten in het onderwijsbeleid zijn het uitvaardigen van wet- en regelgeving, maar ook het sluiten van overeenkomsten of convenanten met sectororganisaties of gemeenten, of het houden van toezicht op het onderwijs door de inspectie. Financiële instrumenten hebben betrekking op doeluitkeringen of subsidies om specifieke beleidsprioriteiten te stimuleren, of om via kortingen of boetes ongewenste activiteiten te ontmoedigen. Communicatieve instrumenten hebben betrekking op een moreel appèl, niet-bindende afspraken of voorlichtingsactiviteiten. Hiervoor wordt vaak de ondersteuningsinfrastructuur ingezet.

De kwaliteitsagenda en het actieplan bevat de maatregelen die de landelijke overheid heeft genomen om opbrengricht werken en excellentie bij scholen te stimuleren en de doelstellingen die daarmee behaald zouden moeten worden. Zo vermeldt het actieplan dat in 2018 90% procent van de PO-scholen opbrengricht zal werken. Om dit te bereiken worden door de overheid zowel juridische, financiële als communicatieve instrumenten ingezet (vgl. Onderwijsraad, 2007). Een voorbeeld van een juridisch instrument is het invoeren van de Wet referentieniveaus Nederlandse taal en rekenen per 1 augustus 2010, en de voorgenomen wijziging van de WPO met het oog op de invoering van een centrale eindtoets en leerling- en onderwijsvolgsysteem in het primair onderwijs.

Door middel van het programma ‘School aan Zet’ probeert de overheid daarnaast verbeterprocessen in de scholen te stimuleren en ondersteunen. Hoewel het hierbij voor een deel om een indirecte financiële prikkel gaat, is het programma voor een groot deel gebaseerd op het overtuigen van scholen om gebruik te maken van het programma, en de middelen in te zetten om de landelijk geformuleerde ambities te realiseren. Dat geldt ook voor de initiatieven die genomen zijn rond het onderwijs aan excellente leerlingen en het predicaat ‘excellente school’. Door scholen als excellent te beoordelen hoopt de overheid dat dit andere scholen zal stimuleren om eveneens de kwaliteit van hun onderwijs te verbeteren opdat zij ook voor het predicaat in aanmerking komen.

Hoewel de overheid gebruik maakt van een variëteit aan beleidsinstrumenten om de beleidsprioriteiten te realiseren, ligt het zwaartepunt daarin op communicatieve beleidsinstrumenten. Deze verschuiving is een direct gevolg van het beleid van deregulering en autonomievergroting dat de overheid sinds de jaren '90 voert (vgl. Onderwijsraad, 2007). Dit betekent dat scholen en schoolbesturen grote vrijheden kennen in de wijze waarop zij verbeteringsprocessen in school vormgeven.

Het onderzoek in deze programmalijs kan niet de doeltreffendheid van verschillende beleidsinstrumenten met elkaar vergelijken aangezien alle scholen in dezelfde mate onder dezelfde beleidsinstrumenten vallen. Bovendien zet, zoals aangegeven, de overheid een mix van beleidsinstrumenten in om doelen te bereiken en zijn de effecten daarvan niet van elkaar te onderscheiden. Het toeschrijven van veranderingen in attitude en handelen van leerkrachten en prestaties van leerlingen is niet ondubbelzinnig toe te schrijven aan beleid, wel is aannemelijk te maken dat verschuivingen toe te schrijven zijn aan bepaalde interventies (allen in de vorm van training en begeleiding van leerkrachten, zie hoofdstuk 3 voor de keuze m.b.t. opzet en uitvoering van de interventies) die binnen de onderzoeken zullen worden uitgevoerd om op die manier inzicht te krijgen in de mogelijkheid om binnen scholen verbeteringen te realiseren gang te zetten op het gebied van het kennen en kunnen gebruiken van referentieniveaus, opbrengstgericht werken en stimuleren van excellente leerlingen. Bovendien levert dergelijk onderzoek inzicht op in belemmerende en bevorderlijke randvoorwaarden voor de implementatie van onderwijsbeleid in de onderwijspraktijk.

De doeltreffendheid van interventies op schoolniveau

Behalve met de doeltreffendheid van overheidsbeleid in het algemeen speelt in deze programmalijs ook de doeltreffendheid van interventies in de school, van professionaliseringsactiviteiten voor leerkrachten een belangrijke rol. Voor het stimuleren van opbrengstgericht werken en excellentie worden aan scholen tal van projecten aangeboden waaraan zij kunnen deelnemen. Ook in deze programmalijs wordt door het uitvoeren van interventies op scholen geprobeerd inzicht te verwerven in of en hoe op scholen een verbetering bewerkstelligd kan worden.

Onderzoek naar effectieve interventies is met name uitgevoerd in de school improvement-traditie, interventies zijn immers altijd gericht op het verbeteren van het onderwijs. Een interventie is maar één onderdeel van een groter geheel aan activiteiten gericht op onderwijsverbetering. In het geval van opbrengstgericht werken bijvoorbeeld is deelname aan een interventie maar een van de manieren waarop een school te maken krijgt met opbrengstgericht werken. De landelijke overheid probeert ook via het Projectbureau Kwaliteit de scholen te bereiken, de landelijke media en de vakbladen besteden aandacht aan opbrengstgericht werken en de inspectie heeft de mate van opbrengstgerichtheid opgenomen in het toezichtkader. Deze aspecten kunnen de effectiviteit van de interventie beïnvloeden, maar dat ligt buiten het bereik van de interventie. Over het algemeen wordt aan een dergelijke ketenbenadering waarin betrokkenen op verschillende niveaus meewerken een positieve uitwerking toegeschreven (zie bijvoorbeeld Fullan, 2006). Het is dan ook aannemelijk dat zowel op de controlescholen als de interventiescholen vooruitgang optreedt, de vooruitgang zou op de interventiescholen groter moeten zijn dan op de controlescholen.

Wat betreft de effectiviteit van de interventie kan onderscheid gemaakt worden in het voortraject en de interventie zelf. Het voortraject mondt uit in de ‘beslissing’ van een school of een leerkracht om aan de interventie deel te nemen. Idealiter kiest een potentiële deelnemer er zelf voor om deel te nemen omdat hij/zij een probleem ervaart waarvan hij/zij vindt dat het opgelost moet worden. Daarbij heeft hij/zij het idee daar zelf iets aan te kunnen doen en het idee dat de interventie daar iets zal bijdragen.

Wat betreft de opzet en de inhoud van de interventie levert onderzoek op verschillende aspecten aanwijzingen voor belangrijke kenmerken:

- Aansluiting bij eigen ervaring en praktijk
- Aandacht voor weerstand, moeilijkheden
- Combineren en verbinden van kennis, vaardigheden en attitude
- Gebruik van activerende werkvormen
- Tijd en ruimte voor ‘uitproberen’
- Begeleid uitproberen in de eigen praktijk en nabespreking
- Uitwisseling met collega’s
- Inbedding in team/school, sturing door de directie
- Aandacht voor borging
- De duur van de interventie
- Deskundigheid van de trainer en begeleider

2.2 Theoretische en empirische achtergrond

De drie thema’s die een hoofdrol spelen binnen het onderwijsbeleid gericht op de onderwijskwaliteit in het primair onderwijs en binnen deze BOPO-onderzoekslijn –

opbrengstgericht werken, de werking van standaarden of referentieniveaus en excellentie hebben alle drie een van eigen onderzoekstraditie en theoretische achtergrond.

Hoewel de term *opbrengstgericht werken* pas kort wordt gebruikt is het achterliggende principe, de werking, van het concept wordt met name binnen het onderwijseffectiviteitsonderzoek al vele jaren onderzocht. Gerelateerde termen die daar gebruikt worden zijn dan bijvoorbeeld evaluerend vermogen, prestatiegerichtheid, hoge verwachtingen. Deze termen verwijzen naar de ‘smalle’ betekenis van opbrengstgericht werken waaronder gerekend worden: het regelmatig toetsen van leerlingen, het analyseren van de resultaten en het trekken van conclusies ten aanzien van het onderwijs aan (individuele) leerlingen. Ook Ledoux, Blok en Boogaard (2009) hanteren een smalle betekenis van opbrengstgericht werken, door hen meetgestuurd onderwijs genoemd, waaronder zij verstaan het vastleggen van doelen en standaarden, het verzamelen van informatie, het registreren van vorderingen, het interpreteren van vorderingen en het nemen van beslissingen.

Met alleen deze smalle inkleuringen van opbrengstgericht werken verbetert echter niet vanzelfsprekend de kwaliteit van het onderwijs en de prestaties van leerlingen. De smalle betekenis beslaat maar een deel van de cyclus van opbrengstgericht werken waarin ook het uitvoeren van (gedifferentieerd) onderwijs en het evalueren daarvan een rol spelen. Zo uitgedrukt is opbrengstgericht werken in te passen in de bekende plan-do-check-act-cirkel, waarin dus niet alleen geëvalueerd en gepland wordt maar ook uitgevoerd. De analysevaardigheden van leerkrachten worden het meest genoemd als hiaat in deze cyclus en vanwege het ontbreken daarvan worden ook vraagtekens geplaatst bij de mate waarin de afstemming van het onderwijs op verschillen tussen leerlingen effectief is (Inspectie van het Onderwijs, 2010).

Internationaal is er veel belangstelling voor het werken met *standaarden*. In veel landen zijn standaarden ingevoerd, veelal met als doel om het onderwijspeil te verhogen. De rationale daarachter komt overeen met de rationale achter opbrengstgericht werken: duidelijkheid over doelen, hoge doelen en verwachtingen zou prestaties van leerlingen verhogen – via het handelen van de leerkracht - en ongelijkheid opheffen omdat iedereen in ieder geval een bepaald minimumniveau behaalt. Ten slotte zou een toegenomen verantwoordingsplicht niveauverhogend kunnen werken. Er is inmiddels ook in meerdere landen en internationaal vergelijkend onderzoek gedaan naar de effecten van standaarden, waaruit meestal geconcludeerd moet worden dat standaarden niet de gewenste effecten hebben. Hoogstens leerlingen die al hoog scoorden presteerden nog iets beter, de zwak presterende leerlingen gingen niet vooruit. Daarvoor worden tal van oorzaken aangedragen:

- standaarden hebben betrekking op allerlei verschillende aspecten van het onderwijs: standaarden voor wat leerlingen aangeboden moeten krijgen en standaarden voor wat ze

moeten beheersen, maar ook standaarden voor de kwaliteit van het onderwijs uitgedrukt in bijvoorbeeld het percentage leerkrachten met een bepaalde opleiding;

- veel standaarden zijn heel vaag en algemeen waardoor ze voor leerkrachten weinig inzicht geven in wat precies bedoeld wordt;
- in veel landen zijn de standaarden zonder implementatiestrategie en toelichting aan het onderwijs opgelegd, waardoor leerkrachten de standaarden in het geheel niet kennen;
- leerkrachten hebben niet de kennis en vaardigheid om de standaarden om te zetten naar de onderwijspraktijk en hun instructie en aanbod aan leerlingen.

Daarnaast wordt door verschillende auteurs in twijfel getrokken of het schoolverbeteringsdoel en het verantwoordingsdoel van standaarden verenigbaar wel zijn.

Of deze onderzoeksresultaten ook van toepassing zijn op de Nederlandse referentieniveaus is vanwege de onvergelykbaarheid van verschillende typen standaarden (zie ook Onderwijsraad, 1999) de vraag. De Onderwijsraad pleit in 1999 voor concrete duidelijke standaarden, waarbij het belangrijkste doel niet is een kwaliteitsverhoging maar het garanderen van toegankelijkheid van het onderwijs, waarvoor dus in het bijzonder het vastleggen van een minimumniveau belangrijk is. In de beleidsdocumenten uit 2007 en 2011 blijkt echter wel dat de referentieniveaus tot doel zouden moeten hebben dat het prestatieniveau van leerlingen omhoog gaat. Ook zullen zij in Nederland uiteindelijk gebruikt gaan worden bij het afleggen van verantwoordelijkheid en is de implementatie een zaak van de school zelf.

Het stimuleren van *excellentie* en (zeer) goed presteren leerlingen heeft in Nederland lange tijd weinig aandacht gehad. Financiële middelen, tijd en aandacht ging naar de zwak presterende leerling, binnen het achterstandenbeleid naar leerlingen die op basis van sociaal-economische situatie of etnische achtergrond een achterstand hadden of, binnen het WSNS- en LGF-beleid naar leerlingen die dat op basis van een leer- of gedragsprobleem of beperking hadden. Pas de laatste jaren zien we de aandacht in onderzoek en theorievorming toenemen. Reviews van (inter)nationaal onderzoek levert niet veel eenduidige conclusies op. Onderzoek naar effecten van onderwijsaanpassingen voor (hoog)begaafde leerlingen is vaak descriptief en kwalitatief van aard. Kwantitatieve onderzoeken zijn in de jaren 90 in meerdere reviews en meta-analyses verzameld en naast elkaar gezet (Vaughn, Feldhusen & Asher, 1991; Renzulli & Reis, 1994; Olszewski Kubilius, 1995). Onderzoek van daarna is in 2008 door Hoogeveen verzameld voor een meta-analyse. Van de 62 experimentele studies die tussen 1993 en 2003 zijn gepubliceerd voldeden er slechts 23 aan de gestelde eisen op het gebied van o.a. design en participanten. Onder de 23 studies waren versnellingsprogramma's, verrijkingsprogramma's binnen de eigen groep, programma's buiten de eigen groep (vgl plusgroep), zomerprogramma's, groepen/scholen speciaal voor gifted students en scholing voor leerkrachten. Hoogeveen vindt net als in de eerdere reviews gemengde resultaten, zowel positieve als negatieve, verschillend voor de verschillende soorten programma's, verschillend voor leerlingen van verschillende leeftijden en verschillend voor verschillende uitkomstmaten.

In het algemeen concludeert zij wel dat onderwijsaanpassingen voor deze leerlingen effectief zijn en met name positief voor het cognitief functioneren. Effecten op niet-cognitieve uitkomsten is wisselend positief en negatief, maar op het zelfbeeld bijna altijd negatief. Dat dit niet altijd onwenselijk is beredeneert Hoogeveen vanuit het principe dat het zelfbeeld van hoogbegaafde leerlingen daalt omdat zij bij deze programma's geconfronteerd worden met werk waarvoor ze moeite moeten doen en met medestudenten die minstens net zo slim zijn als zij. De auteurs eindigen de conclusie met de opmerking dat er niet een beste aanpak voor alle gifted students is en dat elke school meerdere aanpakken zou moeten aanbieden en daarbij monitoren wat de opbrengsten daarvan zijn.

Een bijkomend nadeel van veel onderzoek naar gifted students is dat de andere leerlingen in de groep niet meegenomen worden, dus een mogelijk negatief effect op andere leerlingen van het feit dat de leerkracht meer aandacht moet besteden aan de gifted students wordt niet gemeten. Mogelijk zou extra aandacht voor gifted students ook een positief effect kunnen hebben voor de andere leerlingen, maar ook dat weten we niet. In het algemeen is wel bekend dat homogeen groeperen, wat noodzakelijk is om gifted students met gelijkgestemden aan het werk te laten, wat vaak voorkomt bij verrijkingsmateriaal of bij speciale groepen, positieve effecten heeft voor de sterke leerlingen maar negatieve voor de zwakkere leerlingen. Heterogeen groeperen heeft positieve effecten voor zwakkere leerlingen en geen effecten voor sterkere leerlingen wat een dilemma creëert voor Nederlandse basisschoolleerkrachten die een variatie aan leerlingen in de groep heeft en op zoek moet naar werkwijzen die voor alle leerlingen goed werken.

3 Opzet van Streef

3.1 Inleiding

Zoals in hoofdstuk 1 is aangegeven is het samenhangende onderzoeksprogramma aan de scholen gepresenteerd als Streef, schoolverbeterings- en onderzoeksproject waarin 3 verschillende deelprojecten, interventies aan werden geboden waar scholen met een bouw aan mee konden doen, waarbij ze ook aan de dataverzameling bij de leerlingen en het team mee moesten doen.

In de oorspronkelijke onderzoeksvoorstel is uitgegaan van een groep van 82 scholen, waarbij de omvang van de groep gebaseerd was op de te verwachte grootte van de effecten. De opzet zou experimenteel zijn, met andere woorden: scholen zouden zich opgeven voor Streef, de onderzoekers zouden scholen/bouwen toewijzen aan een van de deelprojecten. Er was sprake van een gekruist design waarbij van elke school een bouw in een interventiestudie participeerde en daarmee in de experimentele groep voor die interventie, terwijl de andere bouw binnen dezelfde school behoorde tot de controlegroep voor een van de andere interventies.

Al vroeg in de werving bleek dat er tegen die voorwaarde niet voldoende scholen zouden zijn die mee wilden werken. Hoewel de deelprojecten inhoudelijk allemaal met onderwijskwaliteit te maken hadden waren ze vanuit het perspectief van de scholen toch te verschillend voor toewijzing. De meeste scholen hadden een duidelijke voorkeur voor een van de projecten, gebaseerd op de inhoud, maar ook op de te investeren tijd. Om de deelprojecten toch uitvoerbaar te krijgen is de random toewijzing losgelaten en is elke school ingedeeld in het deelproject van hun voorkeur. Het gekruiste design is wel grotendeels in stand gebleven, behalve op te kleine scholen waarbij uiteindelijk het hele team aan een interventie heeft deelgenomen. Omdat er van de scholen zowel een voormeting beschikbaar kwam op het niveau van de leerlingen als op het niveau van het team was het achteraf mogelijk om te corrigeren voor eventuele aanvangsverschillen op deze terreinen. Bovendien hebben er uiteindelijk ruim 30 scholen meer aan Streef meegedaan waardoor de analysemogelijkheden in stand bleven.

In dit hoofdstuk wordt in 3.2 eerst een beschrijving gegeven van het verloop van de werving en de uiteindelijke onderzoeksgroep van scholen, leerkrachten en leerlingen. Daarna wordt in 3.3 de specifieke opzet en het verloop van elk van de interventiestudies toegelicht. In 3.4 worden de beide dataverzamelingsprojecten en in 3.5 de opzet van de analyses. In tabel 3.1 wordt eerst een overzicht gegeven van het hele programma zoals het is uitgevoerd.

Tabel 3.1: Opzet van het onderzoek

	Schooljaar 2009-2010	Schooljaar 2010-2011	Schooljaar 2011-2012
Dataverzameling bij leerlingen			
Alle scholen	Groep 1 t/m 7, alleen juni 2010	Groep 1 t/m 8, in januari en juni 2011	Groep 2 t/m 8, in januari en juni 2012
Dataverzameling bij leerkrachten			
Alle scholen	Digitale vragenlijst in voorjaar voor alle leerkrachten		Herhaalde meting (einde schooljaar)
Deelproject 1: Werken met referentieniveaus			
		X	X
Deelproject 2: Gebruik maken van opbrengsten			
		X	
Deelproject 3: Differentiatie en excellentie			
			X

3.2 Werving en respons

Gebaseerd op de te verwachte effectgroottes was vooraf bepaald dat er aan *Werken met referentieniveaus* en *Gebruik maken van opbrengsten* 36 scholen zouden moeten participeren (18 met de onderbouw en 18 met de bovenbouw) en dat de deelname aan *Differentiatie en Excellentie* – casestudie uit 10 scholen moest bestaan om betrouwbare en valide conclusies te kunnen trekken, waarmee er in het totaal minimaal 82 zouden moeten participeren. Voor de werving van Streef is gebruik gemaakt van de scholenbestanden van de beide onderwijsadviesorganisaties die in het deelproject over opbrengstgericht werken de begeleiding van de scholen verzorgden: Marant – in de regio Arnhem – en Cedin – in de regio Noord Nederland. Vooraf zijn de volgende groepen scholen uitgesloten vanwege hun atypische situatie en de belasting van het onderwijsveld:

- Zeer kleine scholen
- Zeer zwakke scholen volgens de Onderwijsinspectie
- Scholen die participeerden in lopend onderzoek van het GION

Dit resulteerde in een bestand van ongeveer 500 scholen en bijbehorende besturen die benaderd zijn met een brief en een folder. Daarnaast was ook een de website project-streef.nl online met aanvullende informatie. Vervolgens zijn eerst alle besturen gebeld, met de vraag of er op bestuursniveau belangstelling was om aan Streef deel te nemen. Afhankelijk van de belangstelling op bestuursniveau zijn de individuele scholen gebeld. Aan zowel besturen als scholen is aangeboden een presentatie van het project te verzorgen. De wervingsperiode heeft gelopen van januari tot en met juni 2010 en heeft uiteindelijk geresulteerd in een groep van 94 deelnemende scholen. Rondom Nijmegen zijn in 2011 nog 8 aanvullende scholen geworven

omdat er aan *Werken met referentieniveaus* in 2010-2011 te weinig scholen hadden deelgenomen en dit project in de tweede helft van het kalenderjaar 2011 nogmaals is uitgevoerd. Ook voor een van de TIER-projecten zijn in 2011 nog 18 aanvullende scholen geworven die ook participeerden in het onderzoeksdeel, waardoor het totaal aantal scholen op 120 uitkwam.

Scholen konden bij aanmelding aangeven wat het project van hun voorkeur was en met welke bouw zij mee wilden doen. Oorspronkelijk was in het onderzoeksvoorstel opgenomen dat scholen random met een boven- (groep 5 t/m 8) of onderbouw (groep 1 t/m 4) aan een van de projecten zou worden toegewezen. Dit bleek in de praktijk niet uitvoerbaar. Elke school had een sterke voorkeur voor een van de deelprojecten en wilde niet deelnemen als zij pas later te horen zouden krijgen aan welk deelproject ze mee konden doen. De keuze voor een bepaald deelproject paste bij de huidige ontwikkeling van de school en deelname aan een ander deelproject niet. Alle scholen zijn vervolgens ingedeeld bij de deelproject van hun keuze met de bouw van hun keuze. Deze quasi-experimentele opzet heeft een belangrijk nadeel ten opzichte van de oorspronkelijke experimentele opzet; het toeschrijven van eventuele effecten aan de interventie is problematischer. Immers, de motivatie om aan het deelproject mee te doen, en het feit dat het project inhoudelijk aansluit bij de schoolontwikkeling kunnen een belangrijke bijdrage hebben gehad aan eventuele effecten. Vasthouden aan de voorgestelde experimentele opzet zou echter betekenen dat het benodigde aantal scholen niet behaald zou worden en dat scholen mee moesten doen aan een project waarvoor zij misschien minder gemotiveerd zouden zijn. De kans op tussentijdse uitval is dan ook aanmerkelijk groter. Om toch een uitspraak te kunnen doen over of eventuele effecten veroorzaakt worden door de interventie is – gebruik makend van informatie over de scholen voorafgaand aan deelname – nagegaan in hoeverre er sprake is van validiteitsbedreigers.

Doordat elke school (in principe) deelneemt met één bouw aan één deelproject, is elke school tegelijkertijd experimentele en controleschool. Doet een school bijvoorbeeld met de bovenbouw mee aan *Werken met referentieniveaus* dan maakt de onderbouw van deze school deel uit van de controlegroep voor de scholen die met de onderbouw meedoen aan *Gebruik maken van opbrengsten*. Deze opzet heeft als voordeel dat de controlescholen waarschijnlijk wat betreft motivatie, attitude ten opzichte van het onderwijsbeleid en betrokkenheid bij het project en het onderzoek niet heel veel afwijken van de experimentele groep. De school doet immers zelf ook mee aan Streef. Tegelijkertijd is er ook een nadeel; er is altijd kan op ‘weglekken’ van de interventie. Onderbouwleerkrachten kunnen zaken oppikken van hun bovenbouwcollega’s waardoor ze niet helemaal ‘onbehandeld’ zijn. Dit wordt deels opgeheven omdat een bepaalde bouw van een school die aan deelproject A meedoet als controlegroep geldt voor andere scholen die aan deelproject B meedoen. Feit is echter wel dat de deelprojecten onderling niet volledig verschillen, allemaal hebben ze immers als doel de kwaliteit van het onderwijs te verbeteren. In de afsluitende vragenlijst aan de scholen is

daarom ook een aantal vragen gesteld over de mate waarin het project bijvoorbeeld in teamvergaderingen is besproken.

‘In principe’ betekent dat er enkele scholen zijn geweest die aan meer dan één deelproject hebben meegedaan en enkele scholen die met meer dan één bouw hebben meegedaan, bijvoorbeeld scholen met een klein team. In het bepalen van de experimentele groep en de controlegroep per deelproject is hier rekening meegehouden. Dit resulteert uiteindelijk in de deelname per deelproject zoals aangegeven in Tabel 3.2.

Tabel 3.2 Overzicht van Streef-scholen per deelproject

		Scholen	Overlap
A	Deelproject 1 (referentieniveaus) 2010 alleen onderbouw	2	1 school (252) heeft met de bovenbouw met project 2 (F) meegedaan
B	Deelproject 1 (referentieniveaus) 2010 alleen bovenbouw	3	1 school (220) heeft ook met project 3 (J) meegedaan
C	Deelproject 1 (referentieniveaus) 2010 onder- en bovenbouw	7	-
D	Deelproject 1 (referentieniveaus) 2011	8	-
E	Deelproject 2 alleen onderbouw	10	1 school (244) heeft ook met project 4 (K) meegedaan
F	Deelproject 2 alleen bovenbouw	17	1 school (252) heeft met de onderbouw met project 1 (A) meegedaan
G	Deelproject 2 onder- en bovenbouw	11	1 school (231) heeft ook met project 3 (J) meegedaan
H	Deelproject 2 MB 2010 (TIER)	6	2 scholen (259, 263) hebben ook aan project 4 (K) meegedaan
I	Deelproject 2 MB 2011 (TIER)	18	1 school (277) heeft ook aan project 4 (K) meegedaan
J	Deelproject 3 (Excellentie)	9	Zie B en G
K	Deelproject 4 (Leren lezen) (TIER)	18	4 scholen (244, 259, 263, 277) hebben ook aan andere projecten meegedaan (E, H, I)
	Totaal aantal deelnemende scholen	109-7 ‘dubbele’ scholen =102	
L	Geen deelname (wel aangemeld maar nooit begonnen of heel snel afgevallen)	18 (14%)	
	Totaal aantal aangemelde scholen	120	

Ten slotte moet wel opgemerkt worden dat onderzoek naar dergelijke actuele onderwerpen altijd beïnvloed wordt door invloeden van andere bronnen, in dit geval bijvoorbeeld stimulans en informatie vanuit het ministerie, maar ook de media, ouders, besturenorganisaties en dergelijke.

3.3 De deelprojecten

3.3.1 Deelproject 1 – Werken met referentieniveaus

Het doel van dit deelproject was leerkrachten kennis te laten maken met de referentieniveaus en hen te leren hoe ze daar in de praktijk, bij het toetsen van leerlingen, mee kunnen werken. De interventie bestaat uit vier bijeenkomsten; twee bovenschoolse bijeenkomsten en twee

bijeenkomsten op schoolniveau. De interventie is gericht op het bepalen van doelen voor rekenen en begrijpend lezen door teams van leerkrachten. De gehanteerde werkwijze is een aangepaste vorm van standaardsetting die ontwikkeld is binnen het project PPO. De leerkrachten kregen in de eerste bovenschoolse bijeenkomst een serie rekenopgaven voorgelegd met de vraag aan te geven in hoeverre de opgaven beheerst moeten worden wil er sprake zijn van een minimum, fundamenteel of streefniveau door leerlingen in hun eigen klas. De resultaten daarvan werden naast de officiële referentieniveaus gelegd. In de tweede bijeenkomst op de eigen school werden de antwoorden van de leerkrachten uit de eerste ronde naast de feitelijk resultaten van hun eigen leerlingen gelegd met als doel hen een realistische visie te geven op de referentieniveaus en aan te geven wat de referentieniveaus betekenen voor de leerlingen in hun eigen groep. In de derde en vierde bijeenkomst lag de focus op begrijpend lezen. In het onderzoek is nagegaan in hoeverre de door de leerkrachten aangegeven niveaus overeenkwamen met de officiële referentieniveaus en in welke mate ze overeenkwamen met het feitelijke niveau van hun leerlingen. Met behulp van de data die verzameld zijn bij leerkrachten en leerlingen is nagegaan of de leerkrachten die aan dit deelproject deel hebben genomen hun attitude en hun handelen hebben aangepast en of het een betekenis heeft gehad voor de resultaten van leerlingen.

Voor dit deelproject zijn de volgende onderzoeksvragen gesteld:

1. Komen teams van leerkrachten tot ongeveer dezelfde doelen aan het eind van de onder- en bovenbouw en hoe verhoudt zich dat tot de wettelijk ingevoerde referentieniveaus?
2. Komen de geformuleerde doelen overeen met wat leerlingen feitelijk presteren?
3. Heeft het formuleren van doelen gevolgen gehad voor het denken van leerkrachten over referentieniveaus en het gebruik van dergelijke niveaus in de klas?
4. Boeken leerlingen op scholen waar doelen geformuleerd zijn meer vorderingen dan op andere scholen?

3.3.2 *Deelproject 2 – Gebruik maken van opbrengsten*

In schooloverstijgende bijeenkomsten en consultaties op school, in het totaal 9 bijeenkomsten, worden leerkrachten geschoold en begeleid bij het optimaal gebruik maken van informatie uit leerlingtoetsen. Omdat Streef een samenwerking betrof tussen de Rijksuniversiteit Groningen en het Cito is het Cito-LOVS in het traject leidend geweest. Leerkrachten leerden, hands on, hoe zij via een stappenplan, met behulp van de computer een goede analyse konden maken en vervolgens welke gevolgen dat had voor het onderwijs daarna.

De bovenschoolse bijeenkomsten werden verzorgd door een RUG-medewerker en hadden meer het karakter van een training, de bijeenkomsten op de eigen school werden verzorgd door een medewerker van een van de onderwijsbegeleidingsdiensten en had meer een begeleidend karakter. Tussen de bijeenkomsten moesten de leerkrachten een aantal opdrachten uitvoeren. De ingeschatte tijdsbesteding voor leerkrachten was 40 uur. Aan het deelproject hebben in het totaal ruim 260 leerkrachten meegedaan. Op de bijeenkomsten op de eigen

school waren steeds alle deelnemende teamleden aanwezig, bij de bovenschoolse bijeenkomsten – waarvoor sommige deelnemers ruim een uur moesten rijden – was elke keer 60 tot 80% van de deelnemers aanwezig.

In het onderzoeksoverzicht is nagegaan wat de gevolgen van deelname waren voor leerkrachten en leerlingen. Wat betreft de leerkrachten is gekeken naar de feitelijke kwaliteit van hun analyses door analyse-rapportages te beoordelen en naar hun attitude en handelen in de klas zoals zij dat zelf rapporteren in de leerkrachtvragenlijst. Wat betreft de leerlingen is gekeken naar hun vorderingen in het jaar dat de leerkrachten aan de interventie deelnamen en naar de situatie een jaar na deelname.

De volgende onderzoeksvragen waren aan de orde:

1. In welke mate en hoe gebruiken de leerkrachten de gegevens die ze kunnen ontleen aan het Cito-leerlingvolgsysteem?
2. Wat zijn de effecten op de vorderingen van leerlingen als leerkrachten als team geschoold worden in een beter gebruik?

De scholen die aan dit deelproject hebben deelgenomen verschillen op schoolgrootte en leerlingpopulatie nauwelijks van de andere scholen die aan Streef meedoen, alleen de scholen die met beide bouwen meededen waren gemiddeld genomen kleiner dan de andere scholen.

3.3.3 Deelproject 3 – Differentiatie en excellentie

Dit deelproject betreft een case-studie waarin 9 scholen gedurende een schooljaar zijn begeleid en gevolgd om de volgende onderzoeksvraag te kunnen beantwoorden:

- Welke activiteiten ondernemen scholen om excellentie (niet alleen cognitief opgevat) bij leerlingen te bevorderen en is er een samenhang met de ontwikkeling van leerlingen?

Waaronder de volgende deelvragen zijn gesteld:

1. Op welke dimensies onderscheiden leerkrachten talenten bij leerlingen en welke instrumenten zetten ze vervolgens in om deze daadwerkelijk te kunnen onderscheiden?
2. Welke talenten onderscheiden ouders bij hun kinderen, en hoe vinden ze dat het onderwijs daarmee omgaat?
3. Hoe vindt differentiatie gericht op het omgaan met bijzondere talenten in de dagelijkse onderwijspraktijk plaats?
4. Hoe ervaren betrokken leerlingen de speciale onderwijsarrangementen?
5. Wat zijn de gevolgen voor de ontwikkeling van de betrokken leerlingen?

Op basis van informatie uit de voormeting en een intakegesprek en observatie aan het begin van het schooljaar 2011-2012 is, afhankelijk van de specifieke situatie van elke school en de gewenste ontwikkeling binnen een globaal vast stramien een planning gemaakt voor het hele schooljaar waarin elke school nog vier keer is bezocht en waarin leerlingen, leerkrachten en directie zijn gevolgd bij hun ontwikkeling. Bij alle scholen is er tijdens de bezoeken regelmatig geobserveerd in een of meerdere groepen. Het laatste bezoek bestond meestal uit

een teambijeenkomst waarin teruggekeken werd op het schooljaar en vooruitgekeken naar het vervolg.

De onderzoekers stelden zich tijdens en na het bezoek op als ‘critical friend’; er is niet slechts data verzameld op de scholen, maar na afloop van elk bezoek is er feedback gegeven en waar mogelijk advies over knelpunten of vervolgstappen. Van elk bezoek is een verslag gemaakt dat is opgestuurd aan de school ter fiattering. Op basis van deze verslagleggingen is per casus een beschrijving gemaakt en zijn vervolgens over de cases heen de onderzoeksvragen beantwoord, aangevuld met resultaten van de nameting van de leerkrachtvragenlijst.

Wat betreft schoolgrootte en leerlingpopulatie (percentage leerlingen met een gewicht van 1,2) verschillen de scholen die aan dit deelproject meedoen niet van de andere scholen die aan Streef deelnemen.

3.4 Dataverzameling

3.4.1 *Dataverzameling bij leerkrachten*

Om de beginsituatie en de veranderingen in leerkrachtgedrag en attitude in kaart te brengen en om te onderzoeken wat de relatie is met vorderingen van leerlingen is in het voorjaar van 2010, voorafgaand aan de verschillende deelprojecten binnen Streef en in het voorjaar van 2012, na afloop van de deelprojecten een digitale vragenlijst bij alle leerkrachten op de deelnemende scholen afgenomen. In 2012 zijn schoolleiders en intern begeleiders apart bevraagd. De vragenlijst bevat vragen op alle terreinen die in de deelprojecten aan de orde komen: referentieniveaus, professionalisering op het gebied van opbrengstgericht werken en differentiatie en excellentie. De selectie en de ontwikkeling van de vragen en de resultaten op de voormeting in 2010 zijn beschreven in Deunk e.a. (2010). In de bijlage is een overzicht opgenomen van de variabele die voor dit deelproject van belang zijn. Daarin is ook kort aangegeven wat de inhoud van de schalen en indexen is.

De volgende onderzoeksvragen zijn voor het onderzoek onder leerkrachten geformuleerd:

1. Wat is de attitude en werkwijze van leerkrachten met betrekking tot aspecten van opbrengstgericht werken, te weten prestatieniveaus en doelen, referentieniveaus en toetsgebruik en –analyse?
2. Wat is de attitude en werkwijze van leerkrachten met betrekking tot differentiatie?
3. Wat is de attitude en werkwijze van leerkrachten met betrekking tot professionalisering en samenwerking in het team?
4. Wat is de samenhang tussen de opbrengstgerichtheid van leerkrachtenteams en de leerwinst van hun leerlingen?
5. Wat is de attitude en werkwijze van schoolleiders en intern begeleiders en hoe verhoudt dit zich tot de attitude en werkwijze van leerkrachten?

3.4.2 Dataverzameling bij leerlingen

Bij de leerlingen van alle deelnemende scholen worden de volgende gegevens uit het Cito-leerlingvolgsysteem verzameld:

- Groep 1 & 2: Ordenen/Rekenen voor Kleuters, Taal voor Kleuters
- Groep 3, 4 & 5: Technisch lezen, Begrijpend lezen (groep 4 en 5), Spelling, Rekenen-wiskunde
- Groep 6,7 en 8: Spelling, Begrijpend lezen, Rekenen-wiskunde

Van leerlingen die doubleren worden de gegevens uit het Cito-leerlingvolgsysteem gewoon verzameld. De scores van doubleurs en onvertraagden kunnen immers op dezelfde schaal worden afgebeeld. Van de leerlingen van groep 8 worden geen data verzameld bij de eerste afname, van de leerlingen van groep 1 niet in het laatste schooljaar.

De scholen nemen de toetsen in eigen beheer volgens de handleiding af, voeren de gegevens in in het digitale Cito-LOVS en verzenden deze gegevens naar het Cito. Scholen die dit systeem niet gebruiken sturen een Excel-bestand of een papieren uitdraai op.

Van de leerlingen zijn daarnaast achtergrondgegevens verzameld over geslacht, leerlinggewicht en leeftijd

Deze dataverzameling resulteerde in een bestand van ongeveer 17000 leerlingen verspreid over 120 scholen. Van maar een beperkt deel van de leerlingen is het bestand helemaal compleet. Veel scholen waren in de dataverzamelingsperiode bezig over te stappen op nieuwe versies van de Cito-LOVS-toetsen bijvoorbeeld van Ordenen naar Rekenen voor kleuters, maar ook bij Taal voor kleuters, Spelling, Rekenen-Wiskunde en Begrijpend lezen waren er verschillende versies in gebruik op de scholen. Waar mogelijk zijn scores omgezet, bij toetsen waar dat niet mogelijk was zijn verschillende analyses uitgevoerd die later in een meta-analyse samen zijn genomen (zie volgende paragraaf).

De dataverzameling bij leerlingen kent geen ‘eigen’ onderzoeksvragen maar wordt gebruikt voor het beantwoorden van de onderzoeksvragen van de andere projecten.

3.5 Opzet analyses

Voor het beantwoorden van de onderzoeksvragen zijn verschillende soorten analyses gebruikt. Voor *Werken met referentieniveaus* zijn bijvoorbeeld door leerkrachten geformuleerde doelen vergeleken met de wettelijk vastgestelde referentieniveaus, voor *Gebruik maken van opbrengsten* zijn analyse-rapportages beoordeeld en voor *Differentiatie en excellentie* de kwalitatieve gegevens uit de cases. Deze deelproject-specifieke analyses worden in de verschillende deelrapporten toegelicht en beschreven. In deze paragraaf wordt ingegaan op de analyse van de leerkrachtvragenlijsten en de leerlingdata die in alle projecten een rol spelen.

De crux van het werken met een vaste groep scholen is hierin gelegen, dat globaal in de helft van deze scholen het onderzoek zich met name richt op de bovenbouw, en op de andere helft op de onderbouw. Desalniettemin worden de vorderingen van de leerlingen in beide

vastgesteld en worden de leerkrachten van alle groepen bevroegd. Daarmee kan afhankelijk van het deelproject steeds een ‘experimentele’ en een controlegroep gevormd worden. Voor de controlegroep worden steeds de bouwen/scholen geselecteerd die nergens aan mee hebben gedaan of die aan een ander deelproject hebben deelgenomen dat inhoudelijk niet lijkt op het betreffende deelproject.

Met het loslaten van de random toewijzing kan feitelijk niet meer van een experimentele groep gesproken worden, beter is te spreken van een behandelde en een niet-behandelde groep. Door op basis van de voormetingen bij leerkrachten en leerlingen en op basis van beschikbare achtergrondgegevens van de scholen te controleren voor de vergelijkbaarheid van de behandelde en de niet-behandelde groep kunnen wel uitspraken gedaan worden over de effecten van de behandeling, zij het minder hard. In paragraaf 3.2 is al meer uitgebreid ingegaan op de voor- en nadelen van een experimentele opzet.

Om de meer inventariserende vragen over attitude en handelen van leerkrachten te beantwoorden zijn eenvoudige frequentiestaten, gemiddelden e.d. berekend, waar relevant, uitgesplitst naar kenmerken van de scholen zoals deelname aan een bepaald deelproject. Om veranderingen tussen 2010 en 2012 te onderzoeken zijn gepaarde t-toetsen uitgevoerd. Om verschillen tussen ‘experimentele’ en controlescholen in veranderingen tussen 2010 en 2012 te onderzoeken is getoetst of de verschilcores voor beide groepen verschillen.

Voor het onderzoeken van de relatie tussen leerkracht- en schoolkenmerken en leerlingvorderingen zijn clusters scholen gevormd die verschillen in hun manier van (opbrengstgericht) werken en is onderzocht welk van deze clusters samenhangt met de beste leerlingprestaties.

Voor het nagaan van de effecten van – met name – *Gebruik maken van opbrengsten* is per leerstofdomein, per cohort een multilevel-analyse uitgevoerd met de voortoets als covariaat. Elke analyse levert een effectgrootte van deelname van de bouw op. Deze effectgroottes zijn vervolgens samengevoegd in een meta-analyse per leerstofdomein. Dit is zowel uitgevoerd voor de directe effecten (in het jaar van deelname) als de retentie-effecten (een jaar na deelname).

4 Resultaten

In dit hoofdstuk worden van alle deelprojecten en van het onderzoek onder leerkrachten de belangrijkste resultaten gepresenteerd. In hoofdstuk 5 wordt daarna ingegaan op de samenhang tussen de resultaten van de projecten en de gevolgen daarvan voor beleid, de onderwijspraktijk en vervolgonderzoek.

4.1 Ontwikkeling bij schoolteams

Samengevat kan op basis van de analyses gesteld worden dat de schoolverbetertrajecten van Streef hebben geleid tot meer en intensiever gebruik van toetsgegevens en -rapportages. Daarnaast hebben de trajecten de attitudes van leerkrachten beïnvloed, vooral op het gebied van opbrengsten en professionalisering. Beter toetsgebruik en een positievere houding tegenover aspecten van opbrengstgericht werken vormen een belangrijke basis voor kwaliteitsverbetering in het onderwijs en zijn hiermee een eerste stap in de richting van prestatieverhoging van leerlingen. Daarmee zijn scholen er echter nog niet: voor het vergroten van de leerwinst van leerlingen lijken een voorttrekkende rol van de schoolleider en intern begeleider en vooral een opbrengstgericht, samenwerkend en differentiërend team essentieel. De meest opvallende resultaten worden hieronder nader toegelicht.

Bij vragen naar het vaststellen van leerdoelen vallen met name de problemen met begrijpend lezen op. Leerkrachten geven aan het minst tevreden te zijn over het begrijpend leesniveau van hun leerlingen. Bovendien stellen zij relatief vaak geen gewenst eindniveau vast. Dit is een gemiste kans, aangezien expliciete leerdoelen kunnen bijdragen aan prestatieverbetering. Slechts rond de helft van de leerkrachten geeft aan de leerdoelen voor de verschillende vakken samen met het team te expliciteren. De schoolleiders en intern begeleiders bevestigen dit beeld, al rapporteren zij iets meer teamoverleg. Hier is nog verbetering mogelijk, al stellen leerkrachten na deelname aan de Streef interventie wel vaker rekendoelen vast in overleg met de intern begeleider of het team dan vooraf.

De Streef-interventies hanger verder samen met de opbrengstgerichte attitude van leerkrachten. Leerkrachten die deel hebben genomen aan een van de deelprojecten zijn in de periode 2012-2012 significant positiever geworden over de haalbaarheid en wenselijkheid van het verbeteren van leerlingprestaties en hebben daardoor in 2012 een significant positievere attitude over opbrengsten dan de leerkrachten in de controlegroep. Schoolleiders en intern begeleiders zijn daarbij in 2012 nog positiever over opbrengsten en basisvaardigheden dan de leerkrachten.

De referentieniveaus zijn een door de overheid ingevoerd middel ter stimulering van een opbrengstgerichte werkwijze op school. In de periode van 2010 tot 2012 is de bekendheid met referentieniveaus bij alle leerkrachten significant toegenomen, vooral bij de leerkrachten die hebben deelgenomen aan een van de Streef-interventies. Bovendien zijn alle schoolleiders en intern begeleiders in 2012 bekend met de referentieniveaus. Dit wil echter niet zeggen dat de referentieniveaus ook leven: nog geen derde van de leerkrachten praat erover op school. Bijna niemand van de leerkrachten, schoolleiders en intern begeleiders verwacht een negatief effect van de referentieniveaus, maar ruim de helft verwacht geen effect of weet het niet. Bovendien zijn de positieve verwachtingen die leerkrachten hadden in 2010 getemperd in 2012, wellicht door de grotere bekendheid. Desondanks staan de leerkrachten redelijk positief tegenover de referentieniveaus. Zij zien de meerwaarde vooral in het expliciteren van doelen en het gebruik bij differentiatie.

Leerkrachten, vooral degenen die hebben deelgenomen aan een Streef interventie, maken in 2012 gebruik van meer toetsrapportages en voeren uitgebreidere analyses uit op de toetsgegevens dan in 2010. De Streef-leerkrachten maken in 2012 significant meer gebruik van uitgebreide analysemogelijkheden dan de leerkrachten in de controlegroep. Ook de schoolleiders en intern begeleiders maken ruimschoots gebruik van de analysemogelijkheden van de Cito LOVS toetsen. De leerkrachten in de Streef-groep zijn daarnaast in de periode 2010-2012 significant meer gaan reflecteren op hun onderwijspraktijk en de eerdere prestaties van de leerlingen naar aanleiding van toetsresultaten. De leerkrachten in de Streef-projecten hebben zich dus positief ontwikkeld in gebruik, analyse en interpretatie van toetsgegevens en zijn hiermee, in ieder geval wat betreft toetsgebruik, meer opbrengstgericht gaan werken.

Differentiatie is een essentieel onderdeel van opbrengstgericht werken. Er zijn nog verbeteringen nodig op gebied van het optimaal gebruik maken van verschillende prestatieniveaus, het aanbieden van verschillende verwerkingsactiviteiten na een toets en de houding tegenover de wenselijkheid en haalbaarheid van differentiatie. Leerkrachten in beide groepen scoren in 2010 en 2012 laag tot gemiddeld op de verschillende differentiatieonderdelen en zouden hier dus meer ondersteuning bij kunnen gebruiken. Ook zouden zij meer kunnen samenwerken. De meeste samenwerking betreft het uitwisselen van lesmateriaal of het praten over de ontwikkeling van nieuw lesmateriaal met collega's. Lesobservaties zijn een belangrijke bron van informatie bij professionele ontwikkeling, maar slechts een-derde van de leerkrachten observeert wel eens bij een collega in de les.

De leerkrachten die de vragenlijst hebben beantwoord, vormen teams in hun scholen. De attitudes en werkwijzen van de leerkrachten in de schoolteams blijkt van invloed op de leerwinst van de leerlingen. Er zijn *opbrengstgerichte* teams, die veel samenwerken, reflecteren en differentiëren en die veel waarde hechten aan rekenen en taal. *Individualistische*

teams hechten wel veel waarde aan de basisvaardigheden en voeren uitgebreide rapportages uit, maar werken weinig samen en reflecteren en differentiëren weinig. *Behoudende* teams hechten weinig waarde aan innovatie en professionalisering, zijn gemiddeld in mate van samenwerking en aandacht voor prestaties en differentiëren veel. *Niet opbrengstgerichte* teams tenslotte hechten weinig belang aan basisvaardigheden en opbrengsten. Leerlingen van *opbrengstgerichte* en *behoudende* leerkrachtenteams boeken significant meer vooruitgang dan leerlingen van *individualistische* en *niet-opbrengstgerichte* teams in meerdere groepen en leerstofdomeinen. De *opbrengstgerichte* en *behoudende* teams hebben gemeen dat zij relatief vaak differentiëren in de klas. Behoudende teams staan echter minder open voor onderwijsvernieuwingen en zien minder noodzaak zichzelf te blijven ontwikkelen op professioneel vlak. Het valt op dat het maken van uitgebreide toets-rapportages, zoals de *individualistische* teams doen, niet genoeg is om extra leerwinst te boeken. Leerkrachten moeten ook naar de informatie uit de rapportages *handelen* en moeten met elkaar als team werken aan en reflecteren op hun onderwijs. Dit geeft het belang aan van ondersteuning bij het schoolbreed implementeren van opbrengstgericht werken, een van de uitgangspunten van de Streefprojecten.

4.2 Werken met referentieniveaus

Uit de resultaten van de standaardsetting blijkt dat de leerkrachtenteams behoorlijk variëren in wat ze als doelen bepalen voor hun leerlingen. Hoewel de gemiddelde waarden voor het minimumniveau nog redelijk in de buurt zitten van het niveau dat landelijk door 90% van de leerlingen op dit moment wordt bereikt, is dat bij de niveaus fundamenteel en streef bepaald niet het geval. Gemiddeld genomen ligt de lat daar veel hoger dan het landelijk bereikte niveau. Daarmee leggen de leerkrachten hun standaard duidelijk hoger vast dan de wettelijke referentieniveaus 1F en 1S.

De leerlingresultaten op de LOVS-toetsen Begrijpend lezen en Rekenen-Wiskunde laten zien dat de minimumdoelen door de meeste leerlingen al (bijna) worden bereikt. Het fundamenteel en streefniveau ligt veel hoger dan het huidige bereikte niveau en ook als in aanmerking wordt genomen dat het hier een groep van wat sterkere scholen betreft, lijken het door de leerkrachten geformuleerde fundamenteel en streefniveau niet realistisch. Het meest omstreden blijkt het streefniveau, vooral voor rekenen. Dat heeft mogelijk te maken dat er bij de wettelijke referentieniveaus verschillende streefniveaus voor rekenen en taal zijn, die van rekenen ligt op een niveau dat minimaal 50% van de leerlingen behaald zou moeten worden, bij taal ligt dat percentage op 25%. Voor leerkrachten is dit verschil in de procedure mogelijk moeilijk te maken.

Uit de analyse van de leerkrachtvragenlijst blijkt dat de leerkrachten die aan dit deelproject hebben meegedaan beter weten wat de referentieniveaus inhouden dan leerkrachten die niet mee hebben gedaan. Bovendien zijn ze significant meer realistisch gaan aankijken tegen mogelijkheden van leerlingen in termen van te bereiken vaardigheidsniveaus dan hun collega's in de controlegroep. Daar hoort bij dat men soms tot meer reële verwachtingen komt, bijvoorbeeld ten aanzien van het perspectief van een E-leerling. Ook zijn de deelnemende leerkrachten significant meer bereid om met de leerlingen te communiceren over verwachtingen. Er zijn maar weinig leerkrachten die een negatief effect verwachten van de referentieniveaus. In 2010 verwachtten de meeste leerkrachten een positief effect van referentieniveaus op hun dagelijks werk en de prestaties van leerlingen. In 2012 is de positieve verwachting echter afgenomen en denken leerkrachten vaker dat de referentieniveaus geen invloed zullen hebben. Deze verandering hangt niet samen met deelname aan dit deelproject maar geldt voor alle leerkrachten.

Uit de analyse van de leerlingresultaten blijkt dat er op dit gebied geen effect van de interventie is. In groep 4 is er een lichte groei van de resultaten maar de controlegroep scoort steeds hoger.

4.3 Gebruik maken van opbrengsten

De eerste stap in de analyse is het onderzoeken van de veranderingen in de analysevaardigheden van leerkrachten. Deze blijken aan het eind van *Gebruik maken van opbrengsten* duidelijk hoger te liggen dan aan het begin van het traject. Een relativering is hier wel noodzakelijk: omdat slechts enkele leerkrachten twee- of zelfs driemaal een analyse hebben ingeleverd was het niet mogelijk om op leerkrachtniveau een voor- en een nameting met elkaar te vergelijken. Het was slechts mogelijk om de ingeleverde analyses aan het begin 'gemiddeld' te vergelijken met de later ingeleverde analyses. Bovendien is er geen sprake van een controlegroep. De vooruitgang is echter zo duidelijk dat er toch gesproken kan worden van een direct effect van *Gebruik maken van opbrengsten*. Deze conclusie bevestigt de conclusie van Visscher et al. (2010) die aangeven dat een beter gebruik van een leerlingvolgsysteem aan te leren is.

Vervolgens is er gekeken naar het meer indirecte effect van *Gebruik maken van opbrengsten* op basis van de leerkrachtvragenlijst. Hierbij was het wel mogelijk om de vergelijking tussen de situatie voor en na *Gebruik maken van opbrengsten* op leerkrachtniveau met elkaar te vergelijken. Uit deze vergelijking blijkt dat de groep leerkrachten die aan *Gebruik maken van opbrengsten* in 2010 al iets meer dan de leerkrachten in de controlegroep ervaring had met het analyseren van toetsresultaten, maar dat desondanks zij op de variabelen die het meest expliciet aan de orde zijn geweest meer vooruitgegaan zijn dan de controlegroep. Dit resulteert in 2012 in een situatie waarbij de leerkrachten die aan *Gebruik maken van*

opbrengsten hebben deelgenomen een uitgebreidere analyse maken van toetsresultaten, daarop in een breder kader reflecteren en meer gebruik maken van groepsplannen waarin vastgelegd is op welke manier zij het onderwijs in de komende periode gaan differentiëren. Wat betreft de keuzes die daarbij gemaakt worden lijken er geen verschillen en geen vooruitgang op te treden.

De vooruitgang in analysevaardigheden volgens de leerkrachten zelf wordt dus bevestigd door de onafhankelijke beoordeling van analyserapportages aan het begin en halverwege/het eind van *Gebruik maken van opbrengsten*.

De tweede onderzoeksvraag betreft het effect van het traject op de vorderingen van leerlingen. De analyses hiervan laten tot nu toe weliswaar een effect zien in de positieve richting maar dit effect is niet significant zodat we feitelijk moeten spreken van het ontbreken van een effect op leerlingniveau. Van de 25 geschatte effectgroottes zijn er – zij het niet allemaal significant - 18 positief en 7, negatief, slechts eenmaal laat een bovenbouwgroep een (niet significant) negatief effect zien en bij begrijpend lezen zijn de effecten allemaal positief. Bij alle leerstofdomeinen (behalve taal in de kleutergroep) is te zien dat het 90%-betrouwbaarheidsinterval zich voor het grootste deel aan de positieve kant van de 0 bevindt. Hoewel er dus in het algemeen een positieve tendens te zien is, is het effect, binnen de gehanteerde grenzen, (net) niet groot genoeg om significant te zijn.

De retentie-effecten zijn zien er minder positief uit. De effectgrootte is even vaak (niet-significant) positief als negatief, bij alle cohorten en beide leerstofdomeinen omvat het betrouwbaarheidsinterval de waarde 0 en het resulterende effect wijkt vanzelfsprekend ook niet af van 0. Samenvattende kunnen we stellen dat het licht positieve – zij het niet significante – effect wat optrad in het schooljaar waarin de leerkrachten deelnamen aan de interventie niet vastgehouden wordt in het schooljaar daarna.

Hierdoor is ook de laatste stap waarbij wordt onderzocht of het effect op leerlingniveau verklaard kan worden door een verbetering in attitude of gedrag van leerkrachten niet aan de orde.

4.4 Differentiatie en excellentie

Het signaleren van excellente leerlingen, het onderwerp van onderzoeksvraag 1 en 2, is voor elke school een moeilijk punt, ook scholen die al veel ervaring hebben zijn er nog niet tevreden over. Alle varianten komen voor en zijn te ordenen van reactief naar actief en deze ordening loopt ongeveer parallel aan de mate waarin de school al ervaring heeft op het terrein van excellente leerlingen. Opvallend is dat het erop lijkt dat scholen die actief signaleren en meer ervaring hebben ook vaker leerlingen mee laten doen die er misschien niet aan toe zijn. Op die scholen waren er bijvoorbeeld meer dan bij de andere scholen leerlingen die de

antwoorden van andere leerlingen overschreven. Het voordeel van deze selectiemaniër is dat er op deze scholen waarschijnlijk geen leerlingen ‘overgeslagen’ worden die wel profijt zouden kunnen hebben van extra stof. Iets wat bij andere scholen die op basis van ‘problemen’ signaleren wel voorkomt. Het bestaan van zo’n ‘tussengroep’ sluit wel goed aan bij de conclusies van de meta-analyses naar wat werkt voor hoogbegaafde leerlingen, namelijk dat er verschillende aanpassingen aangeboden moeten worden om alle leerlingen te bedienen. Afgaand op de vragen van scholen spelen verschillende dimensies waarop talenten te onderscheiden zijn niet zo’n grote rol. De cognitieve dimensie is het belangrijkste, maar de scholen die meer ervaring hebben kijken wel breder; op de meeste van deze scholen bleken ook leerlingen mee te doen die op grond van het toetscores niet opvielen of zelfs onder gemiddeld scoorden, maar van wie bijvoorbeeld een intelligentietest, op initiatief van de ouders, had uitgewezen dat ze hoogbegaafd waren.

Onder de 9 scholen was er geen enkele belangstelling om ouders actief mee te nemen in het onderzoek. Hun voornaamste zorg lag bij de manier waarop een groepsleerkracht binnen de eigen groep op een goede, haalbare manier de (zeer) goede leerlingen voldoende kon stimuleren. In de gesprekken op de scholen is de rol en mening van ouders wel herhaaldelijk aan de orde gekomen.

Alle 9 scholen waren bij aanvang van Streef van mening dat zij al differentieerden voor de (zeer) goede leerlingen; ze boden ze in ieder geval extra werk aan nadat ze de reguliere lesstof af hadden. Deelname aan Streef kwam wel altijd voort uit ontevredenheid daarover. Bij enkele scholen was de belangrijkste ‘openingsvraag’ de vraag om goed, liefst zelfwerkend, materiaal waarmee ze het onderwijs aan de (zeer) goede leerlingen konden verbeteren. Gaande het schooljaar bleek de ‘problematiek’ van het aanbieden van goed onderwijs voor deze doelgroep echter op andere gebieden te liggen, namelijk:

- Voor verrijkingsmateriaal dat voor deze doelgroep in de zone van naaste ontwikkeling ligt is begeleiding nodig. Voor hen is het verrijkingsmateriaal immers net zo moeilijk als reguliere lesstof is voor gemiddelde leerlingen.
- Voor het werken aan verrijkingsmateriaal is tijd nodig, tijd die gewonnen kan worden door in de reguliere lesstof te schrappen, bijvoorbeeld de herhalingsopgaven. Bepalen wat een leerling kan overslaan lijkt soms moeilijker dan aanbieden van extra werk.
- In aansluiting op het vorige punt: leerlingen hadden vaak meerdere keren per dag korte tijd om aan extra werk te werken, steeds als ze hun reguliere werk af hadden. Een dag- of weektaak kan hier oplossing bieden, waarin een leerling zelf tijd kan indelen en gedurende langere tijd aan een opdracht kan werken.
- Aanbieden van ander materiaal is een deel van de aanpassing van het onderwijs, aanpassen van de reguliere klassikale instructiemomenten waar de (zeer) goede leerlingen wel aan meedoen is een ander deel van de aanpassing. Niet alle leerkrachten benutten de

- ‘makkelijke’ mogelijkheden om binnen de reguliere les waarin van alle leerlingen in principe hetzelfde gevraagd wordt, de excellente leerlingen aan te spreken.
- Differentiatie staat of valt met basisvaardigheden van de leerkracht: klassenmanagement, goede instructie (regie in handen) en met de mate waarin (alle) leerlingen zelfstandig aan het werk kunnen zijn en verantwoordelijkheid hebben. Naarmate leerkrachten hier beter toe in staat zijn, zijn leerlingen beter in staat om zelfstandig aan het werk te gaan en creëren leerkrachten tijd om leerlingen individueel of in een klein groepje te begeleiden.
 - Leerkrachten die beginnen met het aanbieden van verrijkingsstof moeten er, nog meer dan anders, rekening mee houden dat ook de voorbereiding tijd kost. Dit lijkt logisch maar blijkt dat toch niet altijd te zijn. Ook het registreren en nakijken kost tijd. Verrijkingsmateriaal zal meer dan reguliere opdrachten bestaan uit het produceren in plaats van reproduceren; dus iets maken of schrijven in plaats van iets in- of aanvullen.
 - Goed extra materiaal is belangrijk, maar geen zelfwerkend toermiddel. Wat goed is, hangt ook deels af van de doelstelling die de school nastreeft. Aan doelstellingen die men voor deze doelgroep heeft ontbreekt het op bijna alle 9 scholen.

Hoewel er hele goede voorbeelden van leerkrachten gevonden zijn, ook in het begin van het schooljaar, was er op alle scholen in alle groepen ruimte voor verbeteringen. Los van het ontbreken van een visie op het onderwijs aan deze doelgroep, waardoor het feitelijk niet mogelijk is om te onderzoeken of het onderwijs succesvol is, waren het creëren van tijd bij leerlingen om aan de verrijkingsstof te werken en het creëren van tijd bij de leerkracht om hier aandacht aan te besteden de aspecten die door de leerkrachten het moeilijkst werden gevonden. Bij enkele leerkrachten stond het gebrek aan basisvaardigheden verdere ontwikkeling in de weg.

In bijna elke school wordt expliciet gemeld dat de ervaringen van leerlingen over het algemeen heel positief zijn, het onderwerp van onderzoeksvraag 4 en 5. Ze vinden het werk dat ze doen leuk en zijn voldaan als ze iets voor elkaar hebben gekregen waarvoor ze moeite hebben moeten doen. Ook leerlingen die al langere tijd met verrijkingsmateriaal werken blijven er enthousiast over.

Daartegenover staan wel ook wat minder positieve uitspraken van leerkrachten over leerlingen waarbij het niet goed gaat en van leerlingen over wat ze niet zo leuk of moeilijk vinden. In een enkele groep was sprake van een leerling voor wie het werk cognitief waarschijnlijk best passend zou zijn maar die niet te motiveren was om het aan te pakken of de vrijheid die vaak gepaard gaat met deze manier van werken niet aankon en bijvoorbeeld rond ging lopen en andere leerlingen storen. Dit zijn echter uitzonderingen waarbij vaak meer aan de hand is dan alleen (hoog)begaafdheid. Soms is er sprake van dubbele labelling waarbij leerlingen (hoog)begaafd zijn maar ook een ander leer- of gedragsprobleem hebben. Soms betreft het ook leerlingen bij wie de school zelf deel van de ‘blaat’ treft, bijvoorbeeld omdat er jarenlang niets is gedaan met het talent van de leerling waardoor de leerling gewend is geraakt

aan vervelen en rondhangen en de motivatie verloren heeft. Overigens leken er op meerdere scholen ook leerlingen met verrijkingsmateriaal te werken waarvan het de vraag is of zij voldoende bagage hebben, m.a.w. of ze begaafd genoeg zijn om mee te doen.

Negatieve ervaringen hebben leerkrachten het meest - op bijna alle scholen – met een ander ‘type’ (hoog)begaafde leerling; het type dat Betts en Niehart (1988) de ‘succesvolle’ leerling noemen. Over deze leerlingen zeggen leerkrachten dat ze ‘gestrest’ raken van het extra werk. Leerlingen zijn perfectionistisch, hebben zelf het gevoel dat ze het werk dat andere leerlingen moeten maken zelf ook allemaal moeten maken omdat ze anders op achterstand komen of iets niet snappen. Leerlingen kiezen ook vaak, in situaties waarin ze een keuze kunnen maken, voor een makkelijk werkje of iets wat ze al eerder hebben gedaan, in plaats van voor iets waarvan ze denken dat het moeilijk zal zijn of waarvan ze niet bij voorbaat weten hoe het moet.

Ook deze leerlingen zijn erbij gebaat dat ze al vroeg in hun schoolloopbaan ervaring opdoen met voor hen moeilijk, maar haalbaar werk, zodat ze leren dat risico’s nemen geen kwaad kan. Ook moet de leerling voldoende tijd hebben voor het werk, systematisch indikken van het reguliere werk is noodzakelijk.

Samenvattend kan geconcludeerd worden dat de ervaringen van leerlingen die systematisch met verrijkingsmateriaal werken positief zijn en dat minder positieve ervaringen deels voortkomen uit het feit dat leerlingen niet vanaf het begin van hun schoolloopbaan voldoende uitgedaagd zijn en pas in hogere leerjaren geconfronteerd worden met hogere eisen. Vanzelfsprekend vinden ook zij niet al het werk leuk en zijn er ook leerlingen waarbij het minder goed werkt. Net als bij leerlingen die op een gemiddeld niveau functioneren of daaronder zijn er onder de leerlingen die cognitief op een hoog niveau functioneren leerlingen die meer zorg van de leerkracht vragen bijvoorbeeld vanwege hun gedrag. Leerkrachten lijken dat van cognitief hoog presterende leerlingen minder goed te accepteren: ‘die moeten toch snappen hoe je je moet gedragen’. Net als dat veel leerkrachten van deze leerlingen verwachten dat ze heel zelfstandig kunnen werken.

Of de leerlingen feitelijk meer vooruitgaan dan op basis van hun eerdere resultaten is te verwachten en zich sociaal-emotioneel goed ontwikkelen is niet te beantwoorden met de beschikbare data. De onderzoeksgroep is heel klein, waarbij de grens van wie er wel en niet onder excellente leerlingen valt van school tot school verschilde, en het leerlingen uit alle verschillende leerjaren betreft. Bovendien hebben ze geen eenduidige interventie ondergaan waaraan een eventueel effect aan toe te schrijven is; hoewel elke school aan dit deelproject meedeed was de vaardigheid en de aanpak van de leerkracht in elke groep anders en betreft het meestal maar een of enkele leerlingen per groep.

5 Integratie en discussie

Het beleid van de landelijke overheid gericht op het verbeteren van de kwaliteit van het primair onderwijs uitgedrukt in prestaties van leerlingen richt zich op drie speerpunten die moeten leiden tot die kwaliteitsverbetering: het invoeren van referentieniveaus zodat leerkrachten een duidelijk doel hebben om naar toe te werken, het stimuleren van opbrengstgericht werken en het stimuleren van onderwijs aan excellente leerlingen. Deze formulering van de speerpunten geeft al aan dat de landelijke overheid daarvoor met name gebruik maakt van communicatieve beleidsinstrumenten waarmee scholen verleid, overreed, gestimuleerd moeten worden om daar aandacht aan te besteden om op die manier hun resultaten te verhogen. Het inzetten van deze beleidsinstrumenten past bij het geldende onderwijsbeleid waarin scholen, in ieder geval wat betreft het onderwijsproces, autonoom zijn. Beleidsondersteunend onderzoek verschaft het beleid en het onderwijsveld kennis over en inzicht in de implementatie en de doeltreffendheid van het beleid en de bevorderende en belemmerende factoren.

Het GION en Cito heeft daartoe een samenhangend onderzoeksprogramma opgezet waaraan ook twee onderwijsadviesorganisatie hebben meegewerkt en dat op de deelnemende scholen is uitgevoerd als schoolverbeteringsprojecten met daaraan gekoppeld dataverzameling bij leerkrachten en leerlingen om achteraf de werking van de interventies vast te stellen. Op deze manier is een situatie gecreëerd waarin scholen ‘extra’ gestimuleerd werden om de beleidsspeerpunten te implementeren zodat inzicht verkregen kon worden in de vraag hoe leerkrachten daar in een optimale situatie mee omgaan en van profiteren. Optimaal omdat ze zelf gekozen hebben voor deelname en er kosteloos bij begeleid werden. Heeft een dergelijke interventie geen effect dan is de kans dat stimulerend overheidsbeleid – op afstand – wel effect heeft klein.

Interessante vragen in dit integrerende hoofdstuk zijn dan dus: welke redenen hadden scholen om te participeren? Hebben de interventies effect gehad en zo nee, welke oorzaken kunnen daaraan te grondslag liggen: oorzaken die hun oorsprong vinden in het beleid, oorzaken die hun oorsprong vinden in de interventies/het onderzoek of andere oorzaken?

Tenslotte komen enkele vervolgvragen voor onderzoek en aanbevelingen aan de orde.

Werving en redenen voor deelname

In hoofdstuk 3 is al aangegeven dat de oorspronkelijke opzet waarin scholen random aan een deelproject toegewezen zouden worden tijdens de werving is losgelaten omdat scholen zelf wilden kiezen aan welk project ze mee zouden doen. Vanuit het perspectief van de scholen een logische keuze, vanuit het perspectief van het onderzoek een noodzakelijke koerswijziging omdat er anders niet voldoende scholen hadden meegedaan. Uit controles op aanvangs- en contextverschillen tussen scholen die aan verschillende deelprojecten meededen

bleek dat er geen verschillen van betekenis waren, waardoor de conclusies geldig bleven, zij het met minder zeggingskracht dan bij een zuiver experimentele opzet.

Tegen deze voorwaarden bleek het wel haalbaar om een voldoende grote groep scholen deel te laten nemen. Anders dan bij werving voor een onderzoek hadden we scholen iets te bieden, namelijk deelname aan een schoolverbeteringsproject, maar tegelijkertijd werd hen ook meer gevraagd, namelijk actieve participatie in het project en deelname aan het onderzoek. In enkele gevallen zijn scholen door hun bestuur (zwaar) gestimuleerd om mee te doen.

Degene die de beslissing nam om deel te nemen was in de meeste gevallen de directie. Op sommige scholen is deelname met het team besproken, op andere scholen is het team ingelicht over deelname. Bij deze scholen kwam het voor dat leerkrachten niet vrijwillig deelnamen.

Tijdens de werving, waarbij er meestal een presentatie of een gesprek op de school plaatsvond, is bijna altijd de motivatie voor deelname aan de orde gekomen, bij de scholen die aan *Differentiatie en excellentie* meededen, een casestudie, is door het intensieve contact nog meer bekend over de reden om mee te doen.

De scholen die aan dit deelproject van Streef hebben meegedaan verschillen op de gemeten leerkrachtvariabelen niet van de scholen die aan een van de andere deelprojecten hebben meegedaan, noch op de kenmerken van opbrengstgericht werken, noch op de kenmerken die meer specifiek betrekking hebben op de manier waarop men met (zeer) goede leerlingen omgaat. In de gesprekken met de scholen waarin ook hun motivatie voor dit deelproject besproken is werden uiteenlopende redenen genoemd. De meest genoemde reden is vanzelfsprekend de erkenning van het feit dat deze groep de laatste jaren onvoldoende aandacht heeft gekregen terwijl scholen wel een verplichting hebben en voelen om alle leerlingen tot hun recht te laten komen. Voor sommige scholen ligt de reden in het feit dat ze bij het aanpassen van het onderwijs aan deze groep leerlingen problemen tegenkomen; wat bieden we aan, waarom en hoe? Dit zijn dus scholen die zich er al wat langer bewust van zijn en al een start hebben gemaakt. Een deel van deze scholen heeft een plusgroep of participeert in een bovenschoolse plusgroep, maar erkent dat dat de eigen groepsleerkracht niet ontslaat van de noodzaak het onderwijs aan te passen aan deze leerlingen. De plusgroep staat op enkele scholen zelfs ter discussie vanwege de kosten en vanwege eerder genoemde reden dat de leerlingen toch het leeuwendeel van de tijd in de eigen groep doorbrengt en de groepsleerkracht uiteindelijk verantwoordelijk blijft voor het hele onderwijs aan de leerling. Andere scholen stonden echt nog aan het begin en wisten niet precies waar en hoe te beginnen. Tenslotte is er een school die aangeeft dat ze hoopt dat door het verhogen van de prestaties van de goede leerlingen het schoolgemiddelde omhoog gaat. Na zoveel jaar aandacht besteed te hebben aan de zwakkere leerlingen zit er daar geen verbetering meer in en bovendien zal een eventuele verbetering, waar heel veel moeite voor gedaan zal moeten worden, niet veel bijdragen aan een verhoging van het schoolgemiddelde.

Deelname aan *Gebruik maken van opbrengsten* dat door de meeste scholen werd gekozen, werd veelal gewoon beargumenteerd vanuit het idee dat de noodzaak werd gevoeld om hier

werk van te maken. Schoolleiders gaven aan dat zij van mening waren dat de vaardigheden van hun leerkrachten op gebied van het analyseren van toetsen en het vervolgens gedifferentieerd lesgeven verbeterd kon worden evenals de leerlingresultaten (in de breedte of op bepaalde deelgebieden). Ook veel leerkrachten gaven aan dit noodzakelijk te vinden en hoewel men over het algemeen wel van mening was daar zelf een rol in te spelen, stuitte het daadwerkelijk tijd besteden aan het analyseren van toetsen en het aanpassen van het eigen gedrag in de groep bij enkele leerkrachten wel op weerstand. Meestal lag het argument dan in de benodigde tijd, een moeilijke klas of een onderwijsopvatting waarin toetsen van (jonge) leerlingen niet past. Vaak kwam deelname (ook) voort uit het feit dat de inspectie vraagt om analyses. Soms was het beschikbaar stellen van het Cito-LOVS en de toetsen die in Streef gebruikt werden ook van belang en het feit dat ze op school begeleid zouden worden door een onderwijsadviseur van een hen bekende organisatie. Scholen die voor *Werken met referentieniveaus* kozen gaven aan dat ze de vaardigheden die in *Gebruik maken van opbrengsten* werden aangeleerd al beheersten of daar recentelijk al aandacht aan te hebben besteed. Uit de analyse van de voormeting bij leerkrachten bleek echter niet dat zij op de relevante variabelen beter scoorden dan de andere scholen.

De meeste scholen die niet mee wilden doen (+/- 80% van de aangeschreven scholen) gaven als belangrijkste redenen het ontbreken van tijd, nooit ergens aan meedoen, pas geleden op een andere manier met hetzelfde onderwerp bezig te zijn geweest en – in enkele gevallen – inhoudelijke problemen met de onderwerpen.

Effectiviteit van de interventies

Uit de resultaten van de afzonderlijke deelprojecten blijken zeker – bescheiden - effecten op de vaardigheden, de attitude en het handelen van leerkrachten. Van de deelprojecten *Gebruik maken van opbrengsten* en - in wat mindere mate - *Werken met referentieniveaus* kan dat met meer zekerheid aangetoond worden dan van de casestudie *Differentiatie en excellentie*, maar ook daar lijkt de kwalitatieve analyse deze conclusie wel te rechtvaardigen. De effecten zijn het grootst op de factoren die het dichtst bij de inhoud van de interventies liggen.

Ook wat betreft de effecten op leerlingen zijn de resultaten van *Differentiatie en excellentie* van een andere orde dan van *Werken met referentieniveaus* en *Gebruik maken van opbrengsten*. Bij *Differentiatie en excellentie* ging het om ervaringen van excellente leerlingen die beter passend onderwijs aangeboden kregen en daar meestal heel enthousiast over waren. Negatieve reacties kwamen wel voor maar dat betrof vaak leerlingen van wie het de vraag was of die het niveau aankonden. Van *Werken met referentieniveaus* zijn er geen effecten op leerlingen gevonden. Van *Gebruik maken van opbrengsten* zijn geen significante effecten gevonden, maar zien we in het algemeen wel een positieve tendens in het schooljaar waarin de leerkracht deelneemt. Daarna ebt het effect grotendeels weg.

Oorzaken liggen waarschijnlijk zowel in de opzet en de intensiteit van de interventie als inhoudelijk in het onderwerp.

In hoofdstuk 2 wordt een opsomming gegeven van kenmerken van effectieve interventies. Om effectief te kunnen zijn moet er door de deelnemers voldoende tijd aan de interventie besteed worden. De totale begeleide tijd van *Gebruik maken van opbrengsten* is minimaal 40 uur ($4 * 2,5$ uur + $5 * 2$ uur voor de bijeenkomsten en voor het maken van de opdrachten is ook minimaal 20 uur nodig). Gezien de reacties van de deelnemers is die tijd door het overgrote deel ook daadwerkelijk aan de interventie besteed. Voor het deelproject over referentieniveaus is 12 uur berekend.

Door de afwisseling van bovenschoolse en schoolgebonden bijeenkomsten die in beide interventies is gehanteerd, die respectievelijk meer op training en meer op begeleiding waren gericht, is gezorgd voor een afwisseling van kennis en vaardigheden en voor een afwisseling van algemene en schoolspecifieke invulling.

Attitude-aspecten, inclusief weerstanden, zijn zowel bovenschools als schoolgebonden in *Gebruik maken van opbrengsten* besproken in discussieopdrachten waarbij zowel gewerkt is in groepjes van leerkrachten van verschillende scholen en verschillende jaargroepen als in teamverband. In *Werken met referentieniveaus* is dit niet zo expliciet aan de orde geweest.

Voor de uitvoering van *Gebruik maken van opbrengsten* is een samenwerking aangegaan met 2 onderwijsadviesorganisaties, waarbij overeengekomen is dat zij personeel zou inzetten dat inhoudelijk kennis had van de verschillende aspecten van opbrengstgericht werken en praktische vaardigheden bezat om leerkracht individueel en teams daarbij te begeleiden. Deze begeleiders zijn vooraf getraind en bovendien bij elke bovenschoolse bijeenkomst aanwezig geweest om er zorg voor te dragen dat zij goed op de hoogte waren van wat daar aan de orde is geweest. De deelnemers aan *Gebruik maken van opbrengsten* hebben zowel wat betreft het analyseren als wat betreft het gedifferentieerd werken in de groep gelegenheid gehad om te oefenen en daarop feedback gekregen. Analyse-rapporten zijn besproken evenals de observatie in de groep.

Samenvattend kan gesteld worden dat *Gebruik maken van opbrengsten* in opzet en inhoud voldoet aan de meeste kenmerken van een effectieve interventie. Voor *Werken aan referentieniveaus* geldt dat in mindere mate, mede vanwege de beperkte tijd die er voor die interventie was. Minder sterke punten van *Gebruik maken van opbrengsten* waren de mate waarin er aandacht is besteed aan de inbedding in het team, de sturing door de directie en de borging in de toekomst. De inbedding in het team is vanwege het onderzoek niet in de interventie aangezet. Het onderzoeksdesign is immers zo opgezet dat de niet-deelnemende-bouw deel uitmaakt van de controlegroep voor andere scholen. Wanneer de verworvenheden van de interventie meteen breed gedeeld zouden worden, verliest het design zijn vergelijkingsmogelijkheden.

Naast de opzet en intensiteit van de interventie kan ook de inhoud, het onderwerp reden zijn voor het ontbreken van resultaten op leerlingniveau. Dit kan simpelweg te maken hebben met

het feit dat referentieniveaus verder van het dagelijks werk in de klas af liggen dan opbrengstgericht werken. In al het effectiviteitonderzoek zien we het gegeven terug dat factoren op leerkracht en klasniveau een grotere rol spelen dan factoren op school of stelselniveau. In *Werken aan referentieniveaus* is wel geprobeerd om juist de abstracte referentieniveau te ‘vertalen’ naar de klassensituatie, maar het onderzoek laat wel zien dat dat moeilijk is; leerkrachten hebben weinig idee van de hoogte en de inhoud van de niveaus. Desgevraagd stellen zij veel hogere doelen dan de wettelijke niveaus. Ten aanzien van opbrengstgericht werken is er voldoende onderzoek dat laat zien dat daar in principe wel een verbetering van leerlingprestaties van te verwachten is. Mogelijk verbeteren leerkrachten ten gevolge van *Gebruik maken van opbrengsten* niet voldoende op de bestandsdelen van opbrengstgericht werken die daarin essentieel zijn: een teamaanpak, kennis over op onderzoek gebaseerde differentiatiemogelijkheden en de vertaalslag van analyse naar onderwijsplanning en uitvoering. Tenslotte zijn de ingezette ontwikkelingen niet doorgezet en geborgd waardoor teams in het volgende schooljaar weer een stok achter de deur nodig hebben.

Vervolgonderzoek

Voor vervolgonderzoek is een wisselwerking nodig tussen de aanpak van deelprojecten. Voor de uitvoering van opbrengstgericht werken in de klas zou een casestudie aanpak zoals in *Differentiatie en excellentie* is gehanteerd meer inzicht geven is waar de cyclus stopt; wat plannen leerkrachten nu feitelijk en wat voeren ze daarvan uit en waarom? Voor het vervolg van *Differentiatie en excellentie* kan een aanpak zoals in *Gebruik maken van opbrengsten* meer zeggingskracht geven door de grotere aantallen en de mogelijkheid om de effecten na te gaan. Waarbij dat overigens bij de atypische doelgroep van excellente leerlingen wel veel complexer is dan bij een reguliere klas. De inhoud van *Werken met referentieniveaus* en *Gebruik maken van opbrengsten* is samengevoegd in een van de twee projecten die aan Streef waren toegevoegd vanuit TIER. Dat project wordt alleen uitgevoerd met leerkrachten van groep 4 en 5 en is zowel inhoudelijke als qua intensiteit een optelsom van *Werken met referentieniveaus* en *Gebruik maken van opbrengsten*. Over de resultaten van dat project wordt in de loop van 2014 gerapporteerd.

Aanbevelingen

De aanbevelingen per deelproject zijn te vinden in de deelrapporten. Gemeenschappelijk is de aanbeveling om benodigde kennis voor opbrengstgericht werken in de Pabo's aan te bieden: analyse-vaardigheden en kennis over op onderzoek gebaseerd differentiatiemogelijkheden. Het uitvoeren van verantwoord gedifferentieerd onderwijs kan niet verwacht worden van een pas afgestudeerde leerkrachten. Maar wanneer en hoe legt een gevorderde leerkracht daar dan een ‘proeve van bekwaamheid’ voor af?

Literatuur

- Betts, G.T. & Neihart, M. (1988). Profiles of the Gifted and Talented. *Gifted Child Quarterly*, 32(2), 248-253.
- Blok, H. (2004). Adaptief onderwijs: betekenis en effectiviteit. *Pedagogische Studien*, 81(1), 5-27.
- Creemers, B.P.M. & Kyriakides, L. (2008). *The Dynamics of Educational Effectiveness: A Contribution to Policy, Practice and Theory in Contemporary Schools*. London/New York: Routledge.
- Dekker, B., Krooneman, P.J., Brekelmans, J. & Groenwoud, M. (2012). *Onderwijsinspanningen taal en rekenen in PO, VO en MBO*. Amsterdam: Regioplan.
- Deunk, M.I. & Doolaard, S. (in voorbereiding). *Attitude en handelen van basisschoolleerkrachten met betrekking tot het verbeteren en borgen van leerlingresultaten*. Groningen: GION.
- Deunk, M. I., Doolaard, S., & Hofman, R. H. (2011). *Attitude en gedrag van basisschoolleerkrachten met betrekking tot het verbeteren van leerlingresultaten. Resultaten van de beginmeting bij leerkrachten*. Groningen: GION.
- Eekelen, I.M. van, Vermunt, J.D., & Boshuizen, H.P.A. (2006). Exploring teachers' will to learn. *Teaching and Teacher Education*, 22(4), 408-423.
- Elmore, R. (2003). *Knowing the right thing to do: School improvement and performance-based accountability*. WashingtonDC: NGA Center for Best Practices.
- Expertgroep Doorlopende Leerlijnen (2008). *Over de drempels met taal en rekenen*. Enschede: SLO.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen. (2009). *Een nadere beschouwing. over de drempels met taal en rekenen*. Enschede: Expertgroep Doorlopende Leerlijnen Taal en Rekenen.
- Fullan, M. (2007). *The new meaning of educational change*. New York/London: Teacher College Press.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Gerven, E. van, e.a. (2011). *Handboek hoogbegaafdheid*. Assen: Van Gorcum.
- Grift, W.J.C.M. van der (2010). *Ontwikkeling in de beroepsvaardigheden van leraren*. Oratie. Groningen: RUG.
- Hattie, J. A. C. (2003). Teachers make a difference. what is the research evidence? *Australian Council for Educational Research Annual Conference on: Building Teacher Quality*. Auckland: University of Auckland.
- Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London, New York: Routledge.
- Hill, H. C. (2007). Learning in the teaching workforce. *The Future of Children*, 17(1), 111-127.
- Hoogeveen, L., Hell, J. van, Mooij, T. & Verhoeven, L. (2004). *Onderwijsaanpassingen voor hoogbegaafde leerlingen*. Nijmegen: ITS.
- Inspectie van het Onderwijs (2012a). *De staat van het onderwijs. Onderwijsverslag 2010/2011*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs. (2010). *Opbrengstgericht werken in het basisonderwijs*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het onderwijs. (2012b). *Monitor verbetertrajecten taal en rekenen 2008/2009, 2009/2010 en 2010/2011*. Utrecht: Inspectie van het Onderwijs.
- Kuipers, J. (2008). *Levelwerk*. Drachten: Cedin/Eduforce.

- Ledoux, G., Blok, H., & Boogaard, M. (2009). *Opbrengstgericht werken. over de waarde van meetgestuurd onderwijs*. Amsterdam: SCO-Kohnstamm Instituut.
- Leune, J.M.G. (2001). *Onderwijs in verandering; reflecties op een dynamische sector*. Groningen: Wolters-Noordhoff.
- Mandinach, E.B., Honey, M. & Light, D. (2006). *A theoretical framework for data-driven decision making*. Paper presented at the AERA: San Francisco.
- Meelissen, M.R.M., Netten, A., Drent, M., Punter, R.A., Droop, M. & Verhoeven, L. (2012) *PIRLS- en TIMSS-2011 : trends in leerprestaties in lezen, rekenen en natuuronderwijs*. Nijmegen: Radboud Universiteit Nijmegen / Enschede: Universiteit Twente
- Ministerie van Onderwijs (2007). *Kwaliteitsagenda Scholen voor morgen*. Den Haag: MOCW.
- Ministerie van Onderwijs (2011). *Actieplan Basis voor Presteren*. Den Haag: MOCW.
- OECD (2010). *The high costs of low educational performance*. OECD.
- Olszewski-Kubilius, P. (1995). A summary of research regarding early entrance to college. *Roeper Review*, 18(2).
- Onderwijsraad (1999). *Zeker weten. Leerstandaarden als basis voor toegankelijkheid*. Den Haag: Onderwijsraad.
- Onderwijsraad (2007). *Presteren naar vermogen*. Den Haag: Onderwijsraad.
- Onderwijsraad (2008). *Partners in onderwijsopbrengst*. Den Haag: Onderwijsraad.
- PO Raad (2012). *Op opbrengsten gericht*. Utrecht: PO Raad
- Renzulli, J. S. & Reis, S. M. (1994). Research related to the Schoolwide Enrichment Triad model. *Gifted Child Quarterly*, 38(1), 7-20.
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart and Winston.
- Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford: Pergamon.
- Scheerens, J., Luyten, H. & Steen, R. (2007). *Review and Meta-analyses of School and Teaching Effectiveness*. Enschede: University of Twente.
- Schildkamp, K., & Kuiper, W. (2010). Data-informed curriculum reform: Which data, what purposes, and promoting and hindering factors. *Teaching and Teacher Education*, 26(3), 482-496.
- Schilt-Mol, T. van & Vijfeijken, M. van (2011). *Taal- en leesverbetertrajecten in het basisonderwijs. De opbrengsten van 3 jaar werken aan taal en lezen*. Tilburg: IVA.
- Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, 17(2), 213-226.
- Vaughn, V.L., Feldhusen, J.F. & Asher, J.W. (1991). Meta-analyses and review of research on pull-out programs in gifted education. *Gifted Child Quarterly*, 35(2), 92-98.
- Verloop, N. (2003). *Onderwijskunde*. Groningen: Noordhoff Uitgevers.
- Visscher, A., & Ehren, M. (2011). *De eenvoud en complexiteit van Opbrengstgericht werken*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- Wahlstrom, K.L. & Louis, K.S. (2008). How teachers experience principal leadership: The roles of professional community, trust, efficacy, and shared responsibility. *Educational Administration Quarterly*, 44(4), 458
- Waterhouse (2006). Inadequate evidence for Multiple Intelligences, Mozart effect, and Emotional Intelligence Theories. *Educational Psychologist*, 41, 247-255.

Bijlage

Relevante variabelen uit de leerkrachtvragenlijst

	Omschrijving	Range en betrouwbaarheid
Samenwerken	Schaal van 10 items op basis van Wahlstrom & Louis (2008) waarin gevraagd is hoe vaak leerkrachten observeren, feedback geven, uitwisselen en overleggen.	1=(vrijwel) nooit tot 5=1x per week <i>In 2010: $\alpha=0,70$</i>
Beroepsattitude	Schaal van 5 items op basis van Van Eekelen e.a. (2006) om de flexibiliteit waarmee leerkrachten hun lesgeven aanpakken en de mate waarin zij zich verantwoordelijk voelen voor de voortgang van de leerlingen in kaart te brengen. Voorbeelditem: Ik trek het mijzelf aan als een leerling zwak presteert.	1=helemaal oneens tot 5=helemaal eens <i>In 2010: $\alpha=0,69$</i>
Metacognitie	Schaal van 8 items die aangeeft in hoeverre de leerkracht belang hecht aan het aanleren van metacognitieve vaardigheden bij leerlingen	1=helemaal oneens tot 5=helemaal eens <i>In 2010: $\alpha=0,75$</i>
Explicitering doelen	Samengestelde variabele die aangeeft of de leerkracht weet welk niveau hij/zij aan het eind van het jaar met de leerlingen wil behalen en in hoeverre dat geëxpliciteerd is op papier, in de methode, in overleg met team of intern begeleider.	1= Nee, niveau is niet geëxpliciteerd. 2= Ja, de leerkracht weet het voor zichzelf of het staat in de methode. 3=Ja, de leerkracht heeft het voor zichzelf of samen met de intern begeleider opgeschreven of besproken. 4=Ja, het is met het team of op een andere manier besproken/opgeschreven.
Attitude opbrengstgericht werken	Index van 4 items die meet hoe opbrengstgericht een leerkracht is. Voorbeelditem: Ik streef altijd naar een hoger niveau voor mijn leerlingen, ook als ze al goed presteren.	1=weinig tot 4=veel opbrengstgericht
Breed kader	Index die aangeeft hoe breed het kader is waarin de leerkracht de toetsresultaten bekijkt.	1=Geen analyse 2=Basale analyse: de Ik maakt een nadere analyse van de resultaten van de zwakke leerlingen en kijkt naar de eerdere toetsresultaten van de leerlingen. 3=Naast de basale analyse vergelijkt de leerkracht de toetsresultaten soms of meestal met de landelijke gegevens en gebruikt minstens 1 van de andere vergelijkingen 4=Uitgebreide analyse
Reflectie	Schaal van 6 items die aangeeft op welke manier de leerkracht terugkijkt na de afname van een toets. Voorbeelditem: Ik kijk terug naar hoe ik instructie	1=nooit, 2=soms, 3=altijd <i>In 2010: $\alpha=0,70$</i>

	heb gegeven	
Basisvaardigheden	Schaal van 3 items die aangeeft hoe belangrijk de leerkracht de nadruk op basisvaardigheden vindt. Voorbeelditem: Ik vind het belangrijk om in het onderwijs een sterke nadruk op rekenen te leggen.	1=helemaal oneens tot 5=helemaal eens <i>In 2010: $\alpha=0,70$</i>
Rapport	Index die aangeeft van hoeveel en welke rapportagemogelijkheden een leerkracht gebruikt.	1=De lk maakt niet systematisch gebruik van de rapportagemogelijkheden. 2=Systematisch gebruik van een van de rapportagemogelijkheden (leerling of groep). 3=Systematisch gebruik van de mogelijkheden op groeps- en leerlingniveau. 4=De lk gebruikt daarnaast de rapportagemogelijkheden van foute antwoorden en/of leerlingen met een eigen leerlijn.
Prestatie-informatie	Schaal van 4 items die meet in welke mate leerkrachten het belangrijk vinden prestatie-informatie te delen met leerlingen. Voorbeelditem: Ik vind het goed om lln te vertellen wat ik hen de komende tijd ga leren.	1=helemaal oneens tot 5=helemaal eens <i>In 2010: $\alpha=0,69$</i>
Dif_niveau	Samengestelde variabele op basis van het aantal niveaugroepen en het aantal leerlingen op een individuele leerlijn. Door de manier van hercoderen geldt <i>niet</i> hoe meer niveaus, hoe beter; bij heel veel individuele leerlijnen en niveaus in de groep krijgt men juist een wat lagere score, omdat blijkt dat men op die manier niet profiteert van de voordelen voor leerlingen en leerkrachten van een (gedeeltelijk) klassikale aanpak.	1=Geen differentiatie in niveaus 2=Matige differentiatie in niveaus 3=Optimale differentiatie in niveaus
Dif_actie	Index van 7 items die meten hoeveel differentiatie-activiteiten een leerkracht neemt als uit een toets blijkt dat een leerling de stof nog niet of helemaal heeft begrepen. Voorbeelditem: ik bied de leerling verrijkingsstof aan	1=nooit tot 3=meestal
Dif_attitude	Index van 4 items die meten hoe de leerkracht over differentiatie denkt. Voorbeelditem: Ik vind dat de leerkracht zijn of haar aandacht moet verdelen tussen zwakke en excellente leerlingen	1=helemaal oneens tot 5=helemaal eens