

WAT SCHOLEN TOEVOEGEN
advies



WAT SCHOLEN TOEVOEGEN

Colofon

De Onderwijsraad is een onafhankelijk adviescollege, opgericht in 1919. De raad adviseert, gevraagd en ongevraagd, over hoofdlijnen van het beleid en de wetgeving op het gebied van het onderwijs. Hij adviseert de ministers van Onderwijs, Cultuur en Wetenschappen en van Landbouw, Natuurbeheer en Visserij. De Eerste en Tweede Kamer der Staten-Generaal kunnen de raad ook om advies vragen. Gemeenten kunnen in speciale gevallen van lokaal onderwijsbeleid een beroep doen op de Onderwijsraad.

De raad gebruikt in zijn advisering verschillende (bijvoorbeeld onderwijskundige, economische en juridische) disciplinaire aspecten en verbindt deze met ontwikkelingen in de praktijk van het onderwijs. Ook de internationale dimensie van educatie in Nederland heeft steeds de aandacht.

De raad adviseert over een breed terrein van het onderwijs, dat wil zeggen van voorschoolse educatie tot aan postuniversitair onderwijs en bedrijfsopleidingen. De producten van de raad worden gepubliceerd in de vorm van adviezen, studies en verkenningen. Daarnaast initieert de raad seminars en websitediscussies over onderwerpen die van belang zijn voor het onderwijsbeleid.

De raad bestaat uit zeventien leden die op persoonlijke titel zijn benoemd.

Advies *Wat scholen toevoegen. Instrumenten voor de bepaling van de toegevoegde waarde van het basisonderwijs*, uitgebracht aan de minister van Onderwijs, Cultuur en Wetenschappen.

Nr. 20030222/732, juli 2003.

Uitgave van de Onderwijsraad, Den Haag, 2003.
ISBN 90-77293-07-8

Bestellingen van publicaties:

Onderwijsraad
Nassaulaan 6
2514 JS Den Haag
email: secretariaat@onderwijsraad.nl
(070) 310 00 00 of via de website: www.onderwijsraad.nl

Ontwerp en opmaak:

Maarten Balyon grafische vormgeving

Drukwerk:

Drukkerij Artoos

© Onderwijsraad, Den Haag
Alle rechten voorbehouden. All rights reserved.

WAT SCHOLEN TOEVOEGEN

**Instrumenten voor de bepaling van de toegevoegde waarde
van het basisonderwijs**

ONS KENMERK
2003022/732

UW KENMERK
PO/KB/03/8463

CONTACTPERSOON

DOORRIJESNUMMER

PLAATS / DATUM
Den Haag, 14 augustus 2003

ONDERWERP
Advies Wat scholen toevoegen

Aan de minister van Onderwijs, Cultuur en Wetenschappen
Mevrouw M.J.A. van der Hoeven
Postbus 25000
2700 LZ Zoetermeer

Mevrouw de minister,

Op 18 maart jongstleden heeft u de Onderwijsraad gevraagd een advies uit te brengen over de vaststelling van de toegevoegde waarde van het onderwijs in basisscholen.


Met genoegen biedt de Onderwijsraad u hierbij zijn advies *Wat scholen toevoegen* aan. Het advies is een uitwerking van de adviesvragen die zijn geformuleerd in bovengenoemde brief.

De raad heeft de adviesvragen geplaatst in de context van nieuwe verhoudingen die zijn ontstaan als gevolg van een grotere autonomie van scholen. De kern van die nieuwe verhouding is dat schoolbesturen in meerdere mate zelf verantwoordelijk zijn voor de kwaliteit van het gegeven onderwijs en voor het zichtbaar maken van de opbrengsten (toegevoegde waarde) van het onderwijs.

In dit advies stelt de raad dat de huidige praktijk van het zichtbaar maken van de opbrengsten niet adequaat is en om een verbetering vraagt. Beter is het om de opbrengsten vast te stellen met een systeem dat ten minste een begin- en een eindmeting omvat. Deze metingen zouden deel moeten uitmaken van het kwaliteitssysteem van de scholen.

De raad realiseert zich dat er nog heel wat ontwikkelwerk verricht moet worden en dat de aanbevelingen niet in één keer gerealiseerd kunnen worden. Voor de verdere uitwerking van de aanbevelingen van de raad stelt hij voor om uit te gaan van een meerjarig ontwikkeltraject.

Namens de Onderwijsraad,



prof. dr. A.M.L. van Wieringen
Voorzitter

drs. A van der Rest
Secretaris

ONDERWIJS raad

MASSAHLAAN 6
2514 JS DEN HAAG
TELEFOON 070 310 00 00
FAX 070 356 14 74
E-MAIL SECRETARIAAT@ONDERWIJSRAAD.NL
WEBSITE WWW.ONDERWIJSRAAD.NL

Inhoudsopgave

Samenvatting		9
1 Inleiding		13
1.1 Inleiding		13
1.2 Adviesvraag		13
1.3 Beleidscontext		15
1.3.1 Besturing		15
1.3.2 Interne schoolverbetering en externe kwaliteitsverantwoording: praktische samenhang		19
1.4 Aanpak		21
1.5 Opzet		22
2 Toegevoegde waarde		23
2.1 Opbrengst van onderwijs in het algemeen		23
2.2 Toegevoegde waarde als opbrengst van onderwijs		24
2.3 Opbrengst van onderwijs: de keuze van ontwikkelingsdomeinen en eindtoetsen		26
2.4 Zijn er alternatieven voor een beginmeting?		30
2.5 Ontwikkelingen in enkele landen		32
2.6 Conclusie		35
3 Onderwijs- en maatschappelijke aspecten		37
3.1 De neveneffecten en bezwaren		37
3.2 Reductie van ongewenste neveneffecten		40
3.3 Conclusies		42
4 Conclusies, aanbevelingen en beantwoording van de adviesvraag		43
4.1 Terugblik, conclusies en aanbevelingen		43
4.1.1 Veertien aanbevelingen		45
4.2 Beantwoording van de adviesvragen		46
Literatuur		50
Bijlage		B.1-53
Bijlage 1 Adviesaanvraag		B.1-53
Bijlage 2 Expertmeeting begintoets basisonderwijs op 20 februari 2003		B.2-57
Bijlage 3 Panelbijeenkomst toegevoegde waarde van het onderwijs in scholen op 6 mei 2003		B.3-59
Bijlage 4 Afkortingen		B.4-61

Samenvatting

Inleiding

Het eerste kabinet-Balkenende formuleerde in het regeerakkoord het voornemen om in het primair onderwijs een begintoets in combinatie met een eindtoets in te voeren, die ouders en leerlingen inzicht moet bieden in de kwaliteit van een school¹. De begintoets wordt daarbij gezien als een element van een meetsysteem voor het bepalen van de opbrengsten van de school. Ook in het regeerakkoord van het tweede kabinet-Balkenende wordt het voornemen van de overheid geuit om zich meer te richten op de kwaliteit van de output. Besturen moeten meer verantwoording afleggen over en inzicht geven in de opbrengsten van het onderwijs².

Huidige situatie

Ook nu al worden de opbrengsten van het basisonderwijs zichtbaar gemaakt aan ouders. Gegevens over de doorstroom, het aantal zittenblijvers, enzovoort worden vaak in de schoolgids vermeld, evenals de score op de Eindtoets Basisonderwijs van de Citogroep. Deze toets wordt bij circa 80% van de leerlingen van de basisschool afgenomen. In een rapport aan de school geeft de Citogroep de schoolscores, dat wil zeggen de gemiddelde prestaties op de verschillende onderdelen van de eindtoets en op de toets in z'n geheel, en vergelijkt deze met het landelijk gemiddelde. Deze schoolscore wordt soms (ten onrechte) gebruikt als een gegeven dat iets zegt over de opbrengsten van de school³. Om 'oneerlijke' vergelijkingen te voorkomen worden scholen ingedeeld in *families of schools*. Dit zijn groepen scholen die vergelijkbaar zijn op het punt van samenstelling van de leerlingenpopulatie. Een nog verdere verfijning van de scores wordt op verzoek eveneens aan de scholen geleverd. Het betreft dan een rapportage over subgroepen binnen een school. Voor de school is deze informatie bruikbaar in het kader van het school-interne kwaliteitsbeleid⁴. Als middel om de leeropbrengsten van de school vast te stellen is deze procedure minder geschikt. Het vaststellen van de netto leeropbrengst is een betere manier. In plaats van netto leerwinst wordt ook wel gesproken over de toegevoegde waarde van het onderwijs.

Adviesvraag

De minister van Onderwijs heeft de Onderwijsraad verzocht een advies uit te brengen over het vaststellen van de opbrengsten (= toegevoegde waarde) van het onderwijs in een school. Gevraagd wordt onder andere of het noodzakelijk is dat de begintoets een rol speelt bij de bepaling hiervan. Ook vraagt de minister een advies over de functie die het vaststellen van de toegevoegde waarde voor overheidssturing kan hebben.

1 TK 2001-2002, 28 375, nr. 5, p. 14.

2 TK 2002-2003, 28 637, nr. 19, p. 8.

3 De Eindtoets Basisonderwijs van de Citogroep is bedoeld als een tweede gegeven naast het advies van de basisschool over het type vervolgonderwijs van de individuele leerling.

4 Zie ook: Website Citogroep.

Beleidscontext

De raad plaatst de adviesvraag in de context van nieuwe verhoudingen die zijn ontstaan als gevolg van (de)regulering en verschuivingen in de autonomie van scholen. De kern van de nieuwe verhouding is dat schoolbesturen in meerdere mate zelf verantwoordelijk zijn voor de opbrengsten van het gegeven onderwijs. "Dit stelt eisen aan de transparantie van de kwaliteit die scholen leveren en vergt een versterking van de positie van ouders, deelnemers en andere partijen in het onderwijs." (Ministerie van Onderwijs, Cultuur en Wetenschappen, 2000, p. 40). Scholen en instellingen dienen in de eerste plaats inzicht te hebben in hun eigen bijdrage aan de ontwikkeling van leerlingen en voorts dienen zij verantwoording af te leggen over de opbrengsten van het onderwijs. Voor wat betreft het afleggen van verantwoording wordt een onderscheid gemaakt tussen de verticale verantwoording (via Inspectie naar de minister) en de horizontale verantwoording (naar ouders, afnemers en de lokale en regionale omgeving). Met betrekking tot deze laatste verantwoordingslijn is transparante informatie over het onderwijs ten behoeve van de schoolkeuze belangrijk. Hiermee wordt gepoogd de positie van ouders te versterken. Het zichtbaar maken van de toegevoegde waarde heeft binnen deze beleidscontext een wezenlijke functie.

Toegevoegde waarde

Toegevoegde waarde heeft betrekking op de bijdrage die een school levert aan de leerprestaties van een leerling. Het bepalen van toegevoegde waarde en het verbinden van een oordeel daaraan is echter een complexe zaak. Kinderen leren immers ook buiten de school. Er zal dus een scherp onderscheid gemaakt moeten worden tussen de totale vooruitgang die een lerende boekt (tussen twee meetmomenten in) en de vooruitgang die kan worden toegeschreven aan door de school ontplooiende activiteiten. Als de overheid de toegevoegde waarde van het onderwijs wil vaststellen, dan zullen een begin- en een eindmeting onderdeel zijn van het instrumentarium voor het vaststellen van de toegevoegde waarde van het onderwijs; ook andere beïnvloedende variabelen van buiten de school zullen zoveel mogelijk moeten worden meegenomen.

Een meetsysteem om opbrengsten vast te stellen kan het beste aansluiten bij hetgeen nu in scholen gebeurt. Toetsen en toetsgegevens worden op veel basisscholen verzameld in een leerlingvolgsysteem. Op basis van gegevens uit een leerlingvolgsysteem kan de school systematisch werken aan de kwaliteit van het onderwijs. De school kan vervolgens de resultaten in verband brengen met kenmerken van het gegeven onderwijs, het schoolbeleid, de aard van de leerlingenpopulatie, enzovoort. Zijn de resultaten beneden verwachting, dan is er alle reden om het schoolbeleid bij te stellen. Dit interne systeem als onderdeel van een uitgebreider kwaliteitszorgsysteem van een school geeft echter niet voldoende informatie over de opbrengsten van het gegeven onderwijs in vergelijking met andere scholen. Voor een adequaat intern systeem voor schoolverbetering en een verantwoordingsmodel richting ouders en overheid is een volgende stap nodig.

Versterken horizontale verantwoording

De invoering van een meetsysteem brengt een wijziging met zich mee van het huidige besturingsinstrumentarium gericht op het afleggen van verantwoording aan bijvoorbeeld overheid, bestuur, leerlingen en ouders. Het gaat in dit advies dus niet over een koppeling tussen de uitkomsten van het onderwijs en bijvoorbeeld de bekostiging, maar primair over de benutting van informatie over de toegevoegde waarde voor de verantwoording aan ouders, bestuur en overheid en secundair voor schoolintern gebruik. Het karak-

ter van de verticale verantwoordingslijn dient te veranderen indien de horizontale verantwoordingslijn wordt verstevigd. In het toezichtkader van de Inspectie van het Onderwijs zal de beleidskeuze van aanbod naar opbrengst moeten doorklinken. Voor de Inspectie betekent dit dat zij het (proportioneel) toezicht kan vereenvoudigen. Tevens kan zij de opbrengstindicatoren gebruiken voor een macrorapportage over de staat van het onderwijs.

Neveneffecten

In dit advies wordt eveneens ingegaan op vermoedens over neveneffecten van het vaststellen van de toegevoegde waarde van het onderwijs en bezwaren hiertegen. De raad concludeert dat ongewenste neveneffecten kunnen optreden indien eenzijdig de nadruk wordt gelegd op toegevoegde waarde. Daarbij kan gedacht worden aan zaken als een toename van selectieverschijnselen, een sterker accent op toetsen en op de groep leerlingen die zich bevindt op de grens tussen twee prestatieniveaus, demotivatie van leraren en een afname van de samenwerkingsbereidheid tussen scholen. De raad acht deze neveneffecten onaanvaardbaar.

Om ongewenste neveneffecten te reduceren stelt de raad voor om aan te sluiten bij datgene wat nu al in scholen gebeurt. Derhalve adviseert de raad, ten eerste, om aan te sluiten bij de bepalingen in de Wet op het primair onderwijs (WPO) over het kwaliteitsbeleid. Een Algemene Maatregel van Bestuur (AMvB) als bedoeld in artikel 13 van de WPO, eerste lid onder a, dient de horizontale verantwoordingslijn te verstevigen. Deze AMvB zal richtlijnen moeten bevatten over de vaststelling van de toegevoegde waarde van het onderwijs en over de presentatie van de toegevoegde waarde, samen met gegevens over de inrichting en het aanbod van de school, in de schoolgids⁵. En ten tweede adviseert de raad om op een inhoudelijk breed terrein toetsmateriaal te (laten) ontwikkelen en beschikbaar te stellen. Scholen kunnen dit gebruiken om aan de verplichtingen voortkomend uit de WPO en de genoemde AMvB te voldoen.

Scholen kunnen de gegevens tevens gebruiken voor het eigen schoolbeleid. Zij zijn zelf verantwoordelijk voor publicatie van de gegevens aan ouders, schoolbestuur en overheid. De publicatie wordt vergezeld van procesgegevens en een verantwoording van de uitslagen.

Aanbevelingen

De raad concludeert dat de huidige praktijk van het zichtbaar maken van de opbrengsten niet adequaat is en derhalve om verbetering vraagt. Nu wordt vaak als een van de indicatoren voor kwaliteit de schoolscore op de Eindtoets Basisonderwijs gebruikt. Het eenvoudige gebruik van dit outputgegeven als indicator van kwaliteit ziet de raad als te beperkt. Een meetsysteem met ten minste een begin- en een eindmeting maakt het beter mogelijk om opbrengsten van het onderwijs vast te stellen. Deze metingen zouden tegelijkertijd deel kunnen uitmaken van het kwaliteitssysteem van de school.

De raad pleit voor een toetsinstrumentarium voor alle leergebieden uit de kerndoelen. Hij realiseert zich dat daarvoor nog heel wat ontwikkelwerk nodig is en dat de uiteindelijke keuze een afweging is van kosten en baten. Vanwege de te verwachte complexiteit adviseert de raad te starten met ontwikkeling en gebruik van toetsen op de cognitieve leergebieden, gevolgd door de overige leergebieden. Verder adviseert de raad om de eerste

⁵ Indien nodig kan ook een bepaling in de WPO overwogen worden.

meting te laten plaatsvinden rondom het zesde levensjaar. Eerder meten kan wel, maar wordt sterk afgeraden.

Voorts stelt de raad voor om een ontwikkeltraject uit te zetten van vier à zeven jaar, waarin toetsen worden ontwikkeld in samenspraak met de betrokkenen en waarin tevens mogelijke neveneffecten worden bestudeerd. Een gedegen communicatie- en implementatieplan dient hieraan ten grondslag te liggen. Een pilot waarin één en ander wordt uitgeprobeerd en neveneffecten in kaart worden gebracht, dient deel uit te maken van dit ontwikkeltraject.

De raad constateert op basis van een secundaire analyse die is uitgevoerd op de data van het PRIMA-cohortonderzoek, dat toetsuitslagen door de jaren heen stabiel blijven. Op grond van deze constatering zijn er geen methodologische argumenten die een tussenmeting rechtvaardigen, wel praktische. De praktische argumenten hebben te maken met de grote uitval die tussen groep 1 en groep 8 plaatsvindt. Tussenmetingen dragen bij aan een verfijning van het systeem voor (interne) kwaliteitszorg. Het opzetten van een meetstelsel ten behoeve van interne schoolverbetering en ten behoeve van externe verantwoording kunnen elkaar dan positief beïnvloeden. Uit de analyses komt ook naar voren dat voor vaststelling van de toegevoegde waarde achtergrondvariabelen van de leerlingen dienen te worden meegenomen. De combinatie van achtergrondvariabelen en de uitkomst van een statistische bewerking van de begin- en de eindmeting geven een indicatie van de toegevoegde waarde van het onderwijs.

Toekomstbeeld

Voor de uitvoering stelt de raad zich in de toekomst een model voor waarbij er sprake is van een databank die gevuld wordt met een variatie aan begin- en eindtoetsen. Voor elk domein en voor elke leeftijdsgroep komen zo verschillende toetsen beschikbaar. Op deze manier wordt scholen een mogelijkheid geboden om toetsen te kiezen die bij de eigen onderwijskundige opvatting passen. Nodig is dat elke meting zogenoemde *kernitems* bevat. Dit zijn items die in elke toets voorkomen en waarvan de relatie met andere items bekend is. Enerzijds kan daardoor elke school zijn eigen toets samenstellen en anderzijds blijft de toegevoegde waarde van scholen vergelijkbaar. Om rekening te houden met levensbeschouwelijke en andere door de gemeenschap belangrijk geachte aspecten, adviseert de raad om een groep bestaande uit ouders, leraren, schoolleiders en vertegenwoordigers van de verschillende besturen- en managementorganisaties bij het opstellen van het implementatietraject te betrekken.

De raad adviseert voor afname van de toetsen een protocol op te stellen dat een correcte uitvoering garandeert. In personeelsformatieve zin zal rekening gehouden moeten worden met de extra taken die scholen moeten uitvoeren. De raad adviseert tevens om scholen te ondersteunen bij de statistische bewerking van testresultaten. Een nog nader te bepalen organisatie berekent de toegevoegde waarde, vergelijkt deze met die van andere scholen en rapporteert dit aan de school. Scholen blijven in een dergelijk model zelf verantwoordelijk voor publicatie van de gegevens. Deze publicatie wordt vergezeld van schooleigen procesgegevens en verantwoording van de uitslagen.

1 Inleiding

In het werkprogramma van de raad voor 2003 is de begintoets primair onderwijs opgenomen als een thema voor advisering. De directe aanleiding hiervoor is het beleidsvoornemen zoals geformuleerd in het Strategisch Akkoord van het eerste kabinet-Balkenende. Het thema begintoets heeft al een langere geschiedenis en moet worden geplaatst in de context van het bepalen van de toegevoegde waarde van het onderwijs in scholen. Het vaststellen van de toegevoegde waarde van het onderwijs komt voort uit de wens van de overheid om zicht te kunnen (blijven) houden op de resultaten die met het onderwijs worden bereikt. Door deregulering en autonomievergroting ligt de uitvoerende verantwoordelijkheid voor het onderwijs meer en meer bij scholen en schoolbesturen, terwijl de overheid wel de eindverantwoordelijkheid voor het onderwijs behoudt.

1.1 Inleiding

Paragraaf 1.2 van dit hoofdstuk beschrijft eerst de adviesvraag van de minister. Vervolgens geeft paragraaf 1.3 de beleidscontext weer waarbinnen de adviesvraag van de minister moet worden geplaatst. Voor een goed begrip van het thema begintoets worden de verantwoordelijkheden van de overheid ten opzichte van het onderwijs kort aangegeven. Het gaat daarbij om een balans tussen overheidsverantwoordelijkheid en verantwoordelijkheden en taken die bij maatschappelijke instituties, waaronder de onderwijsinstellingen, zijn gelegd. Paragraaf 1.4 beschrijft de aanpak die de raad heeft gevolgd om de vraag van de minister te beantwoorden. De opzet van dit advies wordt geschetst in paragraaf 1.5.

1.2 Adviesvraag

Het eerste kabinet-Balkenende formuleerde in het regeerakkoord het voornemen om in het primair onderwijs een begintoets in te voeren, die in combinatie met een eindtoets ouders en leerlingen inzicht moet bieden in de kwaliteit van een school⁶. De begintoets wordt daarin gezien als een element van een meetsysteem voor het bepalen van de opbrengsten van de school. Ook in het regeerakkoord van het tweede kabinet-Balkenende uit de overheid het voornemen om zich meer te richten op de kwaliteit van de output. Besturen moeten meer verantwoording afleggen over en meer inzicht geven in de opbrengsten van het onderwijs⁷.

Ook nu al worden de opbrengsten van het basisonderwijs zichtbaar gemaakt aan ouders en anderen. Gegevens over de doorstroom, het aantal zittenblijvers en dergelijke worden vaak in de schoolgids vermeld, evenals de score op de Eindtoets Basisonderwijs van de

6 TK 2001-2002, 28 375, nr. 5, p.14.

7 TK 2002-2003, 28 637, nr. 19, p.8.

Citogroep. Deze toets wordt bij circa 80% van de basisschoolleerlingen afgenomen⁸. De Citogroep verzamelt de individuele scores en bewerkt deze naar schoolniveau. In een rapport aan de school geeft de Citogroep de schoolscores, dat wil zeggen de gemiddelde prestaties op de verschillende onderdelen van de eindtoets en op de toets in z'n geheel, en vergelijkt deze met het landelijk gemiddelde. Deze schoolscore wordt soms (ten onrechte) gebruikt als een gegeven dat iets zegt over de opbrengsten van de school⁹. Om 'oneerlijke' vergelijkingen te voorkomen worden scholen ingedeeld in *families of schools*. Dit zijn scholen die qua leerlingengroep vergelijkbaar zijn. Een nog verdere verfijning van de scores wordt op verzoek eveneens aan de scholen geleverd. Het betreft dan een rapportage over subgroepen binnen een school. Voor de school is deze informatie bruikbaar in het kader van het interne kwaliteitsbeleid¹⁰. Als middel om de leeropbrengsten van de school vast te stellen is deze procedure minder geschikt.

Het vaststellen van de netto leeropbrengst is een betere manier (zie ook hoofdstuk 2). In plaats van netto leerwinst wordt ook wel gesproken over de toegevoegde waarde van het onderwijs. Toegevoegde waarde zegt iets over de bijdrage die een school levert aan de ontwikkeling van een leerling. Het bepalen van toegevoegde waarde en het verbinden van een oordeel daaraan is echter een complexe zaak. Kinderen leren immers ook buiten de school. Dus zal een scherp onderscheid gemaakt moeten worden tussen de totale vooruitgang die een lerende boekt (tussen twee meetmomenten in) en de vooruitgang die kan worden toegeschreven aan door de school ontplooiende activiteiten. In een systeem om toegevoegde waarde te meten moeten daarom idealiter ook andere beïnvloedende variabelen van buiten de school worden opgenomen¹¹.

De brief van de minister waarin de adviesvraag is verwoord is als bijlage 1 opgenomen. De vragen van de minister zijn de volgende.

1 *Is het wenselijk dat de toegevoegde waarde van een school wordt bepaald en wat is de functie daarvan?*

Paragraaf 1.3 behandelt globaal het mogelijke antwoord op deze vraag. De minister vraagt bij de beantwoording in te gaan op de voor- en nadelen die zijn verbonden aan het vaststellen van de toegevoegde waarde, alsmede op de effecten van meten en bepalen ervan. In het verlengde hiervan vraagt de minister welke domeinen of leergebieden aan de orde zouden moeten zijn bij het vaststellen van de toegevoegde waarde.

Aansluitend zal in dit advies tevens worden ingegaan op de wijze van bepaling van de toegevoegde waarde. De vraag die daarbij aan de orde komt is de volgende:

2 *Is het noodzakelijk dat een begintoets voor vierjarigen een rol speelt in de bepaling van de toegevoegde waarde?*

Om deze vraag te beantwoorden zal onder meer nagegaan worden of er alternatieven voorhanden zijn voor de bepaling van de toegevoegde waarde. Daarnaast wordt in dit

8 En bij circa 10 à 15% van de leerlingen wordt aan het einde van de basisschoolperiode een andere betrouwbare en genormeerde toets afgenomen (zie ook *Inspectie van het Onderwijs, 2003*).

9 De Eindtoets Basisonderwijs van de Citogroep is bedoeld als een tweede gegeven naast het advies van de basisschool over het type vervolgonderwijs van de individuele leerling.

10 Zie ook: Website Citogroep.

11 Het begrip toegevoegde waarde van het onderwijs en het gebruik van dit begrip worden in hoofdstuk 2 beschreven.

advies aangegeven onder welke voorwaarden toegevoegde waarde van het onderwijs kan worden vastgesteld.

Ten slotte handelt dit advies over de functie die het vaststellen van de toegevoegde waarde voor overheidssturing kan hebben. Gegevens over de kwaliteit van het onderwijs krijgen immers pas betekenis indien ze worden ingebed in een beleidscontext met verschillende beleidsinstrumentaties. De vraag die hier betrekking op heeft is:

- 3 *Welke vormen van beleidsinstrumentatie kunnen door de overheid worden verbonden met een systeem van toegevoegde waarde waarvan een begintoets onderdeel is?*

De adviesaanvraag van de minister is te plaatsen in de context van de nieuwe verhoudingen die zijn ontstaan tussen overheid, instellingen en ouders als gevolg van deregulering en autonomievergroting. In paragraaf 1.3 wordt hier verder op ingegaan.

1.3 Beleidscontext

De beleidscontext wordt in deze paragraaf ten eerste gezien tegen de achtergrond van de overheidsverantwoordelijkheid voor het onderwijs. Deze verantwoordelijkheid omvat onder meer:

- het bewaken en stimuleren van de kwaliteit van het onderwijs;
- het voorzien in een basisbekostiging die voldoet voor een kwalitatief goed en toegankelijk initieel onderwijs; en
- het bewaken van een doelmatige en efficiënte besteding van de publieke middelen (Onderwijsraad, 2001d).

Paragraaf 1.3.1 gaat in op de besturingsfilosofie van waaruit deze verantwoordelijkheden worden geoperationaliseerd.

Ten tweede wordt de beleidscontext gezien vanuit het perspectief van de school.

Paragraaf 1.3.2 beschrijft de door de overheid ingezette beleidsinstrumenten en de uitwerking daarvan op schoolniveau.

1.3.1 Besturing

De operationalisering van de overheidsverantwoordelijkheid wordt sinds het begin van de jaren tachtig van de vorige eeuw beïnvloed door opvattingen omtrent deregulering en autonomie van de instellingen. Als gevolg van een serie beleidsmaatregelen is de verhouding tussen overheid, schoolbesturen en instellingen veranderd. Feitelijk heeft er de afgelopen decennia zowel regulering als deregulering plaats gevonden. Zo is in het primair onderwijs het meest gedereguleerd op het financiële en materiële vlak en het meest gereguleerd op het onderwijsinhoudelijke vlak (proces, aanbod en opbrengsten). Op bestuurlijk-organisatorisch gebied is vooral ten aanzien van de thema's 'positie van de deelnemer in relatie tot de kwaliteitszorg' en 'schaalgrootte' gereguleerd. Deze mix van regulering en deregulering heeft geleid tot zowel autonomievergroting als -verkleining. Er wordt dan ook wel gesproken over autonomieverschuivingen (Hooge, 1998).

Tegen de achtergrond van het begrip toegevoegde waarde zal in het toezicht door de overheid ook een accentverschuiving plaats moeten vinden van aanbods- naar opbrengstcontrole. In het recent door de minister goedgekeurde Toezichtkader primair onderwijs van de Inspectie van het Onderwijs zijn de onderwerpen *Onderwijs en leren*, *Zorg voor kwaliteit* en *Opbrengsten* als domeinen van het waarderingskader beschreven. De eerste twee domeinen zijn in het Toezichtkader zorgvuldig uitgewerkt. Voor het domein *Opbrengsten* is dit veel minder het geval. In plaats van een verschuiving is in het Toezichtkader eerder sprake van een cumulatie van onderwerpen.

De kern van de nieuwe verhouding is dat schoolbesturen in meerdere mate zelf verantwoordelijk zijn voor de opbrengst van het gegeven onderwijs. "Dit stelt eisen aan de transparantie van de kwaliteit die scholen leveren en vergt een versterking van de positie van ouders, deelnemers en andere partijen in het onderwijs." (Ministerie van Onderwijs, Cultuur en Wetenschappen, 2000, p. 40). Scholen en instellingen dienen in de eerste plaats zelf inzicht te hebben in hun bijdrage aan de ontwikkeling van leerlingen en voorts dienen zij verantwoording af te leggen over de kwaliteit van het onderwijs.

In het actieprogramma *Kwaliteit aan de basis* (TK 2000-2001, 27 400 VIII, nr. 92) heeft het toenmalige kabinet de ideeën uit eerdere beleidsbrieven over deregulering en marktwerking verder uitgewerkt. Het uiteindelijke doel is een onderwijspraktijk waarin scholen voldoende beleidsruimte hebben en de overheid richting geeft. Scholen leggen verantwoording af over het beleid en de resultaten die men heeft behaald aan de deelnemers van het onderwijs, de directe omgeving van de school en de rijksoverheid. Het afleggen van verantwoording aan de rijksoverheid wordt hieronder verder uitgewerkt onder de subtitel *Toezicht door de overheid*. Het zich verantwoorden tegenover de ouders wordt verder behandeld onder de subtitel *Transparantie voor ouders en besturen*.

De raad plaatst de adviesvraag tegen de achtergrond van de gewijzigde bestuurlijke verhouding en de wens naar transparantie. Het gaat in dit advies dus niet over een koppeling tussen de uitkomsten van het onderwijs en bijvoorbeeld bekostiging, maar over de benutting van informatie over de toegevoegde waarde voor verantwoording aan ouders, bestuur en overheid en voor schoolintern gebruik. De raad maakt onderscheid tussen verticale verantwoording (via Inspectie naar de minister) en horizontale verantwoording (naar ouders, afnemers en de lokale en regionale omgeving)¹².

Om een en ander te realiseren werken scholen, branche en Inspectie verder aan de ontwikkeling van (instrumenten voor) de kwaliteitszorg door scholen. De Inspectie vervult hierin namens de overheid een rol die bijdraagt aan een nieuw evenwicht tussen de (de)regulerende overheid en de zelfregulerende onderwijsinstellingen.

12 De eerste vorm is de verticale verantwoordingslijn: verantwoording van de school tegenover rijk, gemeente en inspectie, die in naam van het algemeen belang optreden. Bij deze verantwoording gaat het om overheden of overheidsinstanties die een verantwoordelijkheid hebben ten aanzien van zowel het openbaar als het bijzonder onderwijs. Ingeval van horizontale verantwoording wordt door de school min of meer rechtstreeks verantwoording afgelegd tegenover de maatschappij of de (plaatselijke) gemeenschap. Aan een dergelijke controlelijn kan op verschillende manieren gestalte worden gegeven en er zijn verschillende vertegenwoordigers van de gemeenschap denkbaar. Voor het openbaar onderwijs kan het gaan om een representant van de staat (de minister) of van de plaatselijke bevolking in de vorm van een representatief orgaan (bijvoorbeeld de gemeenteraad). Voor het bijzonder onderwijs gaat het om een vertegenwoordiger van een lokale particuliere gemeenschap, waartoe uiteraard ook de ouders moeten worden gerekend. Deze gemeenschapsvertegenwoordigers kunnen in het schoolbestuur zitting hebben. Zowel bij het openbaar als het bijzonder onderwijs kunnen ouders ook op andere manieren hun controlefunctie uitoefenen. In het primair onderwijs is de aandacht vooral gericht op de verticale component. De horizontale component blijft, ondanks het bestaan van de schoolgids, achter. (Onderwijsraad, 2002c, p.96).

Toezicht door de overheid

Het toezicht op onderwijs is ingevolge artikel 3 van de Wet op het onderwijstoezicht (WOT) door de overheid opgedragen aan de Inspectie van het Onderwijs¹³. Voor de uitvoering van het toezicht zijn door de Inspectie verschillende instrumenten ontwikkeld. De ontwikkeling van deze instrumenten vindt eveneens plaats tegen de achtergrond van de WOT. Met de invoering van deze wet in 2002 is er een basis gelegd voor toezicht dat is aangepast aan de nieuwe verhoudingen. Zoals hiervoor al is weergegeven is de kern van de nieuwe verhouding dat schoolbesturen zelf verantwoordelijk zijn voor de opbrengsten van het gegeven onderwijs. Voor de Inspectie betekent dit dat haar rol, gelet op de eis van proportioneel toezicht, aansluit bij de verantwoordelijkheid die scholen zelf dragen. Schoolbesturen zijn immers zelf verantwoordelijk voor het formuleren van doelen op het gebied van onderwijskwaliteit, de realisatie en bewaking ervan en de publieke verantwoording erover. Conform artikel 12 van de WOT is de werkwijze voor kwaliteits-toezicht vastgelegd in een of meer toezichtkaders.

Onderdeel van het recent door de minister goedgekeurde toezichtkader is een waarde-ringskader dat wordt gehanteerd voor de beoordeling van scholen. Eén van de domeinen waar de Inspectie naar kijkt is de opbrengsten van het onderwijs in de scholen. Om een beeld te kunnen geven van de opbrengsten is de Inspectie nu nog afhankelijk van de gegevens die door de scholen worden verzameld. Zij merkt hierover op dat zij van circa 30% van de scholen niet kan aangeven of de opbrengsten *gedurende* de basisschoolperiode van voldoende niveau zijn omdat de gegevens hierover ontbreken (Inspectie van het Onderwijs, 2003)^{14,15}.

Transparantie voor ouders en besturen

De veronderstelling achter de hierboven beschreven beleidscontext is dat de transparantie van het onderwijs wordt verbeterd door aan deelnemers objectieve informatie te verstrekken over de opbrengsten van het onderwijs. Ouders krijgen daarmee de mogelijkheid om de schoolkeuze mede op basis van dergelijke informatie te baseren. Voor een deel van de consumenten (ouders, leerlingen) is het onderwijs namelijk (vooralnog) een 'ervaringsgoed'¹⁶. En, belangrijker dan dat, ouders kunnen hun betrokkenheid bij het schoolbeleid baseren op deze kwaliteitsgegevens. Overigens zet de overheid ten aanzien van de kwaliteitsinformatie nu ook al verschillende instrumenten in waarmee de horizontale verantwoordingslijn wordt versterkt (schoolgidsen, kwaliteitskaarten, openbare informatie over inspectiebeoordelingen, enzovoort). Het raadplegen van deze informatie is echter voor veel ouders allerminst vanzelfsprekend¹⁷.

13 Stb. 2002-387.

14 Uit de gegevens van de Inspectie blijkt ook dat ruim 90% van de scholen voldoende gegevens beschikbaar heeft over de opbrengsten aan het einde van de basisschool.

15 Overigens wordt opgemerkt dat scholen die opbrengstgegevens gedurende of aan het einde van de basisschoolperiode beschikbaar hebben, niet per definitie beschikken over een adequaat kwaliteitszorgsysteem. De Inspectie merkt hierover op dat "slechts dertig procent van de basisscholen beschikt over een vorm van kwaliteitszorg die systematisch en cyclisch genoemd mag worden" (Inspectie van het onderwijs, 2003, p. 108).

16 In economische termen wil dit zeggen dat de kwaliteit van een goed (onderwijs in dit geval) pas duidelijk wordt op het moment dat het geconsumeerd wordt. Zie ook: Onderwijsraad, 2001b.

17 Zie ook Onderwijsraad, 2002c. In dit advies constateert de raad dat, ondanks het bestaan van schoolgids, het zwaartepunt van de verantwoordingslijn thans ligt bij de verticale component. Naarmate de horizontale verantwoordingslijn (= de verantwoordingslijn naar ouders) beter is uitgewerkt, kan de verticale verantwoording van karakter veranderen. Dit wil zeggen dat de overheid haar ideeën over deregulering, autonomievergroting en verantwoorden dan eveneens zal moeten doorvoeren in het toezicht dat zij voorstaat. In het toezicht op onderwijs zal derhalve de keuze van controle op het aanbod naar controle op de opbrengsten moeten doorklinken.

Voorbeeld: Weergave gegevens over de opbrengsten van het onderwijs door scholen.

Scholen beschrijven in hun schoolgids hoe zij het onderwijs op hun school vormgeven en welke resultaten zij behalen. Zij kunnen daarbij voor elk leergebied, naast de beschrijving van het aanbod, aan ouders duidelijk maken welke resultaten zij bereiken, ook in vergelijking met scholen met een vergelijkbare populatie. Ouders worden op deze manier enigszins geïnformeerd over de prestaties van een school. Scholen kunnen daarbij ook aangeven wat zij de komende jaren beleidsmatig gaan ondernemen. De beleidsvoornemens kunnen dan met resultaatgegevens worden onderbouwd.

Een voorbeeld uit een schoolgids van een basisschool uit Den Haag

Tussenopbrengsten van het onderwijs: rekenen en wiskunde in groep 4.

Score volgens Citoleerlingvolgsysteem:

Niveau A:	58%
Niveau B:	17%
Niveau C:	17%
Niveau D:	0%
Niveau E:	8%

Niveau C is landelijk gemiddeld, B en A meer dan gemiddeld, bij niveau D is er sprake van enige achterstand, bij niveau E een half jaar of meer achterstand. Leerlingen die lager scoren dan niveau C krijgen extra aandacht middels een handelingsplan en een speciaal programma. Als 10% of meer een E-score heeft in een groep dan is dat onder de norm van de Inspectie.

Commentaar op deze presentatie van gegevens

Deze percentages zeggen weinig over de opbrengst ten gevolge van de kwaliteit van de school. Het beginniveau van de leerlingen en buitenschoolse invloeden zijn immers niet verdisconteerd in de percentages.

Conclusie

De raad is van mening dat de huidige sturing van het onderwijs vooralsnog onvoldoende relatie heeft met de prestaties van een school. Het is mede om deze reden wenselijk om toegevoegde waarde van onderwijsinstellingen vast te stellen. De toegevoegde waarde kan dan dienen als één van de indicatoren van kwaliteit.

Het geven van een valide indicatie van de opbrengsten van het onderwijs kan derhalve een functie vervullen binnen een door de overheid geformuleerd beleidskader. Door deregulering in het onderwijs is de autonomie van onderwijsinstellingen op enkele gebieden toegenomen; gezocht wordt naar een nieuw evenwicht tussen regulering, deregulering en het afleggen van verantwoording. De zorg voor een goede kwaliteit van het onderwijs is een gezamenlijke verantwoordelijkheid van onderwijsinstellingen, burgers en overheid waarbij iedere partij een rol speelt die samenhangt met zijn positie. Het zichtbaar maken van de toegevoegde waarde van het onderwijs door scholen kan binnen deze situatie wenselijk zijn.

1.3.2 Interne schoolverbetering en externe kwaliteitsverantwoording: praktische samenhang

Gegevens over de opbrengsten van scholen hebben geen betekenis op zich: die betekenis krijgen ze pas als ze zijn ingebed in een school- en een beleidscontext. In dit advies gaat het om de vraag of verzameling en eventuele publicatie van gegevens over de opbrengsten van scholen, in de vorm van toegevoegde waarde, een adequaat beleidsinstrument is als het gaat om sturing op de kwaliteit van het onderwijs.

Leerkrachten en scholen willen graag objectief vaststellen hoe de ontwikkeling van hun leerlingen verloopt. Deze ontwikkeling wordt onder andere vastgesteld met behulp van de resultaten uit (diagnostische) toetsen. Met behulp van dergelijke toetsen kan worden vastgesteld wat het niveau op bepaalde ontwikkelingsgebieden is van een leerling. Hierdoor kan een gerichte interventie plaatsvinden, wat leidt tot vergroting van de effectiviteit van het onderwijs.

Toetsen en toetsgegevens worden op basisscholen veelal verzameld in een leerlingvolgsysteem. Met een leerlingvolgsysteem krijgen scholen informatie over de voortgang in het leerproces van hun leerlingen. Een leerlingvolgsysteem kan op drie niveaus gebruikt worden. Ten eerste het niveau van *de leerling*. Leerkrachten kunnen alle leerlingen volgen in hun vorderingen en zo nodig de problemen analyseren. Dit is de diagnostische functie die hiervoor is genoemd. Ten tweede is er het niveau van *de groep*. De verzameling van gegevens op het niveau van de groep maakt zichtbaar hoe de groep ervoor staat. Ook op dit niveau is analyse belangrijk. Als namelijk blijkt dat bepaalde vaardigheden niet goed beheerst worden, dient de oorzaak achterhaald te worden. Dit kan leiden tot wijziging in het didactisch handelen van de leerkracht. Ten derde kunnen op het niveau van de school alle groepsoverzichten naast elkaar een beeld geven van de resultaten van de school. Op basis van deze gegevens kan de school systematisch werken aan de kwaliteit van het onderwijs. De school kan vervolgens de resultaten in verband brengen met kenmerken van het gegeven onderwijs, het schoolbeleid, de aard van de leerlingenpopulatie, enzovoort. Zijn de resultaten beneden verwachting, dan is er alle reden om het schoolbeleid bij te stellen.

In het tekstblok op pagina 18 is een voorbeeld opgenomen van het weergeven van gegevens uit een intern diagnostisch systeem. De diagnostische functie in het kader van het verbeteren of bijstellen van het primaire proces behoort tot de taak van de school. Dit interne systeem als onderdeel van een uitgebreider kwaliteitszorgsysteem van een school geeft echter niet voldoende informatie over de kwaliteit van het gegeven onderwijs door de school in vergelijking met andere scholen. Voor een adequaat intern systeem voor schoolverbetering en een verantwoordingsmodel richting ouders en overheid is een volgende stap nodig.

Zowel ten behoeve van schoolverbetering als voor de verbetering van transparantie van de informatie voor ouders is het noodzakelijk dat de gegevens vergelijkbaar zijn en dat de leerwinst kan worden vastgesteld. Bij dit laatste gaat het om het vaststellen van het deel van de vorderingen dat is toe te schrijven aan de inspanningen van de school (toegevoegde waarde). Eén van de aspecten waarover verantwoording afgelegd zou kunnen worden is over de toegevoegde waarde of netto leerwinst. Zoals hierboven aangegeven zeggen de huidige gegevens uit het interne systeem hier niets over. Het interne meet-systeem kan echter wel dienen als basis.

De invoering van een meetsysteem voor het vaststellen van de opbrengsten van het onderwijs is te beschouwen als een verdere verfijning van het verantwoording afleggen aan leerlingen en ouders. De huidige inzet van beleidsinstrumenten voor het verzamelen en zichtbaar maken van de opbrengsten laat een combinatie van intern en extern gebruik zien.

Zoals gezegd is de zorg voor de kwaliteit van het onderwijs een gezamenlijke verantwoordelijkheid van onderwijsinstellingen, burgers en overheid, waarbij iedere partij een rol speelt die samenhangt met zijn positie. Het zichtbaar maken van de toegevoegde waarde is hierin een gewenste ontwikkeling voor alle betrokken partijen. Het ondersteunt school en schoolbestuur bij het interne beleid van schoolverbetering. Het geeft de overheid een beter beeld van de opbrengst van het onderwijs. Het geeft ouders transparante informatie over de opbrengsten van het onderwijs. En ten slotte biedt het geven van een mogelijkheid tot vergelijking met andere (soortgelijke) scholen een ondersteuning bij het nemen van besluiten door de verschillende actoren.

De praktische samenhang tussen het gebruik van de gegevens voor verantwoording (extern) en schoolverbetering (intern) komt ook tot uiting in de vormgeving van het eerder genoemde proportioneel toezicht door de Inspectie van het Onderwijs. Proportioneel toezicht houdt in dat de Inspectie aansluit bij de verantwoordelijkheid van schoolbesturen. De externe verantwoording via de verticale verantwoordingslijn maakt dan gebruik van de gegevens uit het interne kwaliteitssysteem. Indien de overheid verfijning van het meetsysteem voorschrijft, zal in het toezichtkader van de Inspectie de *toegevoegde waarde* opgenomen moeten worden als één van de belangrijkste indicatoren van het domein *opbrengsten*. De verticale verantwoordingslijn dient immers wel te veranderen indien de horizontale verantwoordingslijn beter is uitgewerkt¹⁸. Zoals gezegd betekent dit dat in het toezichtkader (artikel 12 van de WOT) een verschuiving van aanbod naar opbrengst dient plaats te vinden. Hoofdstuk 3 gaat hier verder op in.

Door vergelijkbaarheid van de gegevens te organiseren kunnen ouders, schoolbestuur en overheid de informatie over de toegevoegde waarde gebruiken voor hun beslissingen. Ook de uitoefening van verantwoordelijkheden van lagere bestuursorganen zoals de gemeente en van (samenwerkende) schoolbesturen vraagt - voorzover zij een dergelijke informatiestroom al niet zelf hebben georganiseerd - om transparante informatie ter onderbouwing van het beleid. Ook hier kan de vaststelling van de toegevoegde waarde van het onderwijs in onderbouw en bovenbouw een rol spelen.

Overigens kan het meten van opbrengsten en het op basis daarvan vergelijken van scholen onbedoelde negatieve neveneffecten met zich mee brengen. Deze zijn mede afhankelijk van de beleidsinstrumenten die aan gegevens over de toegevoegde waarde worden gekoppeld, zoals bijvoorbeeld bekostiging. Op voorhand stelt de raad geen voorstander te zijn van een directe koppeling tussen gegevens over toegevoegde waarde en bekostiging. Indien deze koppeling wel wordt gelegd, treden veel ernstige, onbedoelde en veelal negatieve neveneffecten op. Hoofdstuk 3 gaat hier verder op in.

Conclusie

Uit het voorgaande is naar voren gekomen dat de bepaling van de toegevoegde waarde van het onderwijs een functie kan vervullen binnen een door de overheid geformuleerd

¹⁸ Zie voetnoten 12 en 17.

beleidskader. Door deregulering in het onderwijs is de autonomie van onderwijsinstellingen op enkele gebieden toegenomen; gezocht wordt naar een nieuw evenwicht tussen regulering, deregulering en verantwoording afleggen. Zoals gezegd is de zorg voor de kwaliteit van het onderwijs een gezamenlijke verantwoordelijkheid van onderwijsinstellingen, burgers en overheid, waarbij iedere partij een rol speelt die samenhangt met zijn positie. Het zichtbaar maken van de opbrengsten door het vaststellen van de toegevoegde waarde voor externe verantwoording en praktisch gekoppeld tevens voor interne schoolverbetering is daarbij een gewenste ontwikkeling.

1.4 Aanpak

Tegen de achtergrond van de hierboven geschetste context speelt de adviesvraag van de minister. Om de vragen van de minister te beantwoorden is de raad nagegaan welke methodologische aspecten van belang zijn bij enerzijds het toetsen van kinderen op jonge leeftijd en anderzijds het bepalen van de toegevoegde waarde van het onderwijs.

Hiervoor:

- is een literatuurstudie verricht;
- is in een expertmeeting aan wetenschappers gevraagd naar de actuele stand van zaken;
- is aan het SCO-Kohnstamm Instituut gevraagd om door middel van een secundaire analyse van de PRIMA-bestanden (testresultaten van leerlingen uit het basisonderwijs) te onderzoeken of er alternatieven bestaan voor de begintoets; het onderzoeksrapport is te raadplegen via de website van de raad (www.onderwijsraad.nl); en
- is gekeken naar internationale ervaringen met het vaststellen van de toegevoegde waarde van onderwijs.

Centraal staat de vraag of het mogelijk is de toegevoegde waarde van het onderwijs te bepalen en op welke wijze dat kan gebeuren. Op basis van bovenstaande wordt een voorlopige conclusie getrokken ten aanzien van de wijze van vaststellen van de toegevoegde waarde van het onderwijs.

Vervolgens:

- zijn de neveneffecten en de bezwaren ten aanzien van toetsen en ten aanzien van vaststellen van de toegevoegde waarde van het onderwijs door middel van een literatuurstudie in kaart gebracht;
- is gekeken naar internationale ervaringen; en
- is een panel van praktijkdeskundigen gevraagd naar hun mening over dit onderwerp.

De conclusies hieruit, samen met de conclusies ten aanzien van de methodologische aspecten, geven de ingrediënten voor de beantwoording van de adviesvragen.

1.5 Opzet

Hoofdstuk 2 gaat in op technische aspecten van meten van toegevoegde waarde. Aan de orde komen de conceptuele problemen en de domeinen die in een toets aan de orde dienen te komen. Verder worden mogelijke alternatieven verkend en wordt gekeken naar ervaringen in het buitenland.

Hoofdstuk 3 beschrijft de mogelijke bezwaren en vermoedens van neveneffecten. De wenselijkheid van het bepalen van toegevoegde waarde van het onderwijs is hiervan mede afhankelijk.

In hoofdstuk 4 worden de conclusies naar aanleiding van de voorgaande hoofdstukken beschreven en wordt de adviesvraag van de minister beantwoord.

2 Toegevoegde waarde

Dit hoofdstuk behandelt het begrip toegevoegde waarde tegen de achtergrond van onderwijsopbrengsten in het algemeen. Nagegaan wordt welke mogelijkheden en onmogelijkheden vanuit wetenschappelijk oogpunt vastzitten aan een systeem van vaststelling van toegevoegde waarde waarvan een begintoets onderdeel is. De raad komt tot de conclusie dat onder bepaalde voorwaarden het mogelijk is mede op basis van gegevens uit een begintoets de toegevoegde waarde van scholen te bepalen.

2.1 Opbrengst van onderwijs in het algemeen

Er zijn diverse omschrijvingen van de opbrengst van onderwijs mogelijk, afhankelijk van het accent dat op bepaalde functies van het onderwijs wordt gelegd¹⁹:

- Individueel rendement: het effect van een jaar onderwijs op het netto loon.
- Sociaal of maatschappelijk rendement van onderwijs. Dit begrip wordt in de literatuur op verschillende manieren gebuikt. Namelijk als:
 - het private rendement van onderwijs gecorrigeerd voor loon- en inkomstenbelasting, de directe kosten van onderwijs en onderwijssubsidies;
 - de productiviteitseffecten van onderwijs; en
 - de positieve externe effecten van onderwijs (uitgedrukt in individuele (financiële) baten van onderwijs en baten voor de samenleving als geheel belastingopbrengst, kosten volksgezondheid, enzovoort).
- Leerrendement: Het gaat hier om het verschil in competentie van een lerende bij intrede (instroomniveau) en bij het verlaten van de school (eindniveau).

De omschrijvingen laten zien dat de opbrengst van onderwijs op meerdere manieren aangegeven kan worden en daardoor ook op meerdere manieren kan worden vastgesteld. Metingen en toetsen spelen daarbij een belangrijke rol. Dat is zeker het geval als het gaat om het vaststellen van leerrendementen oftewel om de toegevoegde waarde van onderwijs.

Gebruik van toetsgegevens

De keuze voor (nadruk op) het soort opbrengst en de eenheid hangt onder meer af van het gebruik dat men van de verkregen informatie wil maken. Daarbij kan gedacht worden aan:

- terugkoppeling aan de leerling, ouders en de leraar (gebruik op individueel niveau);
- terugkoppeling aan de school (hetzij voor de schoolleiding, hetzij voor het bevoegd gezag); en

19

Zie verder bijvoorbeeld Groot & Maassen van den Brink (2003); Geerligts, Kops, Den Boer & Van der Veen (2002).

- het geven van een kwaliteitsoordeel over scholen door de overheid (door een absoluut oordeel of door vergelijkende oordelen tussen scholen)²⁰.

Elk van deze functies brengt haar eigen methode van bepaling met zich mee waarvoor andere kenmerken en eisen van belang zijn. Voor diagnostisch gebruik zijn bijvoorbeeld andere psychometrische eisen voor de gebruikte metingen relevant dan voor het maken van vergelijkingen tussen scholen.

Eenheid waarover toegevoegde waarde berekend wordt

De eenheid waarover de opbrengst bekeken wordt, kan betrekking hebben op:

- het individu;
- de leergroep;
- de school of onderwijsinstelling;
- de opleiding of leerweg; en
- het onderwijssysteem als geheel of delen daarvan.

De keuze voor een bepaalde eenheid waarover de opbrengst wordt gerapporteerd hangt van de functie af die bij de rapportage voorop staat. In het beleidsvoornemen voor de introductie van een begintoets staat het oordeel over de kwaliteit van het gegeven onderwijs als doel centraal. De overheid wenst preciezer aan te kunnen geven of en in welke mate een school aan bepaalde kwaliteitscriteria voldoet²¹. In dit advies worden de mogelijkheden daartoe onderzocht, waarbij de begintoets gezien wordt als onderdeel van een systeem om de toegevoegde waarde te meten. Toegevoegde waarde moet opgevat worden als dat deel van de vooruitgang in ontwikkeling van leerlingen dat aan de invloed van de school kan worden toegeschreven. Het bepalen van deze toegevoegde waarde maakt het geven van een kwaliteitsoordeel over scholen mogelijk. In dit advies gaat het daarbij om het leerrendement ofwel de inhoudelijk toegevoegde waarde door de school. De eenheid waarover de opbrengst of toegevoegde waarde handelt is derhalve de school. Voor het meten of de berekening zijn echter meetgegevens op het niveau van het individu (de leerling) nodig.

De bepaling van de toegevoegde waarde staat in het perspectief van de verantwoording. Het gaat er daarbij om of de toegevoegde waarde zoals vastgesteld, acceptabel is. In paragraaf 1.3.2 is hier al op ingegaan en in de slotconclusie komt de raad hierop terug.

2.2 Toegevoegde waarde als opbrengst van onderwijs²²

De achterliggende idee bij het vaststellen van toegevoegde waarde van het onderwijs in scholen is dat de leerling zich hoe dan ook ontwikkelt, maar dat de school deze ontwikkeling beïnvloedt: de school doet ertoe. Dat is ook hetgeen leraren voor hun werk motiveert: vooruitgang zien. Dit basisidee is simpel maar de precieze betekenis en de bepaling ervan is een complexe zaak. Dit wordt hieronder toegelicht.

²⁰ *Andere functies zijn: het op globaal landelijk niveau aangeven van de staat van het onderwijs, voor internationale vergelijkingen en voor het monitoren en het evalueren van min of meer afgebakende beleidsprogramma's.*

²¹ *TK 2002-2003, 28 600 VIII, nr. 116.*

²² *Deze paragraaf is onder meer gebaseerd op Saunders (1999), Sammons (2001) en Goldstein, Huiqi, Rath & Hill (2000), en op de bevindingen uit de expertmeeting die de Onderwijsraad op 20 februari 2003 ten behoeve van dit advies organiseerde (zie bijlage 2 voor de deelnemers aan deze bijeenkomst).*

Meten van toegevoegde waarde van het onderwijs in scholen

Het begrip toegevoegde waarde is ontleend aan economische theorieën. In essentie ligt aan het begrip het *input-output*model ten grondslag: *input* (grondstof, energie en dergelijke) wordt bewerkt en levert een product op, dat als *output* een verkoopwaarde heeft.

Het concept toegevoegde waarde in het onderwijs is verder uitgewerkt in onderzoek naar effectief onderwijs en vergelijkingen tussen scholen. Deze onderzoekstraditie laat zien dat de prestaties van leerlingen sterk worden beïnvloed door hun sociaal-economisch milieu van herkomst. Ook komt duidelijk naar voren dat het beginniveau (niveau bij aanvang van onderwijs) de meest belangrijke enkelvoudige voorspeller van latere prestaties is. Met andere woorden, om scholen goed te kunnen beoordelen op hun bijdrage aan de ontwikkeling van hun leerlingen en om scholen hierop te vergelijken, dan lijkt het nodig te zijn deze beide elementen (achtergrond en beginniveau) in de analyse te betrekken. Het volstaat niet om slechts ruwe, ongecorrigeerde gegevens over de prestaties van leerlingen te gebruiken en te publiceren: nodig is een analyse met gebruikmaking van achtergrond- en beginsituatievariabelen.

De onderzoeksbevindingen laten ook zien dat het bepalen van toegevoegde waarde en het verbinden van een oordeel daaraan een complexe zaak is. Er dient een scherp onderscheid te worden gemaakt tussen de totale vooruitgang die een lerende boekt (tussen twee meetmomenten in) en de vooruitgang die kan worden toegeschreven aan schoolactiviteiten (onderscheiden van andere beïnvloedende factoren). In een systeem om toegevoegde waarde te meten moeten daarom naast begin- en eindmeting idealiter ook andere beïnvloedende variabelen van buiten de school worden opgenomen (zie ook paragraaf 2.4 en Roeleveld, 2003).

Een volgende verfijning heeft te maken met het verschil in effect dat onderwijs kan hebben op verschillende groepen leerlingen, zoals bijvoorbeeld op hoogbegaafden versus 'gewone' talenten. Het onderwijs kan voor de ene groep zeer gunstig uitpakken, terwijl hetzelfde onderwijs voor de andere groep minder goed werkt (men spreekt ook wel van differentiële effecten). De toegevoegde waarde uitgedrukt in één schoolgemiddelde voor de hele groep kan dan tot een zekere vertekening leiden. Idealiter zou daarom de toegevoegde waarde van een school per subgroep moeten worden vastgesteld (Goldstein, Huiqi, Rath & Hill, 2000). Ook dit compliceert uiteraard de manier van opbrengstbepaling.

Het bepalen van de toegevoegde waarde is om verschillende redenen een complexe zaak:

- het is onmogelijk alle relevante beïnvloedende variabelen in de meting en analyse op te nemen;
- de invloed van factoren, bijvoorbeeld gezinsinvloeden, verandert in de loop van de tijd; en
- er worden meetfouten gemaakt.

Het ontwikkelingsproces bij kinderen is te complex om een volledige verklaring van ontwikkelingsvooruitgang in een meet- en analysemodel te vatten. Er ligt hier dus een serieus conceptueel en analyseprobleem. Dit is overigens een probleem dat niet specifiek is voor het meten van toegevoegde waarde; het overgrote deel van het onderwijs- en sociaal-wetenschappelijk onderzoek kampt met dit probleem. Er moet, zoals in deze gevallen gebruikelijk, naar een pragmatische oplossing worden gezocht.

Om zicht te krijgen op de betekenis van toegevoegde waarde van een school, is inzicht in effectieve factoren op school- en klasniveau en inzicht in externe factoren noodzakelijk. Ook is hier van belang de vraag om welke leerlingprestaties het gaat bij het bepalen van de toegevoegde waarde (zie paragraaf 2.3). Gestreefd moet worden naar een *optimaal* verklaringsniveau uitgaande van de vooronderstelling dat de school er wel degelijk toe doet. Kortom, de idee van toegevoegde waarde is in zichzelf simpel en begrijpelijk, maar de nadere bepaling ervan en de beoordeling van de resultaten zijn complex en vergen waardegebonden oordelen.

Als voorlopige bevinding stelt de raad dat er redelijk valide methodes bestaan om de toegevoegde waarde van het onderwijs te bepalen. Tevens stelt hij dat toegevoegde waarde niet het enige gegeven kan zijn wanneer een school op kwaliteit wordt beoordeeld. Voor gestelde metingen voor toegevoegde waarde leveren vooralsnog slechts benaderingen op. Zoals aangegeven is in principe het bepalen van toegevoegde waarde een grote verbetering ten opzichte van het alleen hanteren van ruwe scores op eindprestatiemetingen. Verfijning van het systeem door het rekening houden met mogelijke differentiële effecten zal waarschijnlijk tot een nog betere grondslag voor een kwaliteitsoordeel over scholen leiden.

2.3 Opbrengst van onderwijs: de keuze van ontwikkelingsdomeinen en eindtoetsen

Een belangrijke vraag voor de ontwikkeling van een meetsysteem voor toegevoegde waarde is welke leeropbrengsten in het systeem zouden moeten worden opgenomen. Er zijn immers verschillende ontwikkelingsgebieden te onderscheiden die in het basisonderwijs aan de orde zijn. Ze zijn grotendeels uitgewerkt in de kerndoelen. Deze bestrijken zes kennis- en vaardigheidsgebieden: taal, rekenen/wiskunde, zaakvakken, motorische, sociale en creatieve ontwikkeling.

Moet de opbrengst curriculumdekkend zijn?

Moeten alle ontwikkelingsgebieden die in het curriculum aan de orde komen, gemeten worden? Met andere woorden, moet breed gemeten worden of kan worden volstaan met een smalle meting voor slechts enkele leergebieden. Overwegingen die hierbij een rol spelen zijn: doet de meting recht aan de school, is de meting uitvoerbaar en is deze niet al te belastend voor de organisatie van de school.

Een belangrijk argument bij de keuze voor breed of smal meten is voorts de mate waarin consensus bestaat over het belang van leergebieden in het onderwijs. Voor de leergebieden rekenen/wiskunde en taal is die consensus groot, zowel wat betreft de relevantie van de gebieden als de accenten daarbinnen. Voor andere gebieden als oriëntatie op de natuur en oriëntatie op de wereld is de consensus al minder, en met betrekking tot kunstzinnige vorming en bewegingsonderwijs het geringst. Wordt de toegevoegde waarde gebaseerd op alleen rekenen en taal, dan zullen scholen die zich wensen te profileren op bijvoorbeeld kunstzinnige vorming of op bewegingsonderwijs hier niets van terugzien. Er zijn bepaalde scholen (bijvoorbeeld de zogenoemde vernieuwingscholen zoals Freinet-scholen) die vanuit hun pedagogische visie expliciet werken aan zaken als sociale competentie, zelfstandigheid, creativiteitsontwikkeling, enzovoort. Deze scholen zouden dergelijke domeinen terug willen zien in een meetsysteem. Willen overheid en ouders dit,

dan moet er breed worden gemeten. Dat heeft overzichtelijke praktische gevolgen. Meetproblemen kunnen hierbij toenemen (zie kader).

Consensus

Bij de verschillende kerndoelen en leergebieden horen toetsen waaraan kan worden afgemeten hoe ver leerlingen zijn gevorderd. Het probleem is nu dat om de toegevoegde waarde van het onderwijs met toetsen vast te stellen, een hoge mate van consensus nodig is over de inhoud van de leergebieden (zie ook Onderwijsraad, 2002a).

Hoe groter de consensus, hoe beter het mogelijk is goede toetsen te ontwerpen. Als scholen zeer verschillen op wat zij leerlingen bijbrengen met betrekking tot een bepaald leergebied, dan is het vrijwel onmogelijk een toets voor alle scholen te maken voor dat leergebied. Er is immers geen overeenstemming over de inhoud. Scholen kunnen in dat geval niet of maar zeer ten dele met elkaar worden vergeleken.

De raad constateert overigens dat de maatschappelijke discussie over de inhoud van het onderwijs zich steeds meer richting algemene en brede vaardigheden begeeft. Indien overheid en ouders willen dat de toegevoegde waarde van het onderwijs zichtbaar wordt gemaakt, dan dienen derhalve vele domeinen hierin te worden betrokken. Overigens realiseert de raad zich terdege dat de voorkeur voor 'breed meten' het nodige ontwikkelwerk met zich meebrengt.

Metten in het sociaal domein

Ten Dam et al. (2003) hebben het begrip sociale competentie nader geanalyseerd en een inventarisatie van mogelijk bruikbare instrumenten voor het meten daarvan verricht. De auteurs stellen dat zij met hun begripsverkenning en het ontwikkelde schema van sociale competentie de basis hebben gelegd voor een systematische meetbenadering. Zij concluderen over de huidige instrumentatie het volgende:

- Er is geen instrument beschikbaar dat zonder meer bruikbaar is voor het meten van sociale competentie in brede zin. Wel is er instrumentatie beschikbaar om onderdelen in kaart te brengen.
- Voorts is een aantal instrumenten vooral gericht op problematisch gedrag, hetgeen een te beperkte insteek biedt.
- Een aantal instrumenten heeft in feite meer betrekking op ontwikkelingstaken of sociale taken, dan op het meten van attitudes, (reflectieve) kennis of vaardigheden. Het gaat uiteraard om het laatste wil er sprake zijn van meten van 'sociaal competent kunnen handelen'.

De conclusie kan worden getrokken dat het ontwikkelingswerk op dit terrein nog niet af is, maar wel zover gevorderd dat verdere ontwikkeling van meetinstrumenten kans van slagen heeft.

Relatie begintoets - eindtoets

Voor het vaststellen van de toegevoegde waarde (toename in kennis en vaardigheden als gevolg van het doorlopen van de school) zijn een beginmeting en een eindmeting nodig (eventueel met tussenmetingen, zie paragraaf 2.4). Vanzelfsprekend moeten begin- en eindmeting met elkaar te maken hebben en voor beide metingen zal een (nieuw) toets-

instrumentarium ontwikkeld moeten worden. Het is echter niet zo dat aan het begin op precies dezelfde wijze kan worden gemeten als aan het eind. Een voorbeeld is het bepalen van taalniveau: aan het begin is dit te meten door middel van een woordenschattoets, terwijl aan het eind een tekstbegriptoets kan worden gebruikt. Beide toetsen meten taalniveau, maar met een andere inhoud. De relatie tussen begin- en eindmeting wordt vastgesteld door middel van bepaalde statistische technieken en vervolgens wordt op basis daarvan de toegevoegde waarde bepaald (Bosker, Béguin & Rekers-Mombarg, 2001)²³.

Metten is weten: ontwikkeling

Een metafoor:

Een jonge boom van een bepaalde boomsoort bestaat uit een stam met maar een paar takken. Aan de takken is een overzichtelijke hoeveelheid blaadjes te onderscheiden. Van deze jonge boom is gemakkelijk omtrek, aantal takken, aantal blaadjes, kleur, enzovoort vast te stellen. Dezelfde boom is na vijftien jaar uitgegroeid tot een ranke slanke boom met weinig zijtakken en veel blad. De vraag is of de beginsituatie (dikte van de stam, aantal zijtakken enzovoort) iets zegt over de hoeveelheid takken en bladeren van deze boomsoort op vijftienjarige leeftijd.

Bovenstaande metafoor kan toegepast worden op de relatie van de score van een beginmeting bij leerlingen in groep 1-2 en de score aan het einde van groep 8. We weten dat bij kinderen op jonge leeftijd de ontwikkeling slechts op een beperkt aantal gebieden is te meten. Zo weten we dat in diagnostisch opzicht het niet mogelijk is bij zeer jonge kinderen de algemene ontwikkeling te onderscheiden van taalontwikkeling (Onderwijsraad, 2001c).

De gebruikte boommetafoor maakt duidelijk dat de vraag welke ontwikkelingsgebieden van jonge kinderen relevant zijn om toegevoegde waarde te kunnen meten, niet eenvoudig is te beantwoorden. De geraadpleegde (buitenlandse) literatuur en de gehouden expertmeeting bieden echter voldoende aanknopingspunten voor een pragmatische oplossing.

Praktische problemen bij de afname van een begintoets

Toetsing moet in de praktijk uitvoerbaar zijn. Een belangrijke voorwaarde daarbij is het al bereikte ontwikkelingsniveau van kinderen. Toetsmethoden die bij jonge kinderen niet goed uitvoerbaar zijn, zijn bijvoorbeeld de traditionele gestandaardiseerde papier-en-pen-testen; onder andere vanwege de nog beperkte leesvaardigheid, korte spanningsboog en de neiging tot onvoorspelbare reacties in toetssituaties. Leeftijd geeft een globale indicatie van de mate van ontwikkeling, maar hierbij moet wel bedacht worden dat leeftijd een objectieve, administratieve aanduiding is, die slechts als een ruwe indicator van ontwikkelingen en ontwikkelingsniveaus mag worden opgevat. Gezien de grote variatie in de capaciteiten van kinderen op allerlei terreinen is het verstandig instrumentatie te gebruiken met een groot bereik qua ontwikkelingsniveau (Mooij, 2000). Zo moet een kind van zes jaar waarvan het ontwikkelingsniveau vergelijkbaar is met het ontwikkelingsniveau van het 'gemiddelde' vierjarig kind, de toets ook kunnen maken. Met andere woorden, bij het ontwerp van een meetinstrument dient een groot bereik in ontwikkelingsniveau centraal te staan, uitgaande van het niveau van het gemiddelde zesjarige kind.

23 Om wat preciezer te zijn: met behulp van gegevens over een groot aantal leerlingen, een populatie, kan het verband worden berekend tussen begintoetscores en eindtoetscores. Leerlingen die naar boven of beneden afwijken van de verwachte eindtoetscore, laten een positieve of negatieve toegevoegde waarde zien.

Een factor die toetsen op jonge leeftijd problematisch maakt is dat kinderen gauw zijn afgeleid, waardoor scores per dag kunnen verschillen (Commissie Indicatiestelling Onderwijsachterstanden, 1996). Naarmate kinderen ouder worden spelen deze problemen minder.

Tot slot moet een toetssysteem in schoolse situaties inzetbaar zijn en, afhankelijk van de functie, te hanteren zijn door het onderwijzend personeel dan wel door externe toetsafnemers. Omdat de klassieke grootschalige schriftelijke toetsen voor jonge kinderen niet bruikbaar zijn, zal individueel getoetst moeten worden, hetgeen veel tijd en deskundigheid vergt. Financiën en beschikbaarheid van deskundigheid spelen derhalve een grote rol.

De raad is van mening dat de school verantwoordelijk moet zijn voor de afname van toetsen. Richtlijnen voor afname worden neergelegd in een protocol, waardoor een valide afname zoveel als mogelijk wordt gewaarborgd. Om scholen tegemoet te komen voor de extra inspanningen, zal in formatieve zin compensatie geboden moeten worden.

Instrumenten

Op dit moment zijn er in Nederland diverse instrumenten beschikbaar voor de leeftijdsgroep van 4-7 jaar. Te noemen zijn onder meer de Woordenschattest voor groep 1 en 2, die deel uitmaakt van het DLE-leerlingvolgsysteem (DLE-LVS, DLE staat voor Didactische Leeftijdsequivalent), en de toetsen Ordenen, Taal voor kleuters en Ruimte en tijd, die eveneens deel uitmaken van een leerlingvolgsysteem (LVS van de Citogroep)²⁴. Door het Instituut voor Toegepaste Sociale Wetenschappen (ITS) wordt gewerkt aan de Nederlandse versie van een van oorsprong Engelse beginmetingstoets, het Onderbouwinformatiesysteem (OBIS).

Een overzicht van beschikbare toetsen voor het Nederlandse taalgebied biedt het Testboek voor het Onderwijs, uitgebracht door de Commissie Testaangelegenheden (COTAN) van het Nederlands Instituut van Psychologen (Evers, 2002). De COTAN beoordeelt toetsen, ook voor het onderwijs, op de volgende criteria:

- uitgangspunten bij de testconstructie;
- kwaliteit van het testmateriaal;
- kwaliteit van de handleiding;
- normen;
- betrouwbaarheid;
- begripsvaliditeit; en
- criteriumvaliditeit.

Alle criteria zijn van belang, maar met name de laatstgenoemde is essentieel als het gaat om het inzetten van een test in situaties waarin voorspelling aan de orde is. Over het algemeen geldt dat naarmate de beslissing op basis van toetsgegevens belangrijker is, het belang van genoemde criteria toeneemt.

Een apart probleem

Een aparte kwestie betreft het meten van sociaal-emotionele ontwikkeling, de ontwikkeling van het zelfvertrouwen, van zelfstandigheid, van weerbaarheid, enzovoort. Dit zijn persoonskwaliteiten die moeilijk te kwantificeren zijn en bovendien geldt hier niet: hoe

24

Zie ook Mooij (2000), die enkele testen gebruikte om oordelen van ouders en leerkrachten over beginkenmerken van vierjarigen te valideren.

meer hoe beter. Met andere woorden: naar welke niveaus van kwaliteiten zou gestreefd moeten worden? Er kan beredeneerd worden dat kinderen die problematisch gedrag vertonen, in de evaluatie meegenomen worden²⁵. In principe zijn dergelijke kinderen te signaleren met behulp van klinische instrumenten. Door middel van een leerlingvolgsysteem is een eventuele leerachterstand van kinderen na te gaan. Gegevens over dergelijke leerlingen kunnen meegenomen worden in het waardeoordeel over scholen. De gegevens over problematische leerlingen zijn daarbij dan aanvullend op de bevindingen omtrent de toegevoegde waarde²⁶.

2.4 Zijn er alternatieven voor een beginmeting?

Om na te gaan of er voor een begintoets alternatieven zijn, heeft de Onderwijsraad het SCO-Kohnstamm Instituut verzocht na te gaan of een beginmeting bij jonge kinderen vervangen kan worden door het meten van (benaderende) variabelen (zogenaamde *proxy's*, zoals het opleidingsniveau van de ouders)²⁷. De idee hierachter is dat het makkelijker is de laatstgenoemde variabelen vast te stellen dan jonge kinderen adequaat te toetsen. Uit onderzoek blijkt dat dergelijke variabelen verband houden met en van invloed zijn op (latere) schoolprestaties.

In de gevraagde analyse is nagegaan hoe deze twee manieren om de beginsituatie vast te stellen zich met elkaar verhouden (Roeleveld, 2003). In de analyse is gebruik gemaakt van gegevens uit het PRIMA-cohortonderzoek. Het PRIMA-cohortonderzoek is een tweejaarlijks onderzoek, waarbij steekproefsgewijs toetsen worden afgenomen bij leerlingen van groep 2, 4, 6 en 8. Bovendien worden vragenlijsten afgenomen bij leerkrachten van de betreffende groepen, bij directies en bij ouders van leerlingen uit groep 2. Deze gegevens maken het mogelijk de samenhang tussen de startpositie van leerlingen in groep 2 en hun latere onderwijsresultaten in de groepen 4, 6 en 8 te bepalen. Als representant voor een begintoets zijn de scores op een tweetal toetsen uit het Citoleerlingvolgsysteem (Begrippen en Ordenen) gebruikt. Als *proxy*-variabelen dienden het opleidingsniveau van de ouders, de etnische herkomst, de verblijfsduur in Nederland, het geslacht en de leeftijd in groep 2.

Bevindingen

Uit de analyses blijkt allereerst dat de begintoetscores betere voorspellers zijn voor latere onderwijsresultaten dan de achtergrondkenmerken²⁸. Tegelijk blijkt dat de voorspellingen op basis van toetsen en die op basis van achtergrondvariabelen maar voor een beperkt deel overeenkomen. Elke categorie verklaart voor een belangrijk deel een eigen,

25 *De redenering hier kan zijn dat, mits adequaat gemeten, het zinvol is scholen te vergelijken als pedagogische systemen in hun succes om met deze kinderen om te gaan.*

26 *In Amerikaans onderzoek over de kinderopvang zijn meet- en evaluatiesystemen ontwikkeld om kinderopvanginstellingen onderling te vergelijken op metingen van sociaal-emotionele ontwikkeling, zelfvertrouwen, zelfstandigheid, weerbaarheid, enzovoort (National Institute of Child Health and Human Development, 2002). Daarbij wordt gebruikt gemaakt van een longitudinale onderzoekopzet waarbij onafhankelijke variabelen gedurende langere tijd worden gemeten. Vervolgens wordt de invloed daarvan nagegaan door middel van een hiërarchische lineaire analyse, waarbij de invloed van achtergrondvariabelen als opleiding moeder, gezinsinkomen enzovoort mee in de analyse worden betrokken. Dit onderzoeksmodel biedt een andere invalshoek om de invloed van georganiseerde activiteiten op de ontwikkeling van kinderen vast te stellen. De vraag is echter of een dergelijke invalshoek bruikbaar is voor een onderwijskundig toegevoegdewaardesysteem, vooral omdat een dergelijke longitudinale opzet kostbaar is.*

27 *Het rapport van J. Roeleveld (2003) Herkomstkenmerken en begintoets. Secundaire analyse op het PRIMA-cohortonderzoek is te raadplegen via de website van de Onderwijsraad (www.onderwijsraad.nl). In deze paragraaf worden de belangrijkste bevindingen samengevat.*

28 *Onder achtergrondkenmerken worden hier begrepen: opleidingsniveau ouders, etnische herkomst, verblijfsduur, geslacht, leeftijd en intelligentie.*

uniek deel in de variantie van de verschillende onderwijsresultaten. Dat betekent dat voor een optimale voorspelling *beide* typen informatie (toetsresultaten en achtergrondvariabelen) nodig zijn. De verklaarde variantie varieert van 23% (bij Begrijpend lezen in groep 6) tot 34% (bij de Eindtoets Basisonderwijs in groep 8).

Dat toetsen en achtergronden tot deels verschillende voorspellingen kunnen leiden, kan ook aangetoond worden door per school te kijken naar de gemiddelde toegevoegde waarde voor leerlingen. De toegevoegde waarde bepaald op basis van toetsen en die bepaald op basis van de achtergrondvariabelen, komen weliswaar globaal overeen voor de meeste scholen, maar er zijn ook duidelijke uitzonderingen zichtbaar.

Naast beginbepaling en achtergrondkenmerken speelt ook een meting van algemene intelligentie een rol. Uit de analyses blijkt dat scores op een korte non-verbale IQ-test, afgenomen in groep 4, nog een zekere extra bijdrage aan de verklaarde variantie in de latere onderwijsresultaten kunnen leveren. Het lijkt daarom te overwegen dat een te ontwerpen (of te kiezen) begintoets ook dergelijke elementen meet.

Met betrekking tot de achtergrondvariabelen is gebleken dat het opleidingsniveau van de ouders hier het belangrijkste is en ook steeds effect blijft houden op de latere onderwijsresultaten, ook als gecontroleerd wordt voor scores op de aanvangstoetsen. De etnische herkomst van de leerlingen is in dit opzicht van veel minder gewicht. Wel van belang blijkt de leeftijd in groep 2. Deze weerspiegelt weliswaar slechts ten dele de leeftijd bij instroom, maar het lijkt wel aan te bevelen de leeftijd in de beginmeting op te nemen.

Ten slotte is ook het probleem van tussentijds vertrekkende en instromende leerlingen bij de bepaling van de toegevoegde waarde van scholen aan de orde gekomen. Hoe die toegevoegde waarde ook bepaald wordt, op basis van een begintoets of op basis van achtergrondkenmerken, altijd zullen er leerlingen zijn die uit hun 'normale' loopbaan vallen, door zittenblijven, verwijzing naar een speciale school voor basisonderwijs of door verhuizingen. De systematiek van bepaling van toegevoegde waarde dient dat op een of andere manier te verdisconteren. Uit dit onderzoek blijkt dat zulke mutaties meer voorkomen op scholen in de grote steden met veel allochtone leerlingen dan elders.

Conclusie

De raad trekt de conclusie dat voor het vaststellen van de toegevoegde waarde over een bepaalde onderwijsperiode van een bepaalde school, het absoluut noodzakelijk is voor een beginmeting toetsen te hanteren: alleen achtergrondvariabelen gebruiken volstaat niet. Aangetekend moet wel worden dat niet alle buitenschoolse invloeden op de ontwikkeling kunnen worden verklaard, ook al worden zowel de begintoetsscores als de achtergrondvariabelen in de analyse betrokken. Dit wijst erop dat niet alle factoren die de prestaties beïnvloeden, in de analyse betrokken zijn. Meer toetsen of toetsen die meer meten kunnen ertoe leiden dat de verschillen tussen scholen beter worden verklaard.

Een bevinding is voorts dat leeftijd als verklarende factor van belang is. Immers, de eerste periode van het basisonderwijs kenmerkt zich door ongelijke instroommomenten en daarom is het meenemen van de leeftijd in maanden in de analyse zeker aan te bevelen.

Ten slotte concludeert de raad dat in de begintoets ook elementen van algemene intelligentie moeten worden meegenomen, gelet op de onderzoeksbevindingen.

2.5 Ontwikkelingen in enkele landen

Ook in andere landen staat het onderwerp toegevoegde waarde en het gebruik van begintoetsen volop in de belangstelling. Een beknopte verkenning van ontwikkelingen in diverse Amerikaanse staten, Engeland en Australië laat zien dat gebruik en instrumentatie zeer gevarieerd zijn²⁹. In alle drie landen zijn beginmetingen en het implementeren van meetsystemen van toegevoegde waarde aan de orde. Dit is overigens niet overal het geval. In Vlaanderen bijvoorbeeld lukt het niet de onderwerpen begintoets en toegevoegde waarde op de beleidsagenda te krijgen.

Achtergronden

Achtergronden zijn in de drie landen verschillend. Vooral in sommige Amerikaanse staten wordt erg veel belang gehecht aan het via bepaling van toegevoegde waarde (*value added*) kunnen vaststellen van de prestatie van een school, en zelfs van individuele leraren. In de staat Tennessee heerst de opvatting dat uitsluitend meting van leerlingprestaties volstaat, achtergrondkenmerken van leerlingen worden niet meegenomen. Bij *value added* gaat het om een systeem van analyse van toetsgegevens om de vooruitgang van leerlingen vast te stellen, waarbij het huidige niveau wordt vergeleken met een eerder vastgesteld niveau. Het essentiële kenmerk hier is dat het gaat om verschilcores per individu. De analyse is gericht op de mate van vooruitgang. De claim van deze aanpak is dat de vooruitgang veel meer te maken heeft met de effectiviteit van de leraar dan met demografische factoren van de leerlingen. De vooronderstelling van de beschreven aanpak is dat de leraar de belangrijkste factor is die aan de ontwikkeling van de leerlingen bijdraagt.

In andere staten van Amerika en ook in Engeland en Australië ligt dit anders: daar wordt het wel van belang geacht om achtergrondvariabelen van leerlingen te meten en in de analyse mee te nemen, om aldus tot een betere bepaling van de toegevoegde waarde te kunnen komen.

In Engeland en Australië ligt vooral de nadruk op het gebruik van beginmetingen en bepaling van toegevoegde waarde voor diagnostische doeleinden, hoewel ook rapportage van de gegevens aan bovenschoolse autoriteiten als schoolbesturen en lokale onderwijsautoriteiten voorkomt. In sommige Amerikaanse staten staan beginmetingen en vaststelling van toegevoegde waarde duidelijk in het teken van een 'afrekencultuur'. Aan goede prestaties van de school en de leraar worden ook financiële beloningen gekoppeld. Het beoordelen van leraren op basis van geleverde toegevoegde waarde is overigens niet onomstreden en er wordt ook prudent mee omgegaan (zie verder onder *Gebruik*).

Inbedding in nationaal onderwijsbeleid

Belangrijk is de inbedding van de instrumentatie in nationaal beleid. Dit ligt in de drie landen als volgt.

In Australië is een grootscheeps programma in uitvoering dat betrekking heeft op beoordelingen en beoordelingswijzen in de eerste jaren van het onderwijs. Het programma *Assessment of Literacy and Numeracy in the Early Years of Schooling* is bedoeld om scholen, leraren en onderwijsautoriteiten te ondersteunen. Die ondersteuning is met name

29 Bronnen: *Assessment of Literacy and Numeracy in the Early Years of Schooling – an Overview; School Directors' Handbook. Value Added Assessment; Meisels, Atkins-Burnett, Xue, DiPrima Bickel & Seung-Hee (2003): Foundation stage profile handbook.*

gericht op de kernelementen van het *National Literacy and Numeracy Plan*. Het programma bestaat uit diverse projecten waarin beoordeling (assessment) centraal staat.

In Engeland wordt een duidelijke koppeling gelegd met nationaal vastgestelde leerdoelen, de *Early Learning Goals*. Deze verzameling doelen beslaan zes leer- of ontwikkelingsgebieden: (1) persoonlijke, sociale en emotionele ontwikkeling, (2) communicatie, taal en geletterdheid, (3) wiskundige ontwikkeling, (4) kennis en begrip van de wereld, (5) lichamelijke ontwikkeling, (6) creatieve ontwikkeling. De doelen zijn onderdeel van de zogenoemde *Foundation Stage*, een educatief programma voor de leeftijd van 3 en 4 jaar met een doorloop tot in het vijfde jaar (voorschoolse periode). Het vormt een verplicht onderdeel van het nationale curriculum van Engeland, naast de *keystages 1-4*. De *Early Learning Goals* vormen de doelen voor die educatieve periode. Op basis hiervan is het *Foundation Stage Profile* ontwikkeld. Het gebruik van deze *Foundation Stage Profile* als een meting voor het bepalen van toegevoegde waarde is problematisch vanwege gebreken aan de toetsen. Hierbij zij aangetekend dat de overheid deze functie van toegevoegde waardebeoordeling ook nog niet heeft gesuggereerd, maar er is wel de intentie om bij alle kinderen gegevens te verzamelen voor het ontwikkelen van nationale standaarden.

In Amerika wordt het kader voor beginmetingen en bepalingen van toegevoegde waarde gevormd door *The No Child Left Behind Act* (2001). Binnen dit kader hebben de diverse onderwijsautoriteiten een ruime beleidsruimte om aan de bedoelingen van deze wet gestalte te geven. Een kritisch kenmerk is daarbij de definitie van de jaarlijkse vooruitgang bij leerlingen en de manieren waarop scholen daaraan bijdragen. De federale wet is wat dat betreft weinig voorschrijvend.

Objectief en subjectief

Een belangrijk element in de ontwikkeling en het gebruik van beginmetingen betreft het onderscheid tussen objectieve toetsen en meer subjectieve oordelen op basis van observaties. In Australië bijvoorbeeld zijn de meeste instrumenten schoolgebonden en missen diensgevolge een meer objectief karakter. Wel wordt in een bepaalde regio het in Engeland ontwikkelde PIPS-instrumentarium gebruikt, het *Performance Indicators in Primary Schools Baseline Assessment*. Op basis van deze objectieve toetsen is berekening van toegevoegde waarde mogelijk³⁰.

Verplicht of vrijwillig

Voorts is van belang de vraag of scholen wel of niet verplicht zijn om beginmetingen af te nemen. In Amerika wordt een en ander geregeld door de onderwijsautoriteiten. Hetzelfde geldt voor Australië. In sommige staten zijn beginmetingen voor openbare scholen verplicht, in andere wordt het sterk aangemoedigd. De katholieke scholen en de onafhankelijke scholen in Australië hebben hierin hun eigen beleid, maar volgen veelal de staatsscholen. In Engeland bestaat sinds 1998 de verplichting om leerlingen bij binnenkomst in het primair onderwijs (op vijfjarige leeftijd) te toetsen. Hierbij dient een goedgekeurd beginmetingsinstrumentarium te worden gebruikt. Hiervan zijn er meer dan negentig in Engeland beschikbaar. Over de toegelaten meetinstrumenten hoeft echter tot op

30 In Aberdeen (Schotland) heeft men op basis van toetsing met dit instrumentarium op vijfjarige leeftijd en drie jaar later, kunnen constateren dat de kloof tussen achterstands- en overige leerlingen op het gebied van rekenen en taal sterk is verminderd. De autoriteiten schrijven dit toe aan de combinatie van "robust assessment from the early days in primary 1 and refined teaching approaches" (Website *The Times Educational Supplement*). In Nederland wordt het PIPS-instrumentarium momenteel beproefd op een aantal scholen in een onderzoekssetting (quasi-experiment). De eerste resultaten geven aan dat deze Nederlandse versie van de PIPS, de Obis-test, beschouwd mag worden als een betrouwbaar instrument om beginmetingen te verrichten.

heden nog geen verantwoording over betrouwbaarheid en validiteit te worden afgelegd, hetgeen een belemmering vormt voor gebruik buiten de directe context zoals voor *value added* systemen en vaststellen van leerstandaarden.

Gebruik

In alle landen is de meest geaccepteerde en gewaardeerde functie van beginmetingen en vaststelling van toegevoegde waarde die van diagnostisch middel om zwakke plekken bij leerlingen en onderwijs te ontdekken en verbeteringen aan te brengen door gerichte interventies in de onderwijsaanpak.

Ten aanzien van het gebruik van de verkregen gegevens voor het aanwijzen van 'sterke' en 'zwakke' scholen is er meer terughoudendheid, ofschoon vooral in Amerika de (lokale) onderwijsautoriteiten daar wel sterk toe neigen. Zoals gezegd wordt in Tennessee en Dallas zelfs per leraar de toegevoegde waarde nagegaan.

In Dallas bijvoorbeeld krijgen scholen die boven verwachting presteren, voor de gehele staf een financiële beloning. Opmerkelijk hier is ook dat de doelen die de effectiviteit van de school moeten representeren en worden opgenomen in het verantwoordingsstelsel, worden bepaald door een Taakgroep bestaande uit ouders, leraren, schoolleiders en vertegenwoordigers van de gemeenschap. In het gebruikte systeem wegen de toetscores het zwaarst mee. Aan de afname van de toetsen worden precieze voorwaarden gesteld.

Implementatie en beleidsinstrumenten

In bovengenoemde landen is er sprake van een nationale wet die beoordelen ook op jonge leeftijd sterk aanmoedigt. Nationale overheden ontwikkelen daarom implementatiestrategieën die toetsen en gebruik van toetsgegevens aanmoedigen. In Australië zijn daarvoor de volgende succesfactoren geïdentificeerd.

- Ontwikkeling vindt gedurende een lange periode plaats in samenwerking en samenspraak met leraren, scholen en deskundigen.
- De leraren worden geschoold in het gebruik van de beoordelingsinstrumenten en, in toenemende mate, de analyse van de resultaten.
- Er worden duidelijke verbanden gelegd tussen de beoordelingen, het onderwijsaanbod en de aan te wenden onderwijsstrategieën.
- Er worden speciale analyseprogramma's voor de gegevens ontwikkeld alsmede voor het beheer en communicatie erover. Er wordt daarbij veelal gebruik gemaakt van indicatoren die consistent zijn in het gebruik voor verschillende scholen en sectoren. Er wordt een gemeenschappelijke taal ontwikkeld om de prestaties van leerlingen te beschrijven. Hiermee wordt het ook mogelijk de ontwikkeling van de leerlingen in hun vaardigheden te beschrijven. Scholen en sectoren verkrijgen op basis van de aldus ingekaderde gegevensstroom inzicht in de ontwikkelingstrends bij leerlingen.
- Betrokkenheid van ouders wordt op twee manieren gerealiseerd: bij ouders wordt informatie gevraagd over de ervaringen met taal en rekenen voordat kinderen naar school gegaan zijn, en ouders worden geïnformeerd over de beoordelingen en hun relatie met het onderwijsaanbod.
- Er wordt sterk rekening gehouden met de verschillen tussen leerlingen. Dit leidt soms tot aparte instrumenten en onderwijsprogramma's voor bepaalde groepen leerlingen.

- Ten slotte is het beschikbaar stellen van voldoende materiaal, financiën en ondersteunend personeel van groot belang voor een adequate invoering en gebruik van de beoogde beoordelingsinstrumenten³¹.

Conclusie

De meeste beginmetingen vinden plaats op vijfjarige leeftijd. Nationale overheden vinden het belangrijk dat er gebruik gemaakt wordt van toetsen, maar geen van de in dit korte overzicht opgenomen landen schrijft precies voor welke toetsen te gebruiken en zelfs blijft verplichtstelling van beginmetingen in een aantal staten achterwege. Sterke aanmoediging vanuit het nationale overheidsniveau is er wel.

De meest geaccepteerde functie is de diagnostische. De andere functie, die van verantwoording naar een hogere autoriteit, levert problemen op. Een uitgewerkte implementatiestrategie waarin de betekenis van met name de beginmetingen in een systeem van toegevoegde waarde helder wordt gemaakt, is van groot belang. Ouders, leraren en scholen moeten vertrouwen ontwikkelen in toetsen en het gebruik dat van de gegevens gemaakt wordt. Dat vergt een meerjarige investering in ontwikkeling, gebruik maken van gegevens en kunnen waarnemen welke effecten mogelijk zijn.

2.6 Conclusie

In dit hoofdstuk is ingegaan op de methodologische aspecten van het vaststellen van de toegevoegde waarde van het onderwijs. De algemene conclusie luidt dat het in absolute zin niet mogelijk is om de toegevoegde waarde van het onderwijs vast te stellen. Daarvoor is het vaststellen van de toegevoegde waarde van het onderwijs in scholen voor primair onderwijs te complex. Wel is er een *next best* oplossing voorhanden waarmee een behoorlijke indicatie gegeven kan worden over de toegevoegde waarde van het onderwijs. Deze *next best* oplossing is wetenschappelijk gezien verantwoord en houdt een verbetering in van de thans ontstane praktijk om de schoolscores van de Eindtoets Basisonderwijs te gebruiken als vergelijkende indicator van schoolkwaliteit.

De conclusie ten aanzien van te meten domeinen luidt dat er een voorkeur is voor het meenemen van alle domeinen. De raad realiseert zich dat het gemakkelijker is om alleen te kijken naar de cognitieve domeinen zoals taal en rekenen/wiskunde. De raad sluit zich echter aan bij de gedachte dat onderwijs om meer gaat dan alleen taal en rekenen en dat derhalve ook een verbreding van het toetsinstrumentarium noodzakelijk is. Daarbij zou binnen de domeinen gezocht moeten worden naar vaardigheden die gemeenschappelijk zijn in de curricula van de scholen en die aantoonbaar door scholen te beïnvloeden zijn. Nogmaals, de raad realiseert zich dat er nogal wat ontwikkelwerk is te verzetten.

Voorts lijkt het verstandig een begintoets in te zetten aan het begin van groep 3, rond de zesjarige leeftijd³². De taxatie hierbij is dat omstreeks deze leeftijd, als ruwe aanduiding van ontwikkelingsniveau, er meer mogelijkheden voor de inzet van toetsinstrumenten zullen zijn die eenvoudig zijn af te nemen. Dit maakt het makkelijker dergelijke toetsen in te passen in een systeem voor het vaststellen van de toegevoegde waarde. Bovendien

31
32

Assessment of Literacy and Numeracy in the Early Years of Schooling – an Overview.
Dit voorstel sluit ook aan bij het voorstel van de raad voor het afnemen van taaltoetsen ten behoeve van een nieuwe NT2-regeling (Onderwijsraad, 2002c, p. 22).

is toetsing in een eerder stadium minder geschikt vanwege de ongelijke instroommomenten van leerlingen.

Praktisch biedt een meetmoment omstreeks het zesde jaar ook voordelen:

- toetsing vergt minder tijd;
- toetsen zijn klassikaal afneembaar; en
- de meting is betrouwbaarder dan op vierjarige leeftijd.

Buitenlandse ervaringen wijzen erop dat aan de implementatie van systemen van toegevoegde waarde de uiterste zorg moet worden besteed. De onderzoeksbevindingen uit met name Australië om een gedegen ontwikkelings- en implementatiestrategie te volgen, lijken voor de Nederlandse situatie zeer bruikbaar.

3 Onderwijs- en maatschappelijke aspecten

In het vorige hoofdstuk is aangegeven wat de mogelijkheden zijn voor de vaststelling van de toegevoegde waarde. In dit hoofdstuk wordt nader ingegaan op de mogelijke neveneffecten van invoering van een meetsysteem voor het vaststellen van de toegevoegde waarde van het onderwijs, en de daarmee samenhangende bezwaren. Is het maatschappelijk gezien aanvaardbaar om de toegevoegde waarde die scholen leveren, zichtbaar te maken? Het gaat dan met name om vermoedens over onbedoelde neveneffecten die kunnen ontstaan als gevolg van gebruik van de gegevens. Vervolgens wordt een voorstel gedaan om tegemoet te komen aan de bezwaren door de ongewenste neveneffecten te reduceren.

3.1 De neveneffecten en bezwaren

In deze paragraaf komen aan de orde vermoedens over neveneffecten van het vaststellen van de toegevoegde waarde en het gebruik van toetsen daarvoor en daarmee samenhangende bezwaren. Nadrukkelijk spreekt de raad hier over *vermoedens*. De geuite vermoedens over neveneffecten klinken vaak plausibel, maar zijn niet altijd empirisch bewezen. Soms zijn neveneffecten wel gedocumenteerd met een empirisch bewijs, maar laat nadere bestudering zien dat de mogelijke effecten plaatsvinden door het soort beleidsinstrument dat aan de gegevens is gekoppeld. Zo speelt bijvoorbeeld in de discussie vaak sterk mee dat toetsen en toetsgegevens in het kader van kwaliteitsbeleid en -controle al snel gebruikt kunnen worden voor het maken van vergelijkingen tussen onderwijsinstellingen (zie de 'benchmarking-lijsten' in dagbladen als De Volkskrant en Trouw en in Elsevier). Een ander neveneffect kan ontstaan indien de nadruk alleen ligt op output-meting. Leraren kunnen zich dan gaan gedragen in overeenstemming met de prikkels die van het meetsysteem uitgaan. In dat geval prevaleert het belang van de uitkomsten boven de aandacht voor innovatie en dergelijke (De Bruijn, 2002b). Deze en andere vermoedens van onbedoelde neveneffecten van sturing op resultaten die zijn gevonden in verschillende bronnen worden hieronder beschreven³³.

Beïnvloeding van het imago van de school en de keuze van ouders:

- scholen met een kansarme populatie komen er bij vergelijkingen doorgaans slecht af;
- vergelijken van scholen leidt tot 'shoppende ouders' en tot sterke populatieschommelingen; en
- het vergelijken van scholen bevordert segregatie ('witte vlucht').

Beïnvloeding van leraren en leerlingen:

- leraren werken het liefst op scholen die hoog scoren en ontvluchten andere; en
- het vergelijken en het publiceren van 'ranking' werken sterk motivatieverlagend.

33 Onder meer ontleend aan Boes & Both (2002), De Bruijn (2002b), Onderwijsraad (1999b), Karsten, Visscher & De Jong (1999).

Beïnvloeding van het schoolbeleid:

- scholen zouden goede leerlingen kunnen selecteren bij binnenkomst; voorzover dat niet mogelijk is, zullen scholen een strategie voeren om de slecht presterende leerlingen voortijdig of onnodig te verwijzen naar bijvoorbeeld het speciaal basisonderwijs;
- een sterkere nadruk op de groep leerlingen die zich bevindt op de grens tussen twee prestatieniveaus³⁴;
- vergelijken vergroot de kans op fraude (te verstaan als toetsafnames die onder andere condities dan strikt voorschreven plaatsvinden om de score in gunstige zin te beïnvloeden); scholen die zich daaraan schuldig maken drukken andere weg, want het systeem werkt zoals ‘communicerende vaten’;
- vergelijking leidt gemakkelijk tot curriculumversmalling, overaccentuering van wat toetsbaar is en instructieve vervroeging;
- er treedt een sterk aandacht voor de examens op;
- outputmeting blokkeert innovaties; er zal minder snel gekozen worden om onderwijskundige vernieuwingen in te voeren, omdat invoeringstrajecten op de korte termijn (als gevolg van een invoeringstraject) ten koste kunnen gaan van de toegevoegde waarde van het onderwijs; en
- afname van de samenwerkingsbereidheid tussen scholen.

De raad merkt hierover het volgende op:

Ten aanzien van de beïnvloeding van het imago van de school en van de keuze van ouders zal bij de presentatie van de gegevens op schoolniveau rekening gehouden moeten worden met achtergronden van de leerlingen. Bij de presentatie van de gegevens op schoolniveau zal rekening gehouden moeten worden met achtergronden van de leerlingen. Ook nu al gebeurt dat door bij schoolvergelijkingen uit te gaan van *families of schools* (zie ook pagina 9). Met andere woorden: veel is afhankelijk van de methodiek die wordt gebruikt om een vergelijking te maken.

Ten aanzien van de *shoppende ouders* wordt opgemerkt dat Dronkers in 1999 heeft laten zien dat het effect van de publicatie van kwaliteitsgegevens en van publicatie van de sociale compositie van scholen slechts een klein effect heeft op de leerlingenaantallen van scholen. Van alle waargenomen verschillen in het groeipercentage van scholen kan slechts 2% van de variantie verklaard worden door de publicatie van kwaliteitsgegevens in Trouw (Dronkers, 1999; Dronkers & Veenstra, 2001).

Ten aanzien van de beïnvloeding van leraren en leerlingen geldt dat in tijden van arbeidsmarktkrapte de aantrekkelijkste werkgevers de minste problemen zullen hebben. Dat goed scorende scholen voor werknemers een bepaalde aantrekkingskracht zullen hebben, ligt voor de hand. De raad is overigens van mening dat werkzoekende leraren zich ook door andere aspecten laten leiden als het gaat om keuze voor een arbeidsplaats (bijvoorbeeld faciliteiten, scholingsmogelijkheden, doorstroommogelijkheden, salariëring (bonus voor leraren op achterstandsscholen), professioneel management, enzovoort).

Ten aanzien van de motivatieverlaging bij leraren geldt dat dit gevaar zeker speelt indien de meting en bepaling van de toegevoegde waarde van het onderwijs op een armoedige

34 *De school kan deze laatstgenoemde groep met relatief geringe inspanningen naar een hoger niveau brengen en daarmee tot betere resultaten komen.*

manier tot stand komen, zonder rekening te houden met de complexiteit van het onderwijs. Zoals gezegd kunnen leraren zich gaan gedragen in overeenstemming met de prikkels die van het meetsysteem uitgaan. Uit andere beroepssectoren zijn hiervan illustratieve voorbeelden bekend³⁵. De raad adviseert om een evenwichtig meetsysteem voor toegevoegde waarde te doen ontwikkelen. Het eerder gedane voorstel om alle leergebieden uit de kerndoelen in het systeem te betrekken, draagt daaraan bij. De school zal dan immers op ten minste zes domeinen een score van toegevoegde waarde kunnen rapporteren. Dit draagt bij aan een veelzijdig beeld. Daarnaast kunnen ook andere gegevens worden gebruikt als kwaliteitsindicator. Een en ander draagt ook bij aan vervolmaking van een intern kwaliteitszorgsysteem.

Bovendien dienen scholen de informatie over de toegevoegde waarde te gebruiken in de schoolgids als ondersteuning voor hun verantwoording van het schoolbeleid aan ouders. Overigens realiseert de raad zich dat de verschillende typen informatie die tegelijkertijd en naast elkaar worden gepubliceerd, met elkaar concurreren om aandacht.

Ten aanzien van de beïnvloeding van het schoolbeleid zijn voorbeelden uit Amerika bekend, waarbij scholen vanwege de koppeling van prestaties en financiering hun leerlingen meer gaan selecteren (Zollers & Ramanathan, 1998). Dit is een van de redenen waarom de raad directe koppeling tussen alleen de leerprestaties en bekostiging van het onderwijs afwijst. Verder laat het meetinstrumentarium te veel fouten toe en door tactisch gedrag van scholen kunnen ongewenste neveneffecten ontstaan. Of tactisch gedrag (te verstaan als het afnemen van toetsen onder andere normen dan strikt voorgeschreven) voorkomt in Nederland, is niet bewezen³⁶. Visser (2003) geeft aan dat in Amsterdam de regels voor afname van de Eindtoets Basisonderwijs in het schooljaar 1996-1997 door scholen mogelijk ruim zijn geïnterpreteerd om een betere schoolscore te realiseren. De voorschriften voor afname vanwege de Citogroep liet deze ruimte toe³⁷. Voor wat betreft de scholen die hun leerlingen voorbereiden op de Eindtoets Basisonderwijs met toetstraining, wordt opgemerkt dat dit niet of nauwelijks leidt tot hogere scores op de Cito-eindtoets³⁸.

Ten aanzien van de afname van de samenwerkingsbereidheid tussen scholen wordt opgemerkt dat Fiske en Ladd (2000)³⁹ hebben aangetoond dat prestatiemeting een negatieve invloed heeft op de relaties die scholen met elkaar onderhouden.

35 Een illustratief voorbeeld, ontleend aan De Bruijn (2002b), is afkomstig van de Amerikaanse FBI. De FBI heeft als prestatie-indicator het aantal arrestaties. Na verloop van tijd blijkt dat de FBI tot steeds hogere productiecijfers komt door deserteurs uit het leger te arresteren. Deze zijn eenvoudiger op te sporen dan andere overtreeders van de wet. Deze arrestaties dienen nauwelijks een maatschappelijk belang en worden slechts verricht om aan de prestatienormen te voldoen. Het gevolg is dat het percentage arrestanten dat vervolgens wordt vervolgd, uitermate laag is omdat het hier om triviale prestaties gaat.

36 Omdat de schoolscore op de Eindtoets Basisonderwijs soms impliciet en soms expliciet als externe kwaliteitsindicator wordt gebruikt, is er vrees voor tactisch gedrag en fraude. Met betrekking tot dit laatste punt wordt opgemerkt dat uit de gegevens van de Citogroep blijkt dat er geen aanleiding is om te veronderstellen dat er sprake is van grootschalige fraude met de Cito-toets. De Citogroep gaat er dan ook van uit dat het basisonderwijs de Cito-toets professioneel gebruikt; scholen nemen de toets af omdat ze over onafhankelijke empirische gegevens willen beschikken (TK, 2002-2003, 1009, Antwoord op vragen gesteld door De Vries (CDA)).

37 Visser (2003) geeft aan dat in de handleiding van het Cito in 1997 een zinsnede staat die ruimte biedt voor een eigen interpretatie van de deelname van allochtone leerlingen aan de Eindtoets. Het vermoeden bestaat dat in 1997 door verschillende Amsterdamse basisscholen gebruik is gemaakt van de ruime interpretatiemogelijkheid. De schoolscore zou, door sommige leerlingen niet te laten deelnemen, gunstig beïnvloed kunnen worden. De ruimte voor eigen interpretatie is in 1999 ingeperkt (zie verder Visser, 2003, p. 41).

38 Aldus de Citogroep. Zie ook website citogroep.

39 Fiske & Ladd (2000).

Neveneffecten en bezwaren zijn door de raad in 1999 in kaart gebracht. De raad is in zijn advies Schoolkwaliteit in beeld ingegaan op de vraag hoe op een verantwoorde wijze kan worden voorzien in de behoeften aan gegevens over de kwaliteit van scholen, en heeft daarbij ook gelet op de rol van de rijksoverheid. Om deze vraag te beantwoorden is destijds een studie door Karsten, Visscher & de Jong (1999) uitgevoerd naar buitenlandse ervaringen met het openbaar maken van resultaatgegevens. De conclusie is dat indien de nadruk alleen wordt gelegd op toegevoegde waarde, naast positieve effecten (transparantie, prikkels voor kwaliteit, mogelijkheid voor sturing op prestaties, mogelijkheid tot vergelijking met soortgelijke basisscholen), ook ongewenste effecten kunnen optreden.

De ontwikkelingen in de maatschappij die hebben geleid tot een sterker kwaliteitsbewustzijn en een toegenomen behoefte aan informatie over de kwaliteit van voorzieningen, zijn echter positief en onomkeerbaar (Onderwijsraad, 1999b). Dat steeds meer scholen kwaliteitsbewust zijn, blijkt onder meer uit het gegeven dat in 2002 nog maar circa 30% van de scholen tegenover 48% in 1999 niet beschikt over (volgens de Inspectie van het Onderwijs) bruikbare opbrengstgegevens (Inspectie van het Onderwijs, 2003)⁴⁰. Ongewenste effecten zoals een toename van selectieverschijnselen, een sterkere gerichtheid op examens en op de groep leerlingen die zich bevindt op de grens tussen twee prestatieniveaus, demotivatie van leraren en een afname van de samenwerkingsbereidheid tussen scholen acht de raad echter onaanvaardbaar. In de volgende paragraaf wordt ingegaan op de maatregelen die de overheid dient te nemen om ongewenste neveneffecten tegen te gaan.

3.2 Reductie van ongewenste neveneffecten

Zoals eerder aangegeven wordt er in het onderwijs veel getoetst en hebben de scholen op grond van artikelen 12, 13, 14 en 16 van de WPO een kwaliteitsbeleid geformuleerd⁴¹. Beleidsbeslissingen die scholen en schoolbesturen nemen zijn veelal gebaseerd op informatie uit het eigen kwaliteitssysteem, waaronder een analyse van en oordeel over het onderwijskundig beleid. Een deel van de benodigde gegevens wordt verkregen via methodeonafhankelijke toetsen. Methodeonafhankelijke toetsen maken een vergelijking met andere scholen mogelijk.

Voorgesteld wordt om bij het ontwerp van een systeem voor het meten van de toegevoegde waarde aan te sluiten bij datgene wat nu al in scholen gebeurt. Derhalve adviseert de raad, ten eerste, om aan te sluiten bij de bepalingen in de WPO over het kwaliteitsbeleid en via een AMvB als bedoeld in artikel 13 van de WPO, eerste lid onder a, de horizontale verantwoordingslijn te verstevigen. Dit kan door richtlijnen op te stellen over vaststelling en publicatie van de toegevoegde waarde en deze op te nemen in genoemde AMvB. Ten tweede adviseert de raad om op een breed terrein toetsmateriaal (te laten ontwikkelen en) beschikbaar te stellen, opdat scholen dit kunnen gebruiken om aan de verplichtingen voortkomend uit de WPO en de AMvB te voldoen.

De raad stelt zich in de toekomst een model voor waarbij er sprake is van een databank die gevuld wordt met een brede variatie aan begin- en eindtoetsen. Voor elk domein en voor elke leeftijdsgroep komen zo verschillende toetsitems beschikbaar. Op deze manier

⁴⁰ Zie ook voetnoten 14 en 15.

⁴¹ Genoemde artikelen uit de WPO omvatten bepalingen omtrent respectievelijk het schoolplan, de schoolgids, de klachtenregeling en de vaststelling van deze documenten.

wordt de mogelijkheid voor variatie ingebouwd waardoor een school toetsen kan kiezen die bij de eigen onderwijskundige opvatting passen.

Voorts dienen in de (begin)meting beïnvloedende factoren te worden verwerkt, waardoor er sprake is van netto leerrendement en een vergelijking met elke school mogelijk is. Scholen kunnen de gegevens gebruiken voor het eigen schoolbeleid en zij zijn zelf verantwoordelijk voor publicatie van de gegevens aan ouders, schoolbestuur en overheid. De publicatie wordt vergezeld van gegevens over schoolspecifieke proceskenmerken, zodat een integraal beeld van de school naar voren komt.

Verstevigen horizontale verantwoordingslijn

Door het zichtbaar maken van de toegevoegde waarde van het onderwijs krijgen de verschillende actoren (schoolbestuur, ouders en overheid) belangrijke informatie waarop zij hun beslissingen kunnen baseren. Ten aanzien van het toezicht op het onderwijs en het verantwoorden van de prestaties van scholen aan de overheid merkt de raad op dat het gevaar bestaat dat de toezichthoudende rol van de overheid doorslaat. Als de overheid ervoor kiest dat scholen meer beleidsvrijheid krijgen en scholen zich moeten verantwoorden door de opbrengst van het onderwijs zichtbaar te maken, dan zal de overheid een keuze moeten maken in de wijze waarop zij de verantwoording van en het toezicht op scholen organiseert. De raad is van mening dat in het huidige toezichtkader (verticale verantwoordingslijn; zie ook voetnoot 12) de verschuiving in het beleid (van aanbod naar opbrengsten) nog onvoldoende zichtbaar is gemaakt. Bepaling van de toegevoegde waarde en opnemings hiervan in het verantwoordingssysteem betekent dat in het toezichtkader van de Inspectie van het Onderwijs de beleidskeuze van aanbod- naar opbrengstcontrole meer moet doorklinken. Indien de overheid te veel wil weten en kiest voor één opbrengstverantwoording én aanbodverantwoording, dijt het meetsysteem tot onhanteerbare proporties uit⁴². Concreet stelt de raad voor dat in het toezichtkader het accent meer komt te liggen op de opbrengstenkant en minder op de aanbodkant.

Voor de verantwoording over de opbrengsten van het onderwijs stelt de raad voor dat bij publicatie van gegevens gedifferentieerd wordt naar actoren. De overheid heeft immers andere informatie nodig om haar beslissingen op te baseren dan een schoolbestuur of ouder. Met alleen publicatie van kwantificeerbare gegevens over de toegevoegde waarde van het onderwijs schiet een school voor het schoolbestuur of voor (potentiële) ouders haar doel voorbij. Voor ouders zijn niet alleen resultaatgegevens van belang, maar bijvoorbeeld ook informatie over de wijze waarop het onderwijs is ingericht, het aanbod. Het gaat in dergelijke publicaties om een evenwichtige combinatie van proces- en productprestaties. Voorgesteld wordt derhalve om in de genoemde AMvB de volgende voorschriften op te nemen:

- publicatie van meer gegevens dan alleen over de toegevoegde waarde van scholen, bijvoorbeeld doelmatigheidsgegevens (percentage leerlingen dat een of meer jaren vertraging oploopt en het percentage voortijdige uitvallers);
- publicatie van gegevens over andere kwaliteitskenmerken; en
- publicatie van aanbodgegevens en procesprestaties.

De toegevoegde waarde van het onderwijs in de school en de uitleg van de school over het aanbod zullen op elkaar betrokken moeten worden, met name als het gaat om ver-

42

Het gevaar van uitdijende meetsystemen is door De Bruijn (2002b) aangegeven als een van de wetmatigheden die bij prestatie-meting vaak optreden. In dit geval gaat het om de wet van beleidsaccumulatie: mushrooming. Er ontstaan voortdurend nieuwe productdefinities en dus moet nieuwe informatie worden aangeleverd.

antwoording en informatie naar de directe omgeving van de school. De raad acht de schoolgids hiervoor een geschikt instrument.

De raad geeft als overweging mee om bij het ontwerpproces ouders, leraren, schoolleiders en vertegenwoordigers van de verschillende besturen- en managementorganisaties te betrekken. Levensbeschouwelijke en andere door de gemeenschap belangrijke geachte aspecten kunnen zo worden betrokken bij de toetsontwikkeling; wel moet er tegelijkertijd voor worden gewaakt dat er geen afbreuk wordt gedaan aan de vergelijkbaarheid. Door de gebruikers te betrekken wordt de horizontale verantwoordingslijn versterkt (zie ook Onderwijsraad, 2002c). De raad adviseert met nadruk om voordat tot grootschalige invoering wordt overgegaan, één en ander in een pilot uit te proberen en te vervolmaken.

Niet alleen afspraken over de afname van toetsen, ook afspraken over het gebruik van de gegevens en over wie ze mag gebruiken, zijn van belang. Tevens kan de Inspectie de indicatoren van de opbrengsten gebruiken zijn voor een macrorapportage over de stand van zaken van het onderwijs⁴³.

In hoofdstuk 2 is aangegeven dat in de verschillende landen terughoudend wordt omgegaan met het gebruik van de verkregen gegevens voor het aanwijzen van 'sterke' en 'zwakke' scholen. Alleen in Amerika neigen de (lokale) onderwijsautoriteiten daar soms wel sterk naar. Zoals aangegeven wordt in Tennessee en Dallas zelfs per leraar de toegevoegde waarde nagegaan en worden financiële prikkels gekoppeld aan het niveau van de toegevoegde waarde van het onderwijs in de school. De raad acht een dergelijk gebruik door de overheid, gezien de te verwachten neveneffecten, vooralsnog onwenselijk. In de toekomst kan de raad zich voorstellen dat hij een apart advies uitbrengt over de verschillende componenten in bekostigingsmodellen.

3.3 Conclusies

Uit het voorgaande is gebleken dat het zichtbaar maken van de toegevoegde waarde van het onderwijs onder bepaalde randvoorwaarden mogelijk is. Zichtbaar maken van de toegevoegde waarde van het onderwijs is voor de betrokken actoren van belang en draagt bij aan een betere verantwoording door het onderwijs.

De toename in de behoefte aan informatie over de kwaliteit van voorzieningen acht de raad positief en onomkeerbaar. Hij geeft hierbij aan dat er onbedoelde en ongewenste neveneffecten kunnen optreden. Er zijn maatregelen denkbaar die ongewenste neveneffecten reduceren. De voorgestelde maatregelen hebben onder meer betrekking op de afspraken die worden gemaakt tussen overheid, scholen, ouders en schoolbesturen over welke gegevens op welke manier worden verzameld en voor wie.

In het volgende hoofdstuk zal in algemene zin hierop in worden gegaan. Tevens worden conclusies getrokken op basis van de hoofdstukken 1 tot en met 3 en worden de adviesvragen beantwoord.

⁴³ Voor de volledigheid zij opgemerkt dat binnen het kader van horizontale verantwoording het openbare inspectieoordeel eveneens een rol speelt. In de *Memorie van Toelichting op de WOT* wordt gesteld dat het inspectietoezicht een stimulerende werking dient te hebben op het zelfcorrigerend vermogen van instellingen en het beoogde dynamische evenwicht met actoren in de omgeving (p.9). Verder wordt opgemerkt dat het van groot belang is dat de inspectie haar oordelen actief openbaar maakt. Zie verder onder 1.3.1, subtitel *Transparantie voor ouders en besturen*.

4 Conclusies, aanbevelingen en beantwoording van de adviesvraag

In paragraaf 4.1 van dit hoofdstuk worden op basis van de voorgaande hoofdstukken conclusies geformuleerd en aanbevelingen gedaan. In paragraaf 4.2 worden vervolgens de adviesvragen van de minister beantwoord.

4.1 Terugblik, conclusies en aanbevelingen

In voorgaande hoofdstukken is aangegeven dat de vergroting van de autonomie samen gaat met meer aandacht voor verantwoording en dat dit vraagt om het zichtbaar maken van de opbrengsten van het onderwijs. De raad concludeert dat de huidige praktijk van het zichtbaar maken van de opbrengsten niet adequaat is en derhalve om verbetering vraagt. Nu wordt vaak als een van de indicatoren voor kwaliteit de schoolscore op de Eindtoets Basisonderwijs gebruikt. Het enkelvoudige gebruik van dit outputgegeven als indicator van kwaliteit ziet de raad als te beperkt. Een meetsysteem met ten minste een begin- en een eindmeting maakt het beter mogelijk om opbrengsten van het onderwijs vast te stellen. Deze metingen zouden om praktische redenen tevens deel moeten uitmaken van het kwaliteitssysteem van de school. Om de toegevoegde waarde goed te kunnen vaststellen dienen zowel voor een begin- als voor een eindmeting toetsen te worden ontwikkeld.

De raad pleit ten aanzien van de toetsdomeinen voor een brede benadering. Dit wil zeggen dat het toetsinstrumentarium alle door de kerndoelen bestreken leergebieden dient af te dekken. De raad realiseert zich dat de uiteindelijke keuze een afweging is van kosten en baten. Er kan gekozen worden voor een breed of een smal palet aan toetsen. In het laatste geval ligt het voor de hand en is het ook wenselijk dat ten minste op de leergebieden Nederlands en rekenen/wiskunde getoetst wordt. Binnen de twee uitersten zijn verschillende varianten mogelijk. De raad is er voorstander van om alle leergebieden te meten, maar hij realiseert zich dat daarvoor nog heel wat ontwikkelwerk nodig is. De raad stelt dan ook voor om een ontwikkeltraject uit te zetten van vier à zeven jaar, waarin toetsen worden ontwikkeld in samenspraak met de betrokkenen. Een gedegen communicatie- en implementatieplan dient hieraan ten grondslag te liggen. Een pilot waarin één en ander wordt uitgetoetst en neveneffecten in kaart worden gebracht dient deel uit te maken van dit ontwikkeltraject. Vanwege de te verwachte complexiteit adviseert de raad te starten met ontwikkeling en gebruik van toetsen op de twee eerder genoemde leergebieden, gevolgd door de overige leergebieden.

Verder adviseert de raad om bij de begintoeetsen ook facetten van intelligentie op te nemen. Dit verhoogt in zekere mate de voorspellende waarde van begintoeetsen en derhalve de betrouwbaarheid van de metingen en vaststelling van de toegevoegde waarde. Ten aanzien van de afname van toetsen bij jonge kinderen adviseert de raad om uit te gaan van een groot bereik qua ontwikkelingsniveau. Hiermee wordt bedoeld dat een kind

van zes jaar met een ontwikkelingsniveau vergelijkbaar met dat van het gemiddelde vierjarige kind, de toets ook moet kunnen maken. Bepaling van het beginniveau dient plaats te vinden rondom het zesde levensjaar.

Uit het onderzoek dat, op verzoek van de raad, is uitgevoerd door het SCO-Kohnstamm Instituut blijkt dat de toetsuitslagen door de jaren heen stabiel blijven. Op grond van deze constatering zijn er geen methodologische argumenten die een tussenmeting rechtvaardigen. Er zijn echter wel praktische argumenten. Deze hebben te maken met de grote wisseling in leerlingenpopulatie die tussen groep 1 en 8 plaatsvindt, als gevolg van verhuizing, doubleren, overstap naar een andere school, enzovoort. Een bijkomend voordeel van tussentijdse metingen is dat het systeem voor het vaststellen van de toegevoegde waarde (externe verantwoording) en dat voor de interne kwaliteitszorg elkaar positief kunnen beïnvloeden.

Voor vaststelling van de toegevoegde waarde dienen achtergrondvariabelen van de leerlingen te worden meegenomen. De combinatie van achtergrondvariabelen en de uitkomst van een statistische bewerking van de begin- en de eindmeting geven een indicatie over de toegevoegde waarde van het onderwijs⁴⁴.

Voor de uitvoering adviseert de raad de minister nauw aan te sluiten bij de bepalingen in de WPO over het kwaliteitsbeleid en via een AMvB als bedoeld in artikel 13 van de WPO, eerste lid onder a, de horizontale verantwoordingslijn te verstevigen. Deze AMvB kan voorschriften bevatten om gegevens over de toegevoegde waarde van het onderwijs, samen met gegevens over de inrichting en het aanbod van de school, te presenteren in de schoolgids aan belanghebbenden⁴⁵.

Daarnaast adviseert de raad de minister om het voortouw te nemen bij de ontwikkeling van toetsmateriaal op een breed terrein en dit beschikbaar te stellen aan scholen. Scholen kunnen dit vervolgens gebruiken om aan de verplichtingen te voldoen voortkomend uit de WPO en de genoemde AMvB. Gebruik van toetsmateriaal en het vaststellen van de toegevoegde waarde voor de verantwoording van het schoolbeleid kunnen tevens de ontwikkeling, implementatie en verdere vervolmaking van een schoolintern kwaliteitsstelsel positief beïnvloeden, en vice versa.

Toekomstbeeld

Voor de uitvoering stelt de raad zich in de toekomst een model voor waarbij er sprake is van een databank die gevuld wordt met een brede variatie aan begin- en eindtoetsen. Voor elk domein en voor elke leeftijdsgroep komen zo verschillende toetsen beschikbaar. Op deze manier wordt een mogelijkheid geboden om toetsen te kiezen die bij de eigen onderwijskundige opvatting passen. Nodig is dat elke meting zogenoemde *kernitems* bevat. Dit zijn items die in elke toets voorkomen en waarvan de relatie met andere items bekend is. Enerzijds kan daardoor elke school zijn eigen toets samenstellen en anderzijds blijft de toegevoegde waarde van scholen vergelijkbaar. Om rekening te houden met levensbeschouwelijke en andere door de gemeenschap belangrijk geachte aspecten, adviseert de raad om een groep bestaande uit ouders, leraren, schoolleiders en vertegen-

44 *Op grond van de beginmeting wordt geschat wat het verwachte eindniveau op een bepaald leergebied zal zijn. Het verschil tussen het gemeten en het verwachte eindniveau per leerling is input voor een complexe statistische berekening, die moet leiden tot de bepaling van toegevoegde waarde op het desbetreffende leergebied. In de statistische berekening worden achtergrondvariabelen van de leerling verdisconteerd.*

45 *Indien nodig kan ook een bepaling in de WPO overwogen worden.*

woordigers van de verschillende besturen- en managementorganisaties bij het opstellen van het implementatietraject te betrekken.

De raad adviseert voor afname van de toetsen een protocol op te stellen dat een valide uitvoering garandeert. In formatieve zin zal rekening gehouden moeten worden met de extra taken die scholen moeten uitvoeren. De raad adviseert tevens om ondersteuning voor scholen te organiseren voor de statistische bewerking van de gegevens. Een nog nader te bepalen organisatie krijgt als taak het berekenen van de toegevoegde waarde, het vergelijken van deze waarde met die van andere scholen en het rapporteren aan de school. Scholen blijven in een dergelijk model zelf verantwoordelijk voor publicatie van de gegevens, maar publicatie is wel verplicht. Deze publicatie omvat een evenwichtige combinatie van proces- en productprestaties. Zo kan de toegevoegdewaardescore worden geplaatst in het kader van schoolspecifieke kenmerken. De horizontale verantwoordingslijn wordt op deze manier versterkt en zowel interne als externe actoren (ouders, overheid, bestuur) kunnen de gegevens gebruiken als basis bij het nemen van beslissingen.

Rol Inspectie

Ten aanzien van de verantwoording van de scholen over het onderwijs aan de overheid merkt de raad op, dat de overheid een keuze moet maken in de wijze waarop zij het toezicht organiseert. Zoals al eerder aangegeven, zal de verticale verantwoordingslijn van karakter moeten veranderen indien de horizontale verantwoordingslijn wordt verstevigd. In het toezichtkader van de Inspectie van het Onderwijs zal de beleidskeuze van aanbod naar opbrengst moeten doorklinken. Versobering van het toetsingskader voor wat betreft de inrichting van het onderwijsleerproces (artikel 10, lid 2b van de WOT) zal gefaseerd plaats kunnen vinden parallel aan de voortgang van het door de raad voorgestelde ontwikkeltraject.

4.1.1 Veertien aanbevelingen

De raad geeft in dit advies een aantal aanbevelingen die niet in één keer gerealiseerd kunnen worden. De raad stelt dan ook voor om een en ander verder uit te werken in een ontwikkeltraject van vier à zeven jaar, waarin de aanbevelingen van de raad worden uitgewerkt. Samengevat constateert de raad:

Ten aanzien van de toetsen:

- (1) Voor bepaling van de toegevoegde waarde is ten minste een begin- en een eindmeting noodzakelijk.
- (2) Voor zowel de begin- als een eindmeting dienen toetsen te worden ontwikkeld.
- (3) Ten aanzien van de toetsdomeinen pleit de raad voor een brede benadering. Dit wil zeggen dat het toetsinstrumentarium alle leergebieden van de kerndoelen dient af te dekken. Vanwege de te verwachte complexiteit adviseert de raad te starten met ontwikkeling en gebruik van toetsen op de leergebieden Nederlands en rekenen/wiskunde, gevolgd door de overige leergebieden.

Ten aanzien van de meting en de verwerking van gegevens:

- (4) Bepaling van het beginniveau dient plaats te vinden rondom het zesde levensjaar.
- (5) Voor vaststelling van de toegevoegde waarde dienen achtergrondvariabelen van leerlingen te worden meegenomen.

- (6) Bij de uitvoering stelt de raad zich in de toekomst een model voor waarbij er sprake is van een databank die gevuld wordt een brede variatie aan begin- en eindtoetsen. Voor elk domein en voor elke leeftijdsgroep komen zo verschillende toetsen beschikbaar. Op deze manier wordt een mogelijkheid voor variatie ingebouwd waardoor een school toetsen kan kiezen die bij de eigen onderwijskundige schoolopvatting passen.
- (7) Voor afname van de toetsen wordt een protocol opgesteld dat een valide afname garandeert.
- (8) Om de gegevens statistisch te bewerken wordt ondersteuning aangeboden op een centraal punt. Een nog nader te bepalen organisatie berekent de toegevoegde waarde, vergelijkt deze met andere scholen en rapporteert dit aan de school. Scholen zijn vervolgens zelf verantwoordelijk voor publicatie van de gegevens.

Ten aanzien van wet- en regelgeving:

- (9) Voor de uitvoering adviseert de raad de minister nauw aan te sluiten bij de bepalingen in de WPO over het kwaliteitsbeleid en de horizontale verantwoordingslijn te verstevigen.
- (10) Verder adviseert de raad om bij AMvB als bedoeld in artikel 13 van de WPO, eerste lid onder a, voorschriften op te stellen om gegevens over de toegevoegde waarde van het onderwijs, samen met gegevens over de inrichting en het aanbod van de school, te presenteren in de schoolgids aan belanghebbenden.
- (11) In formatieve zin zal de overheid scholen moeten compenseren voor de extra taken die zij moeten uitvoeren.
- (12) In het toezichtkader van de Inspectie van het Onderwijs zal de beleidskeuzen ten aanzien van het toezicht van aanbod- naar opbrengstcontrole moeten doorlinken.

Ten aanzien van het ontwikkeltraject:

- (13) Om dit alles te realiseren stelt de raad voor om in het beginstadium van het uit te zetten ontwikkeltraject ruimte te maken voor de uitvoering van een pilot om één en ander uit te proberen en te vervolmaken, toetsen te ontwikkelen in samenspraak met de betrokkenen en neveneffecten goed in kaart te brengen. Een gedegen communicatie- en implementatieplan dient aan deze activiteiten ten grondslag te liggen.
- (14) Om rekening te houden met levensbeschouwelijke en andere door de gemeenschap belangrijk geachte aspecten, adviseert de raad om bij het ontwerpproces een groep te betrekken bestaande uit ouders, leraren, schoolleiders en vertegenwoordigers van de verschillende besturen- en managementorganisaties.

4.2 Beantwoording van de adviesvragen

In het voorgaande zijn de antwoorden op de vragen van de minister in brokstukken geformuleerd. In deze paragraaf wordt vraags- en puntsgewijs antwoord gegeven op de vragen en deelvragen.

1 *Is het wenselijk is dat de toegevoegde waarde van een school wordt bepaald en wat is de functie daarvan?*

Ja, de wenselijkheid daarvan vloeit voort uit de huidige besturing van het onderwijs door de overheid. Grofweg kan worden gesteld dat enkele decennia geleden de overheid een weg is ingeslagen waarbij in de sturing van het onderwijs een sterkere nadruk is komen te liggen op de opbrengsten en het resultaat in plaats van op het aanbod. De wens om te komen tot het vaststellen van de toegevoegde waarde, als zijnde de opbrengst van het onderwijs in scholen, is daar een consequentie van. Een andere consequentie is dat scholen zelf verantwoordelijk zijn voor het formuleren van doelen in termen van te realiseren onderwijskwaliteit, de realisatie en bewaking ervan en de publieke verantwoording erover. Gegevens over de toegevoegde waarde kan de school derhalve ook gebruiken voor de interne schoolverbetering.

Indien vanuit deze gedachtengang over sturing van het onderwijs, de overheid overgaat tot het opstellen van voorschriften over de verantwoording van de toegevoegde waarde, dan heeft dit ook gevolgen voor de rol van de Inspectie van het Onderwijs. Deze voorschriften zijn immers gericht op het versterken van de positie van de ouders (horizontale verantwoordingslijn). Daardoor zal de verticale verantwoordingslijn van karakter moeten veranderen en ook meer gericht moeten zijn op aanbod in plaats van opbrengst.

De minister vraagt om bij de beantwoording van deze vraag in te gaan op de voor- en nadelen die zijn verbonden aan het vaststellen van de toegevoegde waarde van scholen, alsmede op de effecten van meten en bepalen van de toegevoegde waarde van het onderwijs. In het verlengde hiervan vraagt de minister de vraag te beantwoorden welke domeinen of leergebieden aan de orde zouden moeten zijn bij het vaststellen van de toegevoegde waarde.

De voor- en nadelen van bepaling van de toegevoegde waarde van het onderwijs alsmede de effecten van meten en bepalen van de toegevoegde waarde zijn uitgebreid aan de orde gekomen in paragraaf 3.2. In het algemeen geldt dat een aantal effecten zoals een toename van selectie, een sterkere focus op examens en op de groep leerlingen die zich bevindt op de grens tussen twee prestatieniveaus, demotivatie van leraren en een afname van de samenwerkingsbereidheid tussen scholen onaanvaardbaar is. Om hier een oplossing voor te bieden wordt voorgesteld om naast gegevens over de toegevoegde waarde van het onderwijs ook gegevens over het aanbod van het onderwijs en procesprestaties door scholen te doen publiceren, zodat een integraal en afgewogen beeld van de school naar voren komt.

Ten aanzien van de domeinen waarop de toegevoegde waarde betrekking moet hebben, stelt de raad dat een brede meting de voorkeur verdient. Hij wijst hierbij wel op problemen bij gebieden waarover geen hoge mate van consensus bestaat als het gaat om het ontwikkelen van toetsen en de vergelijkbaarheid van de scores op deze toetsen.

2 *Is het noodzakelijk dat een begintoeets voor vierjarigen een rol speelt in de bepaling van de toegevoegde waarde?*

Om de toegevoegde waarde te kunnen bepalen is het noodzakelijk dat er een beginmeting plaatsvindt. Deze beginmeting hoeft niet per se op het vierde levensjaar plaats te

vinden, maar tussen het vierde en zevende levensjaar. Concreet stelt de raad voor om de beginmeting rond het begin van het zesde levensjaar af te nemen. De beginmeting moet bestaan uit het met een toets vaststellen van het niveau van beheersing op elk van de zes leergebieden uit de kerndoelen. Hoe de beginmeting moet worden uitgevoerd moet in een ontwikkeltraject worden uitgewerkt. Uit de analyses die zijn uitgevoerd concludeert de raad voorts dat naast de begintoets ook achtergrondvariabelen betrokken moeten worden bij de vaststelling van de toegevoegde waarde, alsmede meting van algemene intelligentie.

Ten slotte handelt dit advies over de mogelijke sturingsinstrumenten. Gegevens over de kwaliteit van het onderwijs krijgen immers pas betekenis indien ze worden ingebed in een beleidscontext met verschillende beleidsinstrumentaties. De vraag die bij advisering beantwoord wordt is:

- 3 *Welke vormen van beleidsinstrumentatie kunnen door de overheid worden verbonden met een systeem van toegevoegde waarde waarvan een begintoets onderdeel is?*

Zoals aangegeven in hoofdstuk 1 is er sprake van een mix van beleidsinstrumenten als het gaat over de sturing van het onderwijs. Zo kan er ten aanzien van het kwaliteitsbeleid worden vastgesteld dat er op het niveau van de overheid *regulerende maatregelen* (bijvoorbeeld de verplichting om een schoolplan en schoolgids op te stellen), *ruilmaatregelen* (bijvoorbeeld bekostigingsvoorwaarden) en *overtuigingsinstrumenten* (bijvoorbeeld rapporten van de Inspectie van het Onderwijs) zijn ingezet.

De raad is van mening dat het voor een versterking van de bij autonomie behorende vrijheid, noodzakelijk is om transparante informatie te verstrekken aan met name ouders. De raad adviseert derhalve om de horizontale verantwoordingslijn te verstevigen en om bij AMvB als bedoeld in artikel 13 van de WPO, eerste lid onder a, voorschriften op te stellen over de opbrengstenverantwoording in de schoolgids. De Inspectie kan vervolgens deze gegevens op macroniveau bewerken en publiceren.

Literatuur

- Assessment of Literacy and Numeracy in the Early Years of Schooling - an Overview.*
Geraadpleegd 2 juli 2003, via Commonwealth Department of Education, Science & Training, <http://www.dest.gov.au/schools/literacy&numeracy/publications/earlyyears/Overviewpaper.pdf>.
- Boes, A. (2002). Toetsen, zin en onzin. Een kritische beschouwing. *Jeugd in School en Wereld*, 87(4), 18-21.
- Boes, A. & Both, K. (2002). *Een streep door de eindtoets. Toelichting en achtergronden*. Valthe: Netwerk SOVO.
- Bosker, R., Béguin, A. & Rekers-Mombarg, L. (2001). Hoe meten we de prestatie van een school? In A.B. Dijkstra, S. Karsten, R. Veenstra & A.J. Visscher (red.), *Het oog der natie: scholen op rapport. Standaarden voor de publicatie van schoolprestaties* (121-135). Assen: Koninklijke Van Gorcum.
- Bruijn, J.A. de (2002a). Outputsturing in publieke organisaties; over het gebruik van een product- en een procesbenadering. *M & O tijdschrift voor organisatiekunde en sociaal beleid*, 32, 5-21.
- Bruijn, J.A. de (2002b). *Prestatiemeting in de publieke sector. Tussen professie en verantwoording*. Utrecht: Lemma.
- Commissie Indicatiestelling Onderwijsachterstanden (1996). *Zo onvoorspelbaar als het leven zelf. Advies*. Zoetermeer: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Coster, A.H.N. (2002). De Cito-eindtoets. De zwakke kanten van een goed instrument. *Jeugd in School en Wereld*, 87(4), 15-17.
- Dam, G. ten, Volman, M., Westerbeek, K., Wolfram, P., Ledoux, G. & Peschar, J. (2003). *Sociale competentie langs de meetlat. Het bestrijden en voorkomen van onderwijsachterstand. Het evalueren en meten van sociale competentie*. Den Haag: Transferpunt Onderwijsachterstanden.
- Dyck, M. van & Leune, H. (2001). Schoolprestatie-indicatoren en de rol van de overheid. In A.B. Dijkstra, S. Karsten, R. Veenstra & A.J. Visscher (red.), *Het oog der natie: scholen op rapport. Standaarden voor de publicatie van schoolprestaties* (190-207). Assen: Koninklijke Van Gorcum.
- Dronkers, J. (1999). Veranderden leerlingaantallen in het voortgezet onderwijs in het schooljaar 1998-1999 door de publicatie van inspectiegegevens en de berekening van het schoolcijfer door trouw in oktober 1997? Een nadere analyse. *Tijdschrift voor Onderwijsresearch*, 24(1), 63-66.
- Dronkers, J. & R. Veenstra (2001). Schoolprestatie-indicatoren in het voortgezet onderwijs: start, reacties en vervolg. In A.B. Dijkstra, S. Karsten, R. Veenstra & A.J. Visscher (red.), *Het oog der natie: scholen op rapport. Standaarden voor de publicatie van schoolprestaties* (21-35). Assen: Koninklijke Van Gorcum.
- Evers, A. (red.) (2002). *COTAN Testboek voor het onderwijs*. Amsterdam: NDC-Boom.
- Fiske, F.B. & Ladd, H.F. (2000). *When schools compete: a cautionary tale*. Washington D.C.: Brookings Institution Press.

- Foundation stage profile handbook. Geraadpleegd op 2 juli 2003 via The Qualifications and Curriculum Authority, <http://www.qca.org.uk/ca/foundation/profiles.asp#handbook>
- Geerligs, J., Kops, Y., Boer, P. den & Veen, F. van der (2002). *Kwalificatiewinst in de beroepsonderwijskolom. Nulmeting*. Wageningen: Stoas Onderzoek.
- Goldstein, H., Huiqi, P., Rath, T. & Hill, N. (2000). *The use of value added information in judging school performance*. Geraadpleegd op 2 juli 2003 via <http://www.ioe.ac.uk/hgpersonal/Using-value-added-information.pdf>
- Groot, W. & Maassen van den Brink, H. (2003). *Investeren en terugverdienen, Kosten en baten van onderwijsinvesteringen*. Den Haag: Sectorbestuur Onderwijsarbeidsmarkt.
- Hiemstra, J. (2002). Als we allemaal hetzelfde weten, dan weten we ook hetzelfde niet. *Jeugd in School en Wereld*, 87(4), 6-9.
- Hooge, E. (1998). *Ruimte voor beleid, autonomievergroting en beleidsuitvoering door basisscholen*. Amsterdam: SCO-Kohnstamm Instituut/Universiteit van Amsterdam.
- Inspectie van het Onderwijs (2003). *Onderwijsverslag over het jaar 2002*. Utrecht: Inspectie van het Onderwijs.
- Karsten, S., Visscher, A. & Jong, T. de (1999). *Handle with care! Buitenlandse ervaringen met het publiceren van schoolprestatiegegevens*. Den Haag: Onderwijsraad.
- Leune, J.M.G. (2000). *Beleid vraagt om goed meten, goed meten vraagt om beleid. Inleiding*. Den Haag: Onderwijsraad.
- Meijnen, G.W. (1997). Indicatiestelling onderwijsachterstanden: een commissie rapporteert. *Tijdschrift voor Onderwijsresearch*, 33(3), 186-189.
- Meisels, S.J., Atkins-Burnett, S., Xue, Y., DiPrima Bickel, D. & Seung-Hee, S. (2003). Creating a system of Accountability: The Impact of Instructional Assessment on Elementary Children's Achievement Test Scores. *Education Analysis Archives*, 11(9). Geraadpleegd op 6 maart 2003 via <http://epaa.asu.edu/epaa/v11n9/>.
- Ministerie van Onderwijs, Cultuur en Wetenschappen (2000). *Onderwijs in stelling. Kracht en creativiteit voor de kennissamenleving*. Zoetermeer: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Mooij, T. (2000). Screening Children's Entry Characteristics in Kindergarten. *Early Child Development and Care*, 165, 23-40.
- National Institute of Child Health and Human Development (2002). Early Child Care and children's Development Prior to School Entry: Results from the Study of Early Child Care. *American Educational Research Journal*, 39(1), 133-164.
- Onderwijsraad (1999a). *Zeker weten. Leerstandaarden als basis voor toegankelijkheid*. Den Haag: Onderwijsraad.
- Onderwijsraad (1999b). *Schoolkwaliteit in beeld. Voorstellen voor een verantwoorde openbaarmaking van gegevens over de kwaliteit van scholen*. Den Haag: Onderwijsraad.
- Onderwijsraad (2000a). *Dereguleren met beleid. Studie naar effecten van deregulering en autonomievergroting*. Den Haag: Onderwijsraad.
- Onderwijsraad (2000b). *Onderwijsbeleid sinds de jaren zeventig. Werkdocument bij het advies 'Dereguleren met beleid, studie naar effecten van deregulering en autonomievergroting'*. Den Haag: Onderwijsraad.
- Onderwijsraad (2001a). *De Markt Meester? Een verkenning naar marktwerking in het onderwijs*. Den Haag: Onderwijsraad.

- Onderwijsraad (2001b). *De rugzak gewogen. Over de bekostiging van WSNS, LGF en de gewichtenregeling basisonderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2001c). *Wat 't zwaarst weegt... Een nieuwe aanpak voor het onderwijs achterstandenbeleid*. Den Haag: Onderwijsraad.
- Onderwijsraad (2001d). *Publiek en Privaat: Mogelijkheden en gevolgen van private middelen in het publieke onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2002a). *De kern van het doel. Reactie op het advies van de commissie Wijnen over de kerndoelen basisonderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2002b). *Over leerlinggewichten en schoolgewichten*. Den Haag: Onderwijsraad.
- Onderwijsraad (2002c). *Wat Scholen vermogen*. Den Haag: Onderwijsraad.
- Roeleveld, J. (2003). *Herkomstkenmerken en Begintoets. Secundaire analyses op het PRIMA-Cohortonderzoek*. Ongepubliceerd onderzoek.
- Sammons, P. (2001). *Fairer Comparisons of Schools: The role of school effectiveness research in promoting improvement*. Geraadpleegd op 2 juli 2003 via http://www.eu2001.se/education/eng/docs/karlstad_sammons.pdf.
- Saunders, L. (1999). A Brief History of Educational 'Value Added': How Did We Get To Where We Are? *School Effectiveness and School Improvement*, 10(2), 233-256. *School Directors' Handbook. Value Added Assessment*. Geraadpleegd op 13 maart 2003 via Evergreen Freedom Foundation, <http://www.effwa.org/pdfs/Value-Added.pdf>.
- Sluijter, C. & Rosier, W. (2002). Goede toetsen maken is een kunst. Toetsen en leerling volgssystemen. *Jeugd in School en Wereld*, 87(4), 12-14.
- Stb. 2002-387. Wet op het onderwijstoezicht.
- TK 2000-2001, 27 400 VIII, nr. 92. Ministerie van Onderwijs, Cultuur en Wetenschappen: *Kwaliteit aan de basis, een actieprogramma voor het primair en voortgezet onderwijs*.
- TK 2001-2002, 28 375, nr. 5. Brief van de informateur aan de voorzitter van de Tweede Kamer, d.d. 3 juli 2002.
- TK 2002-2003, 28 600 VIII, nr. 116. Brief van de minister van Onderwijs, Cultuur en Wetenschappen aan de voorzitter van de Tweede Kamer, d.d. 3 februari 2003.
- TK 2002-2003, 28637, nr. 19. Brief van de informateurs aan de voorzitter van de Tweede Kamer, d.d. 16 mei 2003.
- Velden, R. van der (2001). The proof of the pudding is in the eating. Een heuristisch kader voor de kwaliteitsbeoordeling van scholen. In A.B. Dijkstra, S. Karsten, R. Veenstra & A.J. Visscher (red.), *Het oog der natie: scholen op rapport. Standaarden voor de publicatie van schoolprestaties* (65-78). Assen: Koninklijke Van Gorcum.
- Vijlder, F.J. de & Westerhuis, A. (2002). *Meervoudig publieke verantwoording*. Amsterdam: Max Goote Kenniscentrum voor Beroepsonderwijs en Volwasseneneducatie.
- Visser, M. (2003). *Schoolkeuzeprocedure, kwaliteit en kwaliteitszorg: een project onderwijsvernieuwing in Amsterdam, 1994-2001*. Proefschrift. Groningen: Rijksuniversiteit Groningen.
- Website Citogroep. Geraadpleegd 23 juli 2003 via Citogroep, <http://www.citogroep.nl>.
- Website The Times Educational Supplement. Geraadpleegd op 27 maart 2003 via <http://www.tes.co.uk>.
- Zollers & Ramanathan, (1998). For- Profit Charter Schools and Students with Disabilities: The Sordid Side of the Business of Schooling. *Phi Delta Kappan*, 1998, 297-304.

Bijlage 1

Adviesaanvraag

De voorzitter van de Onderwijsraad,
prof.dr. A.M.L. van Wieringen
Nassaulaan 6
2514 JS Den Haag

Uw brief van	Ons kenmerk	Contactpersoon	Zoetermeer
	PO/KB/03/8463	dhr. Van Dommelen	18 maart 2003
Onderwerp	Bijlage(n)	Doorkiesnummer	
begintoets basisonderwijs en toegevoegde waarde		3682	

Geachte heer Van Wieringen,

In het strategisch akkoord van juli 2002 is het voornemen opgenomen om een verplichte begintoets in te voeren voor het primair onderwijs. Deze begintoets moet in combinatie met een eindtoets zicht kunnen bieden op de toegevoegde waarde van scholen, aldus het strategisch akkoord. Dit voornemen van het nu demissionaire kabinet heeft veel discussie losgemaakt, zowel in het onderwijsveld als in de Tweede Kamer. Bij de behandeling van de onderwijsbegroting voor 2003 heeft de Tweede Kamer een motie aanvaard van de CDA-fractie, waarin de regering verzocht wordt om de begintoets niet te benutten voor externe kwaliteitsoordelen over scholen. Vanuit de kring van het traditioneel vernieuwingsonderwijs heb ik brieven ontvangen met als strekking dat men vreest voor een versmalling van het curriculum als de kwaliteit van scholen uitgedrukt gaat worden in de resultaten op begin- en eindtoetsen. Deze signalen wil ik serieus nemen. Dat betekent echter niet dat alle denken over het gebruik van toetsen voor de kwaliteitsbepaling stopgezet dient te worden. Wel is er behoefte aan een meer fundamentele doordenking van dit onderwerp. Ik verzoek de Onderwijsraad daarom mij hierover te adviseren. Meer in het bijzonder verzoek ik de raad in te gaan op de volgende vragen.

1. Is het wenselijk dat de toegevoegde waarde van een school wordt bepaald en zo ja, wat is de functie daarvan?

Op grond van de Wet op het onderwijstoezicht brengt de inspectie de kwaliteit van afzonderlijke scholen op diverse aspecten van kwaliteit in kaart. Eén van die aspecten betreft de opbrengsten. De wenselijkheid hiervan staat wat mij betreft niet ter discussie, wel de wijze waarop de opbrengsten bepaald worden en de inhoudelijke domeinen waarbinnen opbrengsten bepaald worden.

Momenteel beoordeelt de inspectie de opbrengsten van een school door de resultaten van een school op de Cito-eindtoetsen voor Nederlandse taal en rekenen/wiskunde te vergelijken met de resultaten op scholen met een vergelijkbare leerlingenpopulatie. Daartoe wordt de leerlingenpopulatie getypeerd aan de hand van de categorieën uit de gewichtenregeling. Het bovengenoemde voornemen in het strategisch akkoord van juli 2002 was gebaseerd op de gedachte dat hiervoor beter een meting gebruikt kan worden die aangeeft op welk niveau de leerlingen de basisschool binnenkomen. Naar de 'technische' haalbaarheid hiervan verricht de inspectie momenteel onderzoek (zie mijn brief aan de Tweede Kamer van 12 februari jongstleden, kenmerk PO/OO/03/4313). Voor de advisering door de Onderwijsraad is vooral de vraag van belang welke voor- en nadelen verbonden zijn aan het vaststellen van de toegevoegde waarde van scholen. Als voordelen zijn te noemen: ouders kunnen een vergelijking maken tussen scholen en daar hun schoolkeuze op baseren of een gesprek aangaan met de school van hun kinderen, en scholen worden door dergelijke gegevens (en de reacties van ouders) gestimuleerd om de kwaliteit van het onderwijs te verbeteren. Er kunnen echter ook neveneffecten optreden. Als risico's worden genoemd: scholen versmallen hun curriculum tot datgene dat getoetst wordt; ze besteden onevenredig veel tijd aan het oefenen van toetsen; ze sluiten zwakkere leerlingen uit van deelname aan de toets. Aan de Onderwijsraad wil ik vragen deze (en mogelijk andere) effecten te wegen, en op grond daarvan een advies uit te brengen over de wenselijkheid van het verzamelen en eventueel publiceren van gegevens over de toegevoegde waarde van scholen.

Daarbij verzoek ik de raad tevens de vraag te betrekken welke domeinen of leergebieden aan de orde zouden moeten zijn bij het vaststellen van de toegevoegde waarde: wat zijn de mogelijkheden en risico's van een smalle (taal en rekenen) versus een brede (cognitieve, sociaal-emotionele, kunstzinnige en lichamelijke ontwikkeling) opvatting van opbrengsten? Bij een smalle opvatting gaat het om de toegevoegde waarde bij Nederlandse taal en rekenen/wiskunde, wellicht ook wereldoriëntatie. Bij een brede opvatting wordt de toegevoegde waarde van de school in kaart gebracht bij alle (of zo veel mogelijk) aspecten van de ontwikkeling van kinderen: cognitieve prestaties, normbesef, sociaal-emotionele ontwikkeling, kunstzinnige oriëntatie, lichamelijke ontwikkeling.

2. Is het noodzakelijk dat een begintoets voor vierjarigen een rol speelt in de bepaling van de toegevoegde waarde, en zo ja, onder welke voorwaarden is dat mogelijk? Zijn er alternatieven voorhanden?

Een fundamentele afweging van voor- en nadelen is niet alleen nodig voor de mogelijke effecten van meting en eventuele publicatie van de 'toegevoegde waarde' van scholen. De wijze waarop die toegevoegde waarde wordt bepaald, verdient eveneens nadere doordenking. Ook hierbij gaat het voor de advisering door de Onderwijsraad niet om de 'technische' aspecten, zoals de wijze waarop de gegevens van een begintoets en een eindtoets bewerkt worden tot een maat voor de toegevoegde waarde. Wel is van belang dat de raad nader doordenkt of het voor de bepaling van de toegevoegde waarde noodzakelijk is toetsen af te nemen bij vierjarigen (of vijf- of zesjarigen). Omdat scholen een belang hebben bij de uitkomsten van een begintoets heb ik in de zogenaamde 'enveloppebrief' aan de Tweede Kamer (1 november 2002, kenmerk FEZ/BTA/2002/56157) aangegeven dat de afname van een

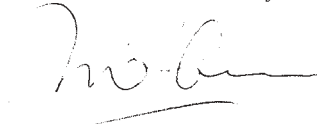
dergelijke toets 'objectief en onafhankelijk dient plaats te vinden'. Dat betekent dat met de afname hoge additionele kosten gemoeid zijn. In de enveloppebrief is aangegeven dat hiervoor 16,5 miljoen euro gereserveerd wordt. Zijn er andere, wellicht goedkopere manieren om de toegevoegde waarde van een school in kaart te brengen?

3. Welke vormen van beleidsinstrumentatie kunnen door de overheid worden verbonden aan een systeem van vaststelling van de toegevoegde waarde van scholen?

Gegevens over de toegevoegde waarde van scholen hebben geen betekenis op zich: die betekenis krijgen ze pas als ze ingebed zijn in een systeem van sturing op de kwaliteit van het onderwijs. De vraag is of verzameling en eventuele publicatie van dergelijke gegevens op zich een toereikend beleidsinstrument is. De veronderstelde werking is tweeledig: ouders oriënteren zich op dergelijke gegevens en baseren hun schoolkeuze erop of spreken de school van hun kinderen erop aan; scholen willen goed voor de dag komen en richten hun kwaliteitsbeleid op verbetering van de leerresultaten. Daarmee kan het vaststellen van de toegevoegde waarde als op zichzelf staande maatregel invloed hebben op het kwaliteitsbeleid van de scholen. Denkbaar is echter ook dat de effectiviteit van deze maatregel wordt verhoogd door de gegevens over de toegevoegde waarde onderdeel te maken van andere mogelijke beleidsinstrumenten: het gesprek tussen schoolbestuur en inspectie over de kwaliteit van de school, de aansturing vanuit de gemeente of het samenwerkingsverband Weer Samen Naar School op de effecten van het achterstands- en zorgbeleid, en dergelijke. Ik verzoek de Onderwijsraad mij te adviseren over de vormen van beleidsinstrumentatie die ingezet kunnen worden om optimale effecten te genereren voor de kwaliteit van het onderwijs.

Over de wenselijkheid en mogelijkheid voor de raad om over deze vragen advies uit te brengen, heeft op 9 januari jongstleden in oriënterende zin ambtelijk overleg plaats gehad. Het verheugt mij dat de Onderwijsraad aangegeven heeft over deze belangwekkende materie advies uit te willen brengen. Ik verzoek de Onderwijsraad het advies in juli 2003 af te ronden.

De minister van Onderwijs, Cultuur en Wetenschappen,



(Maria J.A. van der Hoeven)

Bijlage 2

Expertmeeting begintoets basisonderwijs op 20 februari 2003

Aanwezigen:

De heer prof.dr. R.J. Bosker (Universiteit Twente).

De heer drs. H.E.M. Dobbelaar (Onderwijsraad).

Mevrouw dr. A.A. van der Hoeven - van Doornum (ITS, Wetenschap voor beleid en samenleving).

De heer dr. P.P.M. Leseman (SCO-Kohnstamm Instituut).

De heer prof. dr. G.W. Meijnen (Onderwijsraad).

De heer drs. A. van der Rest (Onderwijsraad).

De heer dr. F.S.J. Riemersma (Onderwijsraad).

De heer dr. F. van der Schoot (Citogroep).

De heer dr. J.H. Uiterwijk (SCO-Kohnstamm Instituut).

Tijdens deze expertmeeting zijn de volgende onderwerpen en vragen aan de orde gekomen:

- De betekenis van toegevoegde waarde.
- De invloed van de school op de ontwikkeling van het kind.
- De domeinen waarop wordt gemeten.
- Het vergelijken van scholen op toegevoegde waarde.
- De meetbaarheid van toegevoegde waarde.
- De relatie tussen begin- en eindmeting en andere beïnvloedende variabelen.
- Differentiële effecten.
- Meten bij jonge kinderen.
- De tijdsspanne tussen twee metingen.

Bijlage 3

**Panelbijeenkomst toegevoegde waarde van het onderwijs in scholen
op 6 mei 2003**

Aanwezigen:

De heer drs. A.W. Boes (Nederlandse Jenaplanvereniging).
De heer J. Boven (Gemeente Schiedam, dienst Educatie, bureau GOA).
De heer drs. H.E.M. Dobbelaar (Onderwijsraad).
De heer drs. A.J.F. Duif (AVS, Algemene Vereniging van Schoolleiders).
De heer T. Gloudemans (Basisschool Bijlmerhorst, Amsterdam).
De heer dr. W.J.C.M. van de Grift (Inspectie van het Onderwijs).
De heer drs. J. de Haan (Besturenraad).
De heer M. van Herpen (Expertisecentrum EGO, ErvaringsGericht Onderwijs)
Mevrouw M.J. Kupers-Verschure (Basisschool Bernadette, Heeten).
De heer prof.dr. G.W. Meijnen (Onderwijsraad).
De heer drs. A. van der Rest (Onderwijsraad).
De heer dr. F.S.J. Riemersma (Onderwijsraad).
Mevrouw drs. P. Rietveld (ABC Amsterdam).
De heer prof.dr. A.M.L. van Wieringen (Onderwijsraad).

Tijdens de panelbijeenkomst zijn onder meer de volgende bespreekpunten aan de orde gekomen:

- De betekenis van deregulering en autonomievergroting voor de bewaking van de kwaliteit van het onderwijs en de rol van toegevoegde waarde daarbij.
- De betekenis van toegevoegde waarde en de bijdrage aan een meer afgewogen oordeel over de kwaliteit van een school.
- De verbetering die door het meten en vaststellen van toegevoegde waarde ten opzichte van de huidige situatie kan ontstaan.
- De wenselijkheid of onwenselijkheid dat de toegevoegde waarde wordt bepaald op alle domeinen waarvoor kerndoelen zijn opgesteld.
- De wijze van vermelding van de toegevoegde waarde in de schoolgids als zijnde een manier voor de school om verantwoording naar de buitenwereld af te leggen.
- De aanspreekbaarheid van scholen door de overheid indien de toegevoegde waarde te laag wordt bevonden.
- De aansluiting van een meetsysteem om opbrengsten vast te stellen, bij hetgeen nu in scholen gebeurt.

Bijlage 4

Afkortingen

AMvB	=	Algemene Maatregel van Bestuur
COTAN	=	Commissie Testaangelegenheden
DLE	=	Didactische Leeftijdsequivalent
ITS	=	Instituut voor Toegepaste Sociale wetenschappen
LVS	=	Leerlingvolgsysteem
OBIS	=	Onderbouwinformatiesysteem
PIPS	=	Performance Indicators in Primary Schools Baseline Assessment
WOT	=	Wet op het onderwijstoezicht
WPO	=	Wet op het primair onderwijs