

Toegevoegde waarde: een instrument voor onderwijsverbetering - niet voor beoordeling

**Toegevoegde waarde:
een instrument voor
onderwijsverbetering -
niet voor beoordeling**

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 Wetenschap en van Economische Zaken. 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.

Advies Toegevoegde waarde: een instrument voor onderwijsverbetering - niet voor beoordeling, uitgebracht aan de Minister en de Staatssecretaris van Onderwijs, Cultuur en Wetenschap.

Nr.20140091/1046, april 2014

Uitgave van de Onderwijsraad, Den Haag, 2014.

ISBN 978-94-6121-000-5

Bestellingen van publicaties:

Onderwijsraad

Nassaulaan 6

2514 JS Den Haag

email: secretariaat@onderwijsraad.nl

telefoon: (070) 310 00 00 of via de website:

www.onderwijsraad.nl

Ontwerp en opmaak:

www.balyon.com

Drukwerk:

Drukkerij Excelsior., Den Haag

© Onderwijsraad, Den Haag.

Alle rechten voorbehouden. All rights reserved.

Aan de Minister en de Staatssecretaris van
Onderwijs, Cultuur en Wetenschap
Mevrouw dr. M. Bussemaker en de heer drs. S. Dekker
Postbus 16375
2500 BJ Den Haag

Nassaulaan 6
2514 JS Den Haag
Telefoon: 070 310 00 00
Fax: 070 356 14 74
secretariaat@onderwijsraad.nl
www.onderwijsraad.nl

Ons kenmerk
20140091/1046

Contactpersoon

Plaats/datum

Den Haag, 7 april 2014

Uw kenmerk

Doorkennummer

Onderwerp

Advies Toegevoegde waarde

Mevrouw de Minister, mijnheer de Staatssecretaris,

Met genoegen biedt de Onderwijsraad u zijn advies *Toegevoegde waarde: een instrument voor onderwijsverbetering* – niet voor beoordeling aan. In dit advies geeft de Onderwijsraad antwoord op de vraag in hoeverre leerwinst en toegevoegde waarde geschikte instrumenten zijn voor verbetering van het onderwijs. Scholen hebben steeds meer gegevens over de vorderingen van leerlingen. Tegelijkertijd wordt daar tot dusver weinig gebruik van gemaakt.

De raad stelt dat leerwinst en toegevoegde waarde wél goede instrumenten zijn voor scholen om hun onderwijs te verbeteren, maar niet voor een oordeel over de kwaliteit van een school. Gegevens over leerwinst geven scholen meer inzicht in de ontwikkeling van hun leerlingen. Met gegevens over toegevoegde waarde kunnen ze hun opbrengsten ook spiegelen aan die van andere scholen. Als beoordelingsmaat is toegevoegde waarde echter ongeschikt. Het is een te beperkte indicator voor de beoordeling van de opbrengsten van een school, omdat alleen de goed meetbare aspecten van het onderwijs erin kunnen worden meegenomen. Ook zijn er verschillende statistische problemen.

Naar aanleiding van zijn bevindingen pleit de raad voor versterking van het eigenaarschap van scholen over leerwinst en toegevoegde waarde door de kennis bij schoolleiders en leraren over de inzet van deze gegevens te vergroten. Verder is bij de beoordeling van de kwaliteit van scholen een betere balans nodig tussen opbrengst- en procesindicatoren.

De raad hoopt dat dit advies bijdraagt aan een verstandige benutting van opbrengstgegevens voor verbetering van de kwaliteit van het onderwijs.

Met beleefde groet,



Prof. dr. G.T.M. ten Dam
Voorzitter



Drs. A. van der Rest
Secretaris

Inhoud

Samenvatting	7
1 Meer informatie over leerprestaties creëert nieuwe instrumenten voor onderwijsverbetering	9
1.1 Aanleiding: discussie over bruikbaarheid van begrippen 'leerwinst' en 'toegevoegde waarde'	9
1.2 Adviesvraag: zijn leerwinst en toegevoegde waarde goede instrumenten voor beter onderwijs?	12
2 Advies: leerwinst en toegevoegde waarde zijn wél geschikt voor onderwijsverbetering, níet voor oordeel over de school	13
2.1 Leerwinst en toegevoegde waarde geven een genuanceerd beeld van prestaties	14
2.2 Toegevoegde waarde is een te smalle indicator voor de kwaliteit van scholen	16
2.3 Vaststellen van toegevoegde waarde vraagt om inhoudelijke keuzes	17
2.4 Er zijn statistische problemen bij het meten van leerwinst en toegevoegde waarde	21
2.5 Het gebruik van toegevoegde waarde als beoordelingsinstrument kan ongewenst strategisch gedrag uitlokken	23
3 Aanbeveling 1: geef de school eigenaarschap over de instrumenten leerwinst en toegevoegde waarde	25
3.1 Vergroot de expertise op scholen over opbrengsten en onderzoek	25
3.2 Stel gegevens beschikbaar voor onderzoek	27
4 Aanbeveling 2: houd bij oordeel over de school rekening met tekortkoming van opbrengstindicatoren	28
4.1 Laat in het toezicht de kwaliteitszorg van scholen zwaarder wegen	28
4.2 Verbeter bestaande opbrengstindicatoren door opleiding van ouders op te nemen	30
4.3 Laat scholen het succes van hun leerlingen in het vervolgonderwijs volgen	31
Afkortingen	33
Literatuur	34
Geraadpleegde deskundigen	36
Bijlage 1	
Adviesvraag	39
Bijlage 2	
Wat meten leerwinst en toegevoegde waarde?	43
Bijlage 3	
Verskillende modellen van opbrengstindicatoren	47

Samenvatting

In hoeverre zijn gegevens over 'leerwinst' en 'toegevoegde waarde' bruikbaar voor het verhogen van de kwaliteit van het onderwijs? Leerwinst is de groei die leerlingen doormaken. Toegevoegde waarde is de bijdrage die scholen leveren aan de gemiddelde groei van hun leerlingen. Om de toegevoegde waarde in een bepaald vak te kunnen vaststellen, moeten toetsresultaten over een langere periode met elkaar én met de achtergrondgegevens van leerlingen in verband worden gebracht.

De raad vindt dat leerwinst en toegevoegde waarde goede instrumenten zijn om scholen te helpen hun onderwijs te verbeteren. Ze zijn echter niet bruikbaar voor een oordeel over de kwaliteit van de school als geheel.

Gegevens over leerwinst geven scholen meer inzicht in de ontwikkeling die leerlingen doormaken bij de gemeten vakken of leerdomeinen. Op basis daarvan kunnen scholen beter maatwerk leveren. Toegevoegde waarde is een goed instrument om de eigen opbrengsten te vergelijken met andere scholen, om er zo nog beter op te kunnen reflecteren. Het is echter een te smalle indicator voor een beoordeling van de school. Toegevoegde waarde zegt alleen iets over die dimensies van het onderwijs die gemeten kunnen worden met valide toetsen, zoals taal en rekenen. De brede ontwikkeling van leerlingen, waarbij ook non-cognitieve vaardigheden een rol spelen, wordt niet meegenomen. Bovendien is er niet één manier om toegevoegde waarde vast te stellen: er zijn inhoudelijke keuzes nodig (bijvoorbeeld voor de achtergrondkenmerken waarvoor gecorrigeerd wordt) en er zijn statistische beperkingen die maken dat de uitkomsten altijd in hun context moeten worden gezien. Tot slot kan toegevoegde waarde als beoordelingsinstrument ongewenst strategisch gedrag uitlokken. Scholen kunnen zich vooral gaan richten op toetsbare onderwijsonderdelen en andere veronachtzamen.

Aanbeveling 1: geef de school eigenaarschap over de instrumenten leerwinst en toegevoegde waarde

Voor een goed gebruik van leerwinst en toegevoegde waarde voor onderwijsverbetering moeten scholen eigenaar worden van de instrumenten. Schoolleiders en leraren hebben meer kennis nodig van statistiek en praktijkgericht onderzoek. De raad pleit voor de inzet van data-teams (waarin schoolleiders en docenten samen met een onderzoeker werken aan oplossingen voor concrete vragen) en voor een meer onderzoekende houding op scholen. Daarnaast pleit hij voor de beschikbaarheid van gegevens over leerwinst en toegevoegde waarde voor onderzoek. Dit bevordert het denken erover en de verdere ontwikkeling van gebruiksmogelijkheden.

Aanbeveling 2: houd bijoordeel over de school rekening met tekortkoming van opbrengstindicatoren

Het toezicht moet gericht zijn op méér dan wat opbrengstindicatoren meten, dus op de brede kwaliteit van het onderwijs, en het moet bevorderen dat deze brede kwaliteit verbetert. Daarvoor is een betere balans nodig tussen opbrengst- en procesindicatoren – nu ligt de nadruk te veel op de eerste. Scholen moeten ruimte krijgen om een inhoudelijke dialoog te voeren met de Inspectie over de omstandigheden die hebben geleid tot de resultaten. Daarnaast vindt de raad dat het opleidingsniveau van ouders beter tot uitdrukking kan komen in bestaande opbrengstindicatoren. Tot slot kunnen scholen zelf het succes van hun leerlingen in het vervolgonderwijs (en in de maatschappij) beter volgen.

Scholen hebben steeds meer informatie over hoe hun leerlingen leren en welke vorderingen ze maken. Hoe kunnen deze gegevens optimaal worden benut om het onderwijs te verbeteren? En zijn ze ook bruikbaar om een oordeel te geven over de school?

1 Meer informatie over leerprestaties creëert nieuwe instrumenten voor onderwijsverbetering

1.1 Aanleiding: discussie over bruikbaarheid van begrippen 'leerwinst' en 'toegevoegde waarde'

Mede dankzij het gebruik van ict (informatie- en communicatietechnologie) is er in het onderwijs steeds meer informatie beschikbaar over de resultaten die leerlingen boeken, en over hoe die zijn behaald. Scholen kunnen met al die gegevens systematisch en doelgericht werken aan beter onderwijs.¹ Voor dit zogenoemde opbrengstgericht werken moeten ze zicht hebben op leerlingresultaten, zowel op individueel niveau als op groeps- en schoolniveau. Dat krijgen ze steeds meer. De invoering van de centrale eindtoets in het basisonderwijs draagt daaraan bij. In het voortgezet onderwijs wordt gewerkt aan de invoering van een diagnostische tussentijdse toets voor Nederlands, Engels en wiskunde, en aan een rekentoets.² Het middelbaar beroepsonderwijs werkt aan centrale toetsen taal en rekenen. Voor opbrengstgericht werken zijn de juiste gegevens alleen niet genoeg; scholen moeten ook leren hoe ze met deze gegevens om moeten gaan en hoe ze deze kunnen inzetten. De mate waarin scholen opbrengstgericht werken varieert nog sterk, maar duidelijk is dat steeds meer instellingen zich in deze richting ontwikkelen. Scholen in het primair en in het voortgezet onderwijs werken, onder andere via projecten als *School aan Zet*, aan de professionalisering van leerkrachten en schoolleiders op het gebied van opbrengstgericht werken.³

De gegevens over leerlingresultaten kunnen dienen als intern hulpmiddel op scholen – om zicht te krijgen op prestaties en om daarmee de leeropbrengsten te verbeteren. De overheid kan ze mogelijk ook gebruiken om beter inzicht te krijgen in de prestaties van afzonderlijke scholen en om daarmee de onderwijskwaliteit als geheel te kunnen verbeteren. Een probleem daarbij is dat toetsscores op zich niet goed vergelijkbaar zijn; het startniveau van leerlingen en hun achtergrond zijn in hoge mate bepalend voor het resultaat. De overheid verkent momenteel de mogelijkheden om de leerwinst van leerlingen en de toegevoegde waarde van scho-

1 Onderwijsraad, 2011a.

2 De rekentoets is in het schooljaar 2013-2014 ingevoerd. Vanaf 2015-2016 telt het cijfer mee bij de beoordeling van het examen, tot die tijd komt het cijfer alleen op een bijlage bij de cijferlijst.

3 Zie www.schoolaanzet.nl.

len te bepalen. Leerwinst is de groei die leerlingen doormaken.⁴ Toegevoegde waarde is de bijdrage die scholen leveren aan de gemiddelde groei van hun leerlingen. Om die te kunnen vaststellen, is méér nodig dan momentopnames van de leerprestaties in de vorm van losse toetsresultaten. Losse toetsresultaten geven zicht op het niveau van leerlingen op een bepaald moment, niet op hun ontwikkeling in een bepaalde periode. Ook als wel meerdere meetmomenten beschikbaar zijn, kan alleen leerwinst gemeten worden als dezelfde inhouden of vaardigheden gemeten worden. Hoe verder twee meetmomenten in de tijd van elkaar af liggen (bijvoorbeeld groep 3 en groep 8), hoe vaker dat niet het geval is. Maar zelfs als leerwinst wel goed gemeten kan worden, is daarmee de toegevoegde waarde van de school nog niet vastgesteld. Leerwinst is namelijk mede afhankelijk van de achtergrond van leerlingen, zoals aanleg en thuissituatie. Om de toegevoegde waarde vast te kunnen stellen, moeten toetsresultaten over een langere periode met elkaar én met de achtergrondgegevens van leerlingen in verband worden gebracht.

Een dergelijke aanpak heeft een aantal praktische en meer fundamentele problemen. Daarom heeft de overheid een verkenning laten uitvoeren naar het gebruik van leerwinst en toegevoegde waarde in het primair en het voortgezet onderwijs, om duidelijk te krijgen wat de mogelijkheden zijn. Op een aantal pilotscholen wordt gebruikgemaakt van toetsgegevens in combinatie met een leerlingvolgsysteem. Ook wordt onderzocht hoe scholen dergelijke opbrengstgegevens kunnen gebruiken voor hun eigen beleid om de kwaliteit op school- en leerlingniveau te verbeteren.⁵

Leerwinst: andere focus dan het leerlingvolgsysteem

Door alle toetsgegevens uit het leerlingvolgsysteem naast elkaar te bekijken, ontstaat in principe een beeld van de ontwikkeling van de leerling. De afzonderlijke toetsen zijn echter primair gericht op een vergelijking van de prestaties van leerlingen onderling of van de prestaties met een specifieke norm voor een toets. Leerwinst richt zich primair op de groei die een leerling doormaakt. Toetsresultaten zullen altijd schommelen van toets naar toets. Leerwinst brengt in beeld in welke mate sprake is van systematische groei.

De definities van leerwinst en toegevoegde waarde aangescherpt

Onder leerwinst verstaat de raad: de toename van vaardigheden of kennis van leerlingen (of van groepen leerlingen) gedurende een bepaalde periode. De winst weerspiegelt dus het verschil in kennis van leerlingen op twee verschillende momenten. Leerwinst kan over bijvoorbeeld een half jaar, een jaar, de gehele schoolperiode, of over een schooloverstijgende periode worden vastgesteld. Conceptueel ligt het voor de hand om leerwinst te zien als het verschil in scores tussen een toets aan het einde en een toets aan het begin van de periode. Uiteraard heeft de leerwinst alleen betrekking op die vakken en leerdomeinen die gemeten zijn.

De toegevoegde waarde van de school is de bijdrage van de instelling aan de leerwinst van zijn leerlingen. Om die vast te stellen, wordt leerwinst eerst gemiddeld per school. Vervolgens moet onderscheid worden gemaakt tussen de bijdrage die de school levert aan de leerwinst en de bijdrage van anderen (het ouderlijk milieu, de samenstelling van de schoolpopulatie, de wijk waarin de school staat). Verder moet men bepalen hoe een stijging van toetscijfers op verschillende delen van een schaal wordt gewaardeerd. Heeft een toename van een 3 naar een 5

4 Ehlert, Koedel, Parsons & Podgursky, 2014.

5 Janssens, Rekers-Mombarg & Lacor, 2014. De uitkomsten van de pilots voor het voortgezet onderwijs worden aan het einde van 2014 verwacht.

dezelfde toegevoegde waarde als een stijging van een 7 naar een 9? Als de nadruk sterk ligt op het basisniveau van leerlingen, dan kan de toename van 3 naar 5 zwaarder wegen; als de nadruk sterk ligt op excellentie, dan zou de stijging van 7 naar 9 meer gewicht kunnen krijgen.

Toegevoegde waarde (zoals hierboven gedefinieerd) is een indicator voor leeropbrengsten, maar het is niet de enige. Er bestaan ook andere opbrengstindicatoren; bijvoorbeeld een eindmeting waarbij gecorrigeerd wordt voor een aantal achtergrondkenmerken van leerlingen. De Inspectie maakt hiervan gebruik. Een uitgebreide beschrijving van de verschillende modellen om de opbrengst van een school in kaart te brengen is te vinden in bijlage 3.

Het nut van het meten van leerwinst en toegevoegde waarde

Leerwinst kan een beter beeld geven van de ontwikkeling van een leerling dan de ruwe score op de eindtoets. De gedachte is verder dat als de toegevoegde waarde van scholen op een adequate manier kan worden gemeten, scholen een beter beeld krijgen van hun opbrengsten in vergelijking met andere scholen – rekening houdend met de achtergronden en omstandigheden van hun leerlingen. Voor basisscholen met relatief veel kinderen van hoogopgeleide ouders zou daardoor bijvoorbeeld kunnen blijken dat de opbrengst in termen van toegevoegde waarde lager is dan de opbrengst in termen van de leerwinst. Anders gezegd: scholen met een dergelijke leerlingenpopulatie hoeven doorgaans relatief weinig toe te voegen om de ontwikkeling van hun leerlingen te stimuleren en zo hoge toetsscores en een hoge leerwinst te behalen. Andersom kan blijken dat de inspanningen die een school levert om achterstandsleerlingen naar een bepaald niveau te brengen, juist meer opleveren dan in eerste instantie naar voren komt. Hoge eindscores betekenen dus lang niet altijd een hoge leerwinst en een hoge leerwinst betekent niet altijd een hoge toegevoegde waarde; lage eindscores lang niet altijd een lage leerwinst en een lage toegevoegde waarde.⁶

Nog te starten: pilot toegevoegde waarde in het middelbaar beroepsonderwijs

In het middelbaar beroepsonderwijs start de Inspectie van het Onderwijs in 2014 een proef waarin een opbrengstindicator is: het diplomarendement in relatie tot de vooropleiding van deelnemers. De minister van OCW (Onderwijs, Cultuur en Wetenschap) wil onderzoeken of deze indicator ook bruikbaar is als onderdeel van prestatiebekostiging vanaf 2016.⁶ Bij deze indicator in het middelbaar beroepsonderwijs is geen sprake van een begin- en eindmeting. Daarom valt hij niet onder de in dit advies gehanteerde definitie van toegevoegde waarde. Deze nieuwe opbrengstindicator wordt vooralsnog alleen vastgesteld op instellingsniveau en niet op opleidingsniveau.

Discussie rondom toepassing van toegevoegde waarde voor betere onderwijskwaliteit

Er is in het onderwijs veel discussie over de betekenis van toegevoegde waarde en over de kansen en risico's van toepassing van het begrip. Voorstanders zien het meten van toegevoegde waarde vooral als een stimulans voor scholen om hun onderwijs te verbeteren. Zij zien toegevoegde waarde dus als een middel om de opbrengsten van een school beter in beeld te brengen en opbrengstgericht werken te bevorderen. De term toegevoegde waarde is ook sterk verbonden met het begrip 'accountability' in het onderwijs. Dat wil zeggen dat indicatoren van toegevoegde waarde kunnen worden gebruikt om de kwaliteit van het onderwijs te beoordelen, of zelfs om de bekostiging te bepalen. Tegenstanders van het meten van toegevoegde waarde wijzen op de risico's van ongewenst strategisch gedrag van scholen (bijvoorbeeld bewust aansturen op lage beginscores), en ook op de onmogelijkheid om alle aspecten van onderwijskwaliteit mee te nemen in de toegevoegde waarde. Deze discussie beperkt zich niet

⁶ Ministerie van Onderwijs, Cultuur en Wetenschap, 2013b.

tot Nederland maar speelt ook in andere landen, waar soms al ervaringen met het gebruik van toegevoegde waarde zijn opgedaan. In de Verenigde Staten is het meten van onderwijsresultaten bijvoorbeeld een belangrijk onderdeel van het omstreden 'No Child Left Behind'-beleid.

Het gebruik van de begrippen leerwinst en toegevoegde waarde in het onderwijs roept verschillende vragen op. Hoe verhouden de opbrengsten voor taal en rekenen zich tot de resultaten in andere vakken? En ook sociale opbrengsten, bijvoorbeeld op het terrein van persoonlijke vorming, sociale competenties en burgerschap, maken deel uit van wat een school bijdraagt aan de ontwikkeling van leerlingen. Is het wenselijk en doenlijk om ook op al die terreinen de leerwinst van leerlingen vast te stellen – en daarmee de toegevoegde waarde van scholen? En als er consensus bestaat over wat leerwinst is, hoe kan de toegevoegde waarde van de school dan precies worden bepaald? Welk deel van de leerwinst kan worden toegeschreven aan de school of leraar, en in welke mate spelen bijvoorbeeld de motivatie en thuissituatie van leerlingen een rol?

1.2 **Adviesvraag: zijn leerwinst en toegevoegde waarde goede instrumenten voor beter onderwijs?**

Tot dusver blijven leerwinst en toegevoegde waarde in het onderwijs nog voor een groot deel buiten beeld bij het gebruik van opbrengstgegevens door en over scholen.

De minister en staatssecretaris van OCW hebben de Onderwijsraad gevraagd te adviseren over het gebruik van leerwinst en toegevoegde waarde voor het verhogen van de kwaliteit van het onderwijs. De raad beantwoordt in dit rapport de volgende vraag (de volledige adviesvraag is te vinden in bijlage 1).

Op welke wijze kunnen gegevens over leerwinst en toegevoegde waarde bijdragen aan de kwaliteit van het onderwijs?

De raad is zich ervan bewust dat de begrippen spanning oproepen tussen 'verbeteren van onderwijs' en 'verantwoorden over het gegeven onderwijs'. Hij besteedt in dit advies dan ook aandacht aan de verschillende doelen die toegevoegde waarde en leerwinst kunnen hebben, aan hoe die doelen al dan niet gecombineerd kunnen worden en aan de gevolgen van toepassing van leerwinst en toegevoegde waarde voor de schoolpraktijk.

Het advies heeft betrekking op het primair onderwijs, het voortgezet onderwijs en het middelbaar beroepsonderwijs. De raad maakt bij de beantwoording van de adviesvraag gebruik van internationaal onderzoek en ervaringen van binnen- en buitenlandse scholen. Ter voorbereiding op dit advies is een overzichtsstudie uitgevoerd naar de internationale literatuur op het gebied van leerwinst en toegevoegde waarde door dr. M. Ehren. Deze studie is te raadplegen via de website van de Onderwijsraad. Daarnaast zijn gesprekken gevoerd met diverse deskundigen. Ook zijn panelbijeenkomsten georganiseerd met leraren, schoolleiders en bestuurders uit de verschillende onderwijssectoren. Een overzicht hiervan is te vinden in de lijst met geraadpleegde deskundigen achter in dit advies.

De raad vindt dat leerwinst en toegevoegde waarde goede instrumenten zijn om scholen meer zicht te geven op hun opbrengsten bij de gemeten vakken en leerdomeinen. Ze kunnen scholen helpen hun onderwijs te verbeteren. De instrumenten zijn echter (net als andere opbrengstindicatoren) níet bruikbaar voor een oordeel over de kwaliteit van de school als geheel.

2 **Advies: leerwinst en toegevoegde waarde zijn wél geschikt voor onderwijsverbetering, níet voor oordeel over de school**

Gegevens over leerwinst zijn een goede aanvulling op andere opbrengstindicatoren: ze geven meer inzicht in de ontwikkeling die leerlingen doormaken bij bepaalde vakken of leerdomeinen. Met dit inzicht kunnen scholen hun onderwijs verbeteren en zijn ze beter in staat maatwerk te leveren. De gegevens over de ontwikkeling van de prestaties maken niet alleen duidelijk of een leerling afwijkt van het gemiddelde, maar ook of zijn resultaten volgens verwachting verlopen – of deze leerling nu onder- of bovengemiddeld presteert gezien zijn eerdere leerprestaties. De raad vindt dat scholen gegevens over leerwinst beter kunnen benutten. Daarvoor is nodig dat zij het belang ervan inzien en de instrumenten zelf goed kunnen hanteren. Ze moeten er, kortom, eigenaar van worden.

Op basis van leerwinst kan een indicator van de toegevoegde waarde van een school worden berekend, om de school een instrument in handen te geven voor reflectie op de eigen opbrengsten (in vergelijking met andere scholen). De raad vindt echter dat de beoordeling van de kwaliteit van de school door de Inspectie hier niet op gebaseerd mag zijn. Daarvoor is een dergelijke indicator te smal en zijn er statistische beperkingen die maken dat de uitkomsten altijd in hun context moeten worden gezien. De raad vindt dat er meerdere indicatoren nodig zijn voor de beoordeling van scholen. Ook ziet hij ruimte voor verbetering van de bestaande opbrengstindicatoren.

Dit hoofdstuk licht toe waarom leerwinst en toegevoegde waarde wel geschikt zijn voor onderwijsverbetering, maar niet gebruikt kunnen worden als grond voor de beoordeling van scholen. In hoofdstuk 3 doet de raad aanbevelingen om scholen te ondersteunen in het gebruik van deze instrumenten. In hoofdstuk 4 doet hij enkele aanbevelingen om andere indicatoren voor de beoordeling van scholen beter te benutten.

Leerwinst en toegevoegde waarde geven een genuanceerd beeld van prestaties

Gegevens over leerwinst bieden kansen om beter te differentiëren

Met behulp van gegevens over leerwinst kunnen scholen inzicht krijgen in de ontwikkeling van al hun leerlingen. Ze krijgen niet alleen een genuanceerd beeld van de gemiddelde leerling, maar ook zicht op zwak en excellent presterende jongeren. Vanuit dat inzicht kunnen scholen hun onderwijsaanbod differentiëren (bijvoorbeeld door de instructie op verschillende niveaus af te stemmen). Zo vergroten ze de leeropbrengsten voor een verscheidenheid aan leerlingen.

Leerwinst kan onderwijsgeevenden helpen in hun dagelijkse praktijk: elke leraar wil graag weten hoe het ervoor staat met zijn leerlingen, hoe ze de afgelopen periode vooruit zijn gegaan en waar mogelijkheden liggen om ze verder te helpen. De raad denkt dat hier de sleutel ligt tot eigenaarschap van schoolleiders en leraren over het instrument leerwinst. Het is een begrijpelijk concept, waarvan de berekening door de school zelf kan worden gedaan, eventueel met behulp van de software die scholen gebruiken voor hun leerlingvolgsysteem. Verder is het voor een school zinvoller om nauwkeurig te kijken naar de ontwikkeling die een kind van jaar op jaar doormaakt dan naar de prestaties aan het einde van de school: in het eerste geval heeft de school nog de mogelijkheid om bij te sturen.

Ervaringen van basisscholen met leerwinst ⁷

Uit de pilots in het basisonderwijs komt naar voren dat 60% van de scholen informatie over leerwinst inderdaad als waardevol ervaart en die ook denkt te gebruiken als het in het leerlingvolgsysteem beschikbaar is. Tegelijkertijd is er de nodige ruimte voor verbetering in de wijze waarop gegevens worden gerapporteerd: twee derde van de basisscholen vindt de specifiek voor de pilot ontwikkelde rapportages over leerwinst inzichtelijk, 52% van de scholen is ook positief over de bruikbaarheid. De helft geeft aan meer opbrengstgericht te werken sinds de start van de pilot. Daar staat tegenover dat respectievelijk 9% en 15% van de scholen de rapportages over leerwinst niet inzichtelijk of bruikbaar vindt. ⁸

Leerwinst omgerekend naar toegevoegde waarde maakt vergelijking van prestaties mogelijk

Met de leerwinst van alle kinderen samen kan de gemiddelde leerwinst op schoolniveau worden berekend (voor de gemeten vakken taal en rekenen). Rekening houdend met de achtergrond van de leerlingen – en de gegevens zorgvuldig interpreterend – kan de school vervolgens zijn toegevoegde waarde vaststellen: wat heeft de instelling bijgedragen aan de prestaties op de gebieden taal en rekenen? Doel hiervan is de leerprestaties in perspectief te plaatsen. De prestaties kunnen namelijk hoger of lager uitvallen dan verwacht zou mogen worden op grond van beginmetingen en achtergrondkenmerken van leerlingen. Dit inzicht kan een prikkel tot verbetering zijn. Via ‘benchmarking’ kunnen scholen met behulp van informatie over leerwinst en toegevoegde waarde beter inzicht krijgen in hun eigen opbrengsten. Door vervolgens te leren van goede voorbeelden, en bijvoorbeeld onderling gerichte feedback te geven, kan het onderwijs verder verbeterd worden.

⁷ De pilots in het basisonderwijs hebben betrekking op 40 basisscholen, waarvan 7 in het speciaal basisonderwijs. In de rest van dit advies wordt verwezen naar de uitkomsten bij de deelnemende reguliere scholen. Gezien het kleine aantal sbo-scholen is het lastiger om over deze scholen een uitspraak te doen. Bovendien kent het sbo een aantal aanvullende complicaties bij het bepalen van leerwinst. In dit licht is het mogelijk ook weinig verrassend dat sbo-scholen de inzichtelijkheid en bruikbaarheid van gegevens over leerwinst en toegevoegde waarde minder positief beoordelen dan reguliere scholen. Uitgebreidere informatie over de sbo-scholen is te vinden in Janssens, Rekers-Mombarg & Lacor, 2014.

⁸ Janssens, Rekers-Mombarg & Lacor, 2014.

Ervaringen in verschillende sectoren met toegevoegde waarde

Uit de pilots blijkt dat toegevoegde waarde voor basisscholen een ingewikkelder concept is dan leerwinst. Niettemin vindt ongeveer de helft van de deelnemende scholen dat de rapportage over toegevoegde waarde inzichtelijk (55%) en bruikbaar (49%) is.⁹ Naast het bestaan van meerdere meetmomenten is een belangrijk verschil met de bestaande opbrengstindicatoren dat toegevoegde waarde rekening houdt met het opleidingsniveau van ouders, etniciteit en het type zorgleerlingen. In het basisonderwijs is het opleidingsniveau van ouders zeer relevant voor de schoolprestaties van leerlingen, maar ook de etnische achtergrond is nog steeds van invloed op met name de taalprestaties.¹⁰ Wanneer deze beide factoren worden meegewogen, ontstaat een genuanceerder beeld van de bijdrage die een school levert aan de leerwinst van de leerlingen bij taal en rekenen.

In het voortgezet onderwijs vinden ook pilots plaats met leerwinst en toegevoegde waarde, maar deze zijn later gestart, zodat daarvan nog geen uitkomsten beschikbaar zijn. Het bepalen van toegevoegde waarde zal zich hier waarschijnlijk ook beperken tot opbrengstgegevens in een klein aantal basisvakken.

In het middelbaar beroepsonderwijs maakt de grote verscheidenheid aan opleidingen zinvolle vergelijkingen lastiger. Bovendien heeft het middelbaar beroepsonderwijs volgens de WEB (Wet educatie en beroepsonderwijs) een drievoudige kwalificatieopdracht: uitoefening van een beroep, deelname aan de maatschappij (burgerschap) en doorstroom naar vervolgonderwijs. Er zijn in het middelbaar beroepsonderwijs nog geen ervaringen met indicatoren voor toegevoegde waarde op basis van leerwinst opgedaan.

Scholen kunnen de informatie beter gebruiken

Uit de internationale literatuur komt naar voren dat scholen lang niet altijd gebruikmaken van de beschikbare opbrengstgegevens.¹¹ Enerzijds worden scholen niet altijd geprikkeld om ze te benutten, anderzijds is het voor veel instellingen lastig om de beschikbare informatie goed te interpreteren.

Gegevens moeten in elk geval aan de volgende kenmerken voldoen om bruikbaar te zijn voor scholen:¹²

- de informatie moet actueel zijn (om het onderwijs aan de desbetreffende groep leerlingen bij te kunnen sturen);
- gegevens moeten op het juiste aggregatieniveau beschikbaar zijn (per leerjaar, klas, afdeling of vak); en
- de informatie moet toepasbaar zijn (waar is bijsturing van het onderwijs gewenst en op welke manier).

Leerwinst en toegevoegde waarde leveren vooral een bijdrage aan kwaliteitsverbetering wanneer het gebruik ervan is ingebed in de onderwijspraktijk. Daarvoor moeten schoolleiders en leraren het als zinvolle maten beschouwen voor feedback, en geoefend zijn in het gebruik ervan. De raad doet in hoofdstuk 3 enkele aanbevelingen om dit te bevorderen.

⁹ Janssens, Rekers-Momberg & Lacor, 2014.

¹⁰ Onderwijsraad, 2013b.

¹¹ Ehren, 2014.

¹² Ehren, 2014.

Toegevoegde waarde is een te smalle indicator voor de kwaliteit van scholen

De raad vindt dat toegevoegde waarde geen goed instrument is voor de beoordeling van de kwaliteit van de school. In de eerste plaats is toegevoegde waarde daarvoor een te smalle indicator.

Minder goed meetbare dimensies van onderwijs blijven buiten beeld

Gebaseerd op leerwinst, zegt toegevoegde waarde alleen iets over die dimensies van het onderwijs die gemeten kunnen worden en waar valide toetsen voor beschikbaar zijn. Taal en rekenen zijn onderwijsdomeinen die zich met toetsen relatief betrouwbaar in beeld laten brengen. De meeste toetsen, waaronder die van het leerlingvolgsysteem, zijn ook op deze onderdelen gericht. De berekening van toegevoegde waarde wordt dus vrijwel uitsluitend gebaseerd op leerresultaten in deze vakgebieden. De raad is van mening dat de ontwikkeling van leerlingen en de bijdrage daaraan van scholen uit méér bestaat dan alleen het aanleren van (Nederlandse) taal- en rekenvaardigheden. Dit is evident in het voortgezet onderwijs en in het middelbaar beroepsonderwijs, maar zelfs in het basisonderwijs nemen taal en rekenen slechts de helft van de onderwijstijd in beslag.¹³ Scholen bieden daarnaast andere vakken aan en beogen een brede ontwikkeling van leerlingen, met aandacht voor bijvoorbeeld studievvaardigheden, de sociaalemotionele ontwikkeling, burgerschapsvorming en culturele ontwikkeling.¹⁴ Er zijn echter nauwelijks instrumenten voorhanden om de toegevoegde waarde op met name de domeinen van sociale en persoonlijke ontwikkeling betrouwbaar te meten.¹⁵ Bovendien is het van belang dat scholen op de verschillende leergebieden ruimte hebben voor een eigen invulling. Uniforme toetsing voor alle domeinen is dus ook ongewenst.

Leren is geen eendimensionaal en lineair proces

Onderwijs is bovendien geen eendimensionaal proces waarbij continu vooruitgang op een specifieke vaardigheid wordt gemeten. Zelfs binnen een vak komen verschillende vaardigheden aan bod (een leerling kan bij rekenen bijvoorbeeld goed zijn in vermenigvuldigen, maar niet in delen). Daarnaast geeft een maat voor leerwinst alleen een positie op een ordinale schaal weer. Niet elke stap vooruit op een vaardigheidsschaal hoeft echter even groot te zijn. Is de vooruitgang van een 4 naar een 5 evenveel waard als de vooruitgang van een 7 naar een 8? Hiervoor zijn externe ijkpunten voorhanden, maar de interpretatie van de gegevens door schoolleiding en leraren is uiterst belangrijk. Daarom is gebruik van leerwinst (en daaraan gekoppeld toegevoegde waarde) wel bruikbaar voor onderwijsverbetering in specifieke leer-domeinen, maar niet voor de beoordeling van de school. Bijlage 2 beschrijft hoe het meten van leerwinst en toegevoegde waarde in zijn werk gaat en gaat nader in op de mogelijke gevolgen van deze inherente beperking. In bijlage 3 wordt een gedetailleerder overzicht gegeven van veelgebruikte modellen om schoolopbrengsten te berekenen.

Hoe belangrijk de indicator toegevoegde waarde is voor de kwaliteit van het onderwijs, hangt bovendien af van het perspectief van waaruit onderwijs wordt beschouwd. Ook in dit opzicht is er sprake van meerdere dimensies. Uit onderzoek komt bijvoorbeeld naar voren dat in het middelbaar beroepsonderwijs deelnemers en beleidsmakers veel waarde hechten aan het slagsingspercentage als opbrengstindicator, terwijl leraren en werkplekbegeleiders dat een veel minder belangrijke factor vinden bij het beoordelen van de kwaliteit van een opleiding. Zij

¹³ Inspectie van het Onderwijs, 2013b.

¹⁴ Onderwijsraad, 2011b.

¹⁵ Dijkstra, 2012.

hechten meer aan de structuur van de opleiding en aan de mate waarin studenten worden uitgedaagd.¹⁶ Daarnaast wordt de tevredenheid van deelnemers over de opleiding door instellingen ook als belangrijk indicator van onderwijskwaliteit ervaren, terwijl dit aspect niet als zodanig door de Inspectie wordt meegenomen in de beoordeling.¹⁷ De overheid zal geïnteresseerd zijn in de totale toegevoegde waarde op stelselniveau. Verder zullen ouders vooral geïnteresseerd zijn in de optimale ontwikkeling van hun eigen kind, en minder in de maximale gemiddelde ontwikkeling van alle kinderen. Verschillende ouders blijken bovendien belang te hechten aan verschillende zaken: publicatie van schoolgegevens op een diversiteit aan terreinen biedt ouders de mogelijkheid om goed geïnformeerd een afweging te maken. In dat licht is het onwenselijk dat de overheid de opbrengsten van scholen reduceert tot één maat van toegevoegde waarde voor toezicht of bekostiging.

Begin- en eindmeting worden uitgevoerd bij verschillende groepen leerlingen

Niet elke leerling die de school binnenstroomt, maakt deze school ook af. Leerlingen verhuizen, stoppen met hun opleiding of wisselen om andere redenen tussentijds van school. Met name in bepaalde wijken van de grote steden kunnen de percentages leerlingen die de school tussentijds verlaten of binnenkomen aanzienlijk zijn. In Amsterdam wisselt 35% van de leerlingen die in groep 1 beginnen gedurende de basisschoolperiode van school.¹⁸ Voor het berekenen van leerwinst en daarmee toegevoegde waarde, is dat een lastig probleem. Moet de toegevoegde waarde alleen berekend worden over de leerlingen die de hele tijd op dezelfde school zijn gebleven? Dit geeft dan geen informatie over hoe goed een school het doet voor leerlingen die tussentijds uit- of instromen.

2.3 Vaststellen van toegevoegde waarde vraagt om inhoudelijke keuzes

Nog een reden om toegevoegde waarde niet te gebruiken voor de beoordeling van scholen, is dat niet-objectieve factoren een rol spelen bij de vaststelling. Om de bijdrage van een school te kunnen bepalen zijn inhoudelijke keuzes nodig, bijvoorbeeld over de achtergrondkenmerken waarvoor gecorrigeerd wordt. Moet in de berekening van toegevoegde waarde rekening worden gehouden met gemiddeld lagere prestaties van jongens of werkt dit lage verwachtingen van jongens juist in de hand? Moet rekening gehouden worden met de ruimere financiële middelen die scholen met meer achterstandsleerlingen krijgen? Er bestaat op deze vragen geen 'juist' antwoord: in feite kan voor elke school een aantal toegevoegde waardes worden berekend, waarbij de discussie blijft welke daarvan de beste indicator is. Ook bij leerwinst, dat de basis vormt voor toegevoegde waarde, geldt dat verschillende factoren een rol spelen. Leerwinst is niet alleen afhankelijk van de inspanningen van de school, maar bijvoorbeeld ook van familieomstandigheden en de kwaliteit van (eventuele) huiswerkbegeleiding.¹⁹ Kinderen van hoogopgeleide ouders hebben meer hulpbronnen tot hun beschikking en weten daar ook meer profijt uit te halen.²⁰ Als over de gehele linie scholen minder energie steken in leerlingen met een bepaalde achtergrond, dan zal dit in de gangbare methodes voor het bepalen van leerwinst of toegevoegde waarde niet worden toegeschreven aan de school. Alleen als individuele scholen bij bepaalde groepen afwijken van de gemiddelde resultaten van andere scholen, zal dit tot uiting komen in een indicator van toegevoegde waarde.

¹⁶ Van der Sluis, Reezigt & Borghans, 2013.

¹⁷ Inspectie van het Onderwijs, 2012b.

¹⁸ Gemeente Amsterdam DMO, 2010.

¹⁹ Onderwijsraad, 2011b.

²⁰ Janssens, Rekers-Mombarg & Lacor, 2014.

Pilotscholen niet eensgezind over bruikbare achtergrondkenmerken

Bij de pilots met leerwinst en toegevoegde waarde in het basisonderwijs is rekening gehouden met een, in eerste instantie lange, lijst met 'fairnesskenmerken'. De lijst is vervolgens ingekort. Een aantal schoolkenmerken (zoals denominatie) werd geschrapt omdat ze niet statistisch relevant bleken. Verder gaven scholen aan dat zij geen behoefte hadden aan een correctie voor het geslacht van de leerling. Uiteindelijk is voor het bepalen van de toegevoegde waarde van de school gekeken naar het hoogste opleidingsniveau van de ouder, de etniciteit van de leerling en het type zorgleerling. De helft van de betrokken scholen (47%) vindt vervolgens dat dit een eerlijke vergelijking oplevert van de toegevoegde waarde ten opzichte van andere scholen. Opvallend genoeg is 28% van de scholen het daar niet mee eens. Een aantal scholen pleit bijvoorbeeld voor het betrekken van extra fairnesskenmerken.²¹ Het rapport geeft niet aan welke dat zijn.

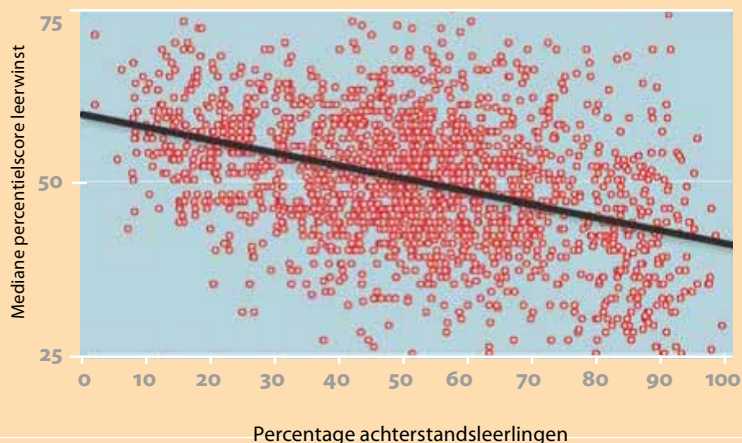
Totstandkoming van indicatoren van toegevoegde waarde is niet transparant

Voor veel basisscholen is de totstandkoming van de indicatoren voor het beoordelen van hun opbrengsten onvoldoende transparant.²² De keuze voor een bepaalde correctie op de leerwinst kan belangrijke gevolgen hebben voor de uitkomst van de toegevoegde waarde.

Onderstaande grafieken, gebaseerd op gegevens van scholen in de Amerikaanse staat Missouri, laten zien dat de correctie op achtergrondkenmerken verschillend kan uitpakken. De grafieken geven schattingen van leerwinst en toegevoegde waarde voor het vak wiskunde. De schattingen verschillen afhankelijk van het aantal correcties voor achtergrondkenmerken van leerlingen dat erin wordt meegenomen.

Figuur 1 laat zien dat de middelste waarde van de gemeten leerwinst van alle leerlingen (de mediaan) sterk varieert tussen al deze scholen, maar dat deze gemiddeld lager is op scholen met veel achterstandsléerlingen.

Figuur 1: Vergelijking van leerwinst (bij rekenen) met het percentage achterstandsléerlingen op schoolniveau



Bron: Ehlert, Koedel, Parsons & Podgursky, 2014.

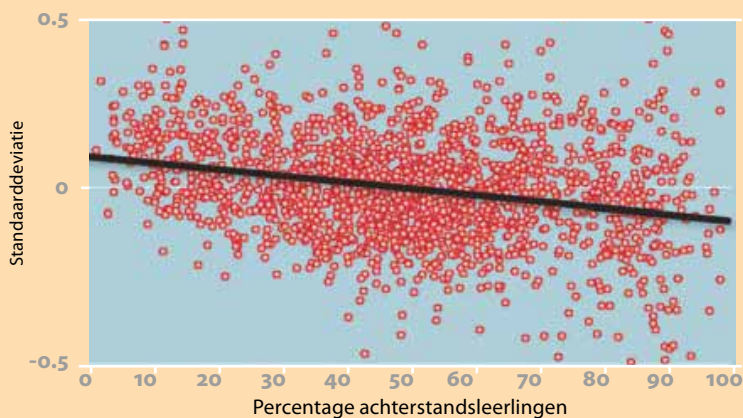
Toelichting: In deze grafiek wordt het percentage leerlingen dat in aanmerking komt voor een gratis schoolmaaltijd (of een korting daarop) gebruikt als maat voor het aandeel achterstandsléerlingen op een school. De mediane percentielscore leerwinst (in rekenen) op de verticale as is een maat voor de leerwinst van scholen.

21 Janssens, Rekers-Mombarg & Lacor, 2014.

22 Ministerie van Onderwijs, Cultuur en Wetenschap, 2013a.

Figuur 2 geeft schattingen van de toegevoegde waarde wanneer gecorrigeerd wordt voor leerlingkenmerken. Dan blijken de verschillen tussen verschillend samengestelde scholen klete zijn. Scholen met veel achterstandsleerlingen scoren echter gemiddeld nog steeds lager, bijvoorbeeld doordat zij mogelijk meer moeite hebben met het aantrekken van goede leraren.

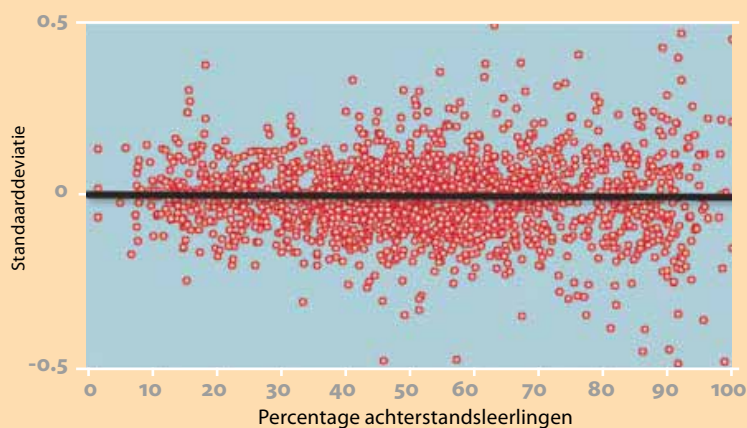
Figuur 2. Vergelijking van een indicator van toegevoegde waarde, waarbij leerwinst (bij rekenen) is gecorrigeerd voor leerlingkenmerken, met het percentage achterstandsleerlingen op schoolniveau



Bron: Ehlert, Koedel, Parsons & Podgursky, 2014.

Toelichting: Ook hier wordt het percentage leerlingen dat in aanmerking komt voor een gratis schoolmaaltijd (of een korting daarop) gebruikt als maat voor het aandeel achterstandsleerlingen op een school. De standaarddeviatie geeft een indicator van de toegevoegde waarde waarbij leerwinst is gecorrigeerd voor leerlingkenmerken.

Figuur 3. Vergelijking van een indicator van toegevoegde waarde waarbij leerwinst (bij rekenen) is gecorrigeerd voor verschillen in samenstelling tussen scholen, met het percentage achterstandsleerlingen op schoolniveau



Bron: Ehlert, Koedel, Parsons & Podgursky, 2014.

Toelichting: Ook hier wordt het percentage leerlingen dat in aanmerking komt voor een gratis schoolmaaltijd (of een korting daarop) gebruikt als maat voor het aandeel achterstandsleerlingen op een school. De standaarddeviatie geeft een indicator van de toegevoegde waarde waarbij leerwinst is gecorrigeerd voor verschillen in samenstelling tussen scholen.

In figuur 3 wordt volledig gecorrigeerd voor de verschillen in samenstelling tussen scholen. De gemiddelde toegevoegde waarde van scholen met eenzelfde percentage achterstandsleerlingen is op nul gesteld. De resterende indicator geeft aan of de toegevoegde waarde van een school hoger of lager is dan een gemiddelde school met eenzelfde percentage achterstandsleerlingen.²³

Keuze voor een leerlingenpopulatie

Daar komt bij dat toegevoegde waarde afhankelijk is van beleidskeuzes die een school maakt. Op welke leerlingenpopulatie richt men zich? Scholen met dezelfde gemiddelde toegevoegde waarde kunnen verschillen in de toegevoegde waarde die zij voor individuele leerlingen hebben. Een school kan bijvoorbeeld vooral goede resultaten boeken bij initieel laag presterende leerlingen, terwijl een andere school juist goed is in het verbeteren van de opbrengsten van hoog presterende leerlingen. Een complicerende factor daarbij is dat de leerwinst van individuele leerlingen kan verschillen naar gelang de samenstelling van de leerlingenpopulatie. Leerlingen kunnen zich optrekken aan vergelijkbare of betere leerlingen in hun klas, maar kunnen ook worden afgeleid door klasgenoten. De optimale klassensamenstelling voor individuele leerlingen hoeft niet samen te vallen met die op school- of stelselniveau. Dat roept ook de vraag op in hoeverre de leerwinst die leerlingen boeken een verdienste is van de school of van bijvoorbeeld de aanwezigheid van goede medeleerlingen. De school kan daar overigens wel een sturende rol in spelen, bijvoorbeeld door klassen zo samen te stellen dat de meeste synergie ontstaat. Daar staat tegenover dat scholen de maatschappelijke opdracht hebben om bijvoorbeeld leerlingen met problemen op te vangen – ook als dat ten koste gaat van de leerwinst van andere leerlingen.²⁴ Uiteindelijk blijft het daarom een inhoudelijke afweging wat wel en wat niet een onderdeel van de toegevoegde waarde van de school is.

Hoog niveau in de klas en goede leraren spelen ook een rol

Onderzoek onder vo-scholen in Engeland laat zien dat er een samenhang bestaat tussen schoolcompositie en schoolprestaties.²⁵ Een mogelijke verklaring voor deze samenhang is dat leerlingen meer leren wanneer zij bij betere leerlingen in de klas zitten ('peer-effects').²⁶ Een tweede verklaring is dat scholen met betere leerlingen gemiddeld vaak ook betere leraren hebben. Dat wil niet zeggen dat er geen goede leraren in achterstandsscholen zijn, maar wel dat deze scholen meer moeite hebben om goede leraren aan te trekken en vast te houden.²⁷

Ook uit onderzoek naar de mobiliteit van leerkrachten in het Nederlandse basisonderwijs komt naar voren dat zij een voorkeur hebben voor scholen met hoge Cito-scores, kleine klassen en weinig (allichtone) achterstandsleerlingen. Hoewel scholen met veel achterstandsleerlingen kleinere klassen hebben, blijken deze nog onvoldoende klein te zijn om aantrekkelijk te worden.²⁸ Ook het feit dat scholen met veel gewichtenleerlingen meer moeite hebben om leraren aan te trekken, betekent dat de kwaliteit op deze scholen vaker kwetsbaar is. Uit onderzoek van de Inspectie van het Onderwijs (2009) naar 'sterke basisscholen' komt bijvoorbeeld naar voren dat dit minder vaak scholen met veel gewichtenleerlingen zijn.

23 Ehlert, Koedel, Parsons & Podgursky, 2014.

24 Carrell & Hoekstra, 2010.

25 Allen & Burgess, 2013.

26 Jackson, 2013.

27 Clotfelter, Ladd & Vigdor, 2006.

28 Bonhomme, Jolivet & Leuven, 2012.

Keuze voor type toetsen, doel ervan en moment van afname

Tot slot zijn doel en kwaliteit van de gekozen toetsen bij het meten van leerwinst van belang. Toetsscores worden in veel gevallen ook beïnvloed door andere factoren dan de beheersing van de leerstof en vaardigheden.²⁹ Dit kan te maken hebben met de vragen die toevallig in een bepaalde toets aan de orde komen. Er kunnen bijvoorbeeld vertekeningen ontstaan doordat toetsvragen qua onderwerp of niveau meer of minder aansluiten bij bepaalde leerlingen. Een toets kan bewust ontworpen zijn voor het nemen van een zak-slaagregeling, maar daardoor minder betrouwbaar zijn bij de beoordeling van een leerling die ver boven (of onder) de zak-slaaggrens presteert. Ook het moment van afname of de omstandigheden kunnen toetsresultaten sterk beïnvloeden. Dit geldt niet alleen bij toetsen waarbij de druk hoog is omdat er veel van afhangt, maar ook bij toetsen waar juist niets op het spel staat: sommige leerlingen zijn meer, andere minder intrinsiek gemotiveerd om een toets goed te maken.

2.4 Er zijn statistische problemen bij het meten van leerwinst en toegevoegde waarde

Het meten van leerwinst en toegevoegde waarde kent verder een aantal statistische problemen.³⁰

Meetproblemen bij het bepalen van leerwinst

Het meten van leerwinst veronderstelt dat op een school een begin- en eindmeting plaatsvindt, met een aantal jaren ertussen. De vergelijkbaarheid van toetsen voor jongere en toetsen voor oudere leerlingen (bijvoorbeeld op een basisschool) is echter niet eenvoudig. Er is betrouwbare toetsing mogelijk vanaf ongeveer zes jaar, maar het verschil in taalvaardigheid tussen bijvoorbeeld leerlingen van zes en van twaalf jaar maakt een valide vergelijking tussen deze twee meetmomenten moeilijk. Een leerling kan beter presteren omdat zijn kennis op het gemeten vakgebied is toegenomen, maar ook omdat hij bij de eindmeting gewoon beter kan lezen en de vragen beter begrijpt. Kennis in een bepaald vak kan sterk samenhangen met de kennis van andere vakken. Dit is geen onoverkomelijk probleem voor diagnostisch gebruik van toetsgegevens binnen de school in het kader van een leerlingvolgsysteem (inclusief een vergelijking met de toetsscores van andere scholen ten behoeve van de eigen kwaliteitszorg). Schoolleiders en leraren kunnen de gegevens interpreteren. Het is wél een probleem voor het gebruik van de toetsgegevens voor het nauwkeurig en eenduidig bepalen van de leerwinst: hiervoor is de vergelijkbaarheid van de resultaten van de begintoets met die van de eindtoets noodzakelijk.³¹ Daarnaast kunnen ook verschillen in meetmomenten van invloed zijn. Gebeurt dit in het ene schooljaar op een ander moment dan in het andere, dan heeft de gemeten leerwinst betrekking op een andere periode en een ander tijdpad.

Er is bovendien sprake van vele verschillende vormen van leerwinst, waarvan een aantal mogelijk niet goed in beeld wordt gebracht. Oorzaak hiervan zijn de grote verschillen in lesstof tussen onderwijssectoren. Waar in het primair onderwijs taal en rekenen belangrijke – maar niet de enige – vakken zijn, kennen het voortgezet onderwijs en het middelbaar beroepsonderwijs een grotere variatie aan vakken met grote verschillen in inhoud en looptijd. Verder heeft het middelbaar beroepsonderwijs expliciet een dubbele doelstelling: leerlingen opleiden voor

²⁹ Ehren, 2014.

³⁰ Voor een overzicht zie: Ehren, 2014.

³¹ Onderwijsraad, 2011b.

arbeidsmarkt én hoger beroepsonderwijs, terwijl het basis- en voortgezet onderwijs leerlingen primair voorbereiden op vervolgonderwijs.

Meetproblemen bij het bepalen van toegevoegde waarde

Bovengenoemde problemen bij het meten van leerwinst hebben uiteraard ook gevolgen voor de (daarvan afgeleide) toegevoegde waarde. Het bepalen van toegevoegde waarde kent daarbovenop enkele specifieke moeilijkheden.

Ten eerste is de toegevoegde waarde van een school geen eenduidig begrip. De toegevoegde waarde kan op veel manieren worden berekend, waarbij de uitkomsten van de verschillende modellen verschillende resultaten opleveren. Leerlingprestaties kunnen bijvoorbeeld worden vergeleken met een absolute norm, maar ook met een relatieve. Bijlage 3 geeft een overzicht van een aantal veelgebruikte modellen voor opbrengstindicatoren. Onderzoek laat zien dat de uitkomsten van de modellen weliswaar samenhangen, maar toch tot sterk verschillende classificaties van individuele scholen kunnen leiden (ineffectief, gemiddeld, effectief). Dit geldt in het bijzonder voor scholen rondom het gemiddelde. Daar vinden veel verschuivingen in rangorde plaats afhankelijk van de opbrengstmaat die wordt gebruikt. Alleen de classificatie van de uitschieters naar boven en naar beneden is redelijk stabiel.³²

Een ander probleem is dat indicatoren voor toegevoegde waarde (net als de meeste andere prestatie-indicatoren) slechts een beperkte groep scholen in beeld brengen die daadwerkelijk significant hoger of lager bijdragen aan de leerprestaties van hun leerlingen dan gemiddeld.³³ Dat komt door de doorgaans kleine leerlingenaantallen die in metingen van toegevoegde waarde worden opgenomen. De meeste scholen hebben minder dan honderd leerlingen per jaarlaag. Dit is nadelig voor de stabiliteit van de gegevens en maakt de uitkomsten gevoeliger voor toevallige factoren, zoals een cohort goed scorende leerlingen, maar ook een griep epidemie of storende factoren tijdens de afname van een toets. Kleine scholen hebben daarom statistisch gezien de grootste kans om tot de groep scholen met de kleinste of grootste toegevoegde waarde te behoren, omdat toevallige factoren op deze scholen een relatief grote impact hebben. Alleen de waarden van de grotere scholen zijn dus betrouwbaar.³⁴ Voor kleinere scholen moet een aantal jaren worden samengevoegd om een vergelijkbare betrouwbaarheid te realiseren.

Op het moment dat de toegevoegde waarde wordt bepaald op het niveau van individuele leraren worden de leerlingenaantallen nog kleiner en daarmee de statistische nauwkeurigheid geringer.³⁵ Alleen daarom al acht de raad het vaststellen van de waarde op dit niveau onwenselijk voor beoordelingsdoeleinden. De toegevoegde waarde van een leraar is verder niet alleen zijn individuele verdienste, maar mede afhankelijk van de inzet en kwaliteit van zijn collega's. Indicatoren voor toegevoegde waarde kunnen hooguit een bescheiden rol spelen in evaluatiesystemen van docenten en als feedback op het eigen functioneren, mits de informatie zorgvuldig geïnterpreteerd wordt.³⁶

³² Ladd & Walsh, 2002; Timmermans, 2012.

³³ Timmermans, 2012.

³⁴ Kane & Staiger, 2002.

³⁵ Sass, Semykina & Harris, 2013.

³⁶ Haertel, 2013.

Het gebruik van toegevoegde waarde als beoordelingsinstrument kan ongewenst strategisch gedrag uitlokken

Tot slot kan het gebruik van toegevoegde waarde als indicator voor schoolkwaliteit ongewenst strategisch gedrag uitlokken van scholen.³⁷

Het onderwijsaanbod kan verschromen

Toegevoegde waarde laat, zoals in paragraaf 2.2 is betoogd, slechts een deel van de onderwijskwaliteit zien. Zou deze waarde tóch worden ingezet om de kwaliteit van scholen te beoordelen, dan kan dit ertoe leiden dat scholen vakken en vaardigheden die niet worden gemeten, veronachtzamen. Er gaat bijvoorbeeld minder aandacht en geld naar onderdelen die in de toetsen niet aan bod komen. Te verwachten valt dat bijvoorbeeld in het basisonderwijs het accent eenzijdig komt te liggen op taal en rekenen. Dit fenomeen is in de literatuur al veelvuldig gerapporteerd.³⁸ Het kan leiden tot een verschroming van het aanbod.

'Teaching to the test' wordt aantrekkelijk

Ook binnen de getoetste vakken als taal en rekenen kan een verschuiving optreden richting makkelijke toetsbare onderdelen. Ervaringen in verschillende Amerikaanse staten bevestigen dat dit gebeurt. Daar waren als gevolg van 'incentives' voor kwaliteitsverbetering de resultaten van 'high-stakes tests' beter dan die van 'low-stakes tests' – hoewel ze vergelijkbare vaardigheden beoogden te meten.³⁹

Dit fenomeen wordt 'teaching to the test' genoemd. Als een test alle beoogde aspecten van het onderwijs omvat is dit geen probleem. Maar in veel gevallen zijn niet alle onderdelen van een vak even goed toetsbaar en kent een toets de nodige beperkingen. Dit is met name een risico bij toetsen met een grote voorspelbaarheid en beperkte validiteit en betrouwbaarheid. Bij teaching to the test worden onderwijs en instructie versmald naar de vakken, onderwerpen en vaardigheden die worden getoetst, en naar de item-formats waarin dat gebeurt. Het perkt het leren van de leerling in. Bovendien geven de scores op de toets een minder nauwkeurig beeld van de daadwerkelijke vaardigheid die is geleerd in het hele domein.⁴⁰

Teaching-to-the-test is in het bijzonder een risico wanneer toetsresultaten een belangrijke rol gaan spelen bij de beoordeling en bekostiging van scholen. De vertekening kan ten koste gaan van de waarde van de scores voor interne (diagnostische) doeleinden. Als het Cito-leerlingvolgsysteem en de bijbehorende toetsen belangrijker worden voor de beoordeling van scholen, dan krijgen scholen een incentive om de resultaten te beïnvloeden. Ze hebben daar ook ruime mogelijkheden voor: ze kunnen de toetsen afnemen op momenten die ze zelf bepalen. Om dit te voorkomen zouden scholen elk jaar op een vast moment een (kostbare) nieuwe Cito-toets moeten afnemen, wat scholen tegelijkertijd een nog sterkere incentive geeft tot teaching to the test.

³⁷ Ehren, 2014.

³⁸ Voor een overzicht zie: Ehren, 2014.

³⁹ Figlio & Loeb, 2011.

⁴⁰ Ehren, 2014.

Aansturen op lage scores op begintoetsen

Het gebruik van toetsscores bij het bepalen van de kwaliteit van een school kan altijd leiden tot strategisch gedrag van scholen. Toegevoegde waarde als indicator is daar echter extra gevoelig voor vanwege de mogelijkheid om leerlingen op de begintoets laag te laten scoren. Op die manier wordt de gemeten toegevoegde waarde kunstmatig verhoogd. Ook kan de instructie selectief worden gericht op leerlingen waarmee de meeste leerwinst valt te behalen. Amerikaans onderzoek laat zien dat sommige leerlingen door deze strategie (ook op de lange termijn) veel profijt hebben van aangescherpte accountability-maatregelen, terwijl anderen er juist nadeel van ondervinden.⁴¹

In de praktijk blijkt met name de betrouwbaarheid van de eerste beginmeting problematisch. Scholen hebben meer mogelijkheden om deze scores laag te houden dan om hogere scores te bewerkstelligen op de eindmeting. Ze kunnen bijvoorbeeld leerlingen in de aanloop naar een toets verwarren door andere stof aan te bieden. Of ze kunnen de toets plannen op een moment dat leerlingen minder goed voorbereid zijn. Het is ook mogelijk leerlingen minder tijd te geven om de toets te maken, of om ze te zeggen dat de toets niet zo belangrijk is. Ervaringen in bijvoorbeeld Amerika – waar door het No Child Left Behind-beleid in veel staten sterke prikkels voor scholen zijn ontstaan – laten zien dat deze voorbeelden niet denkbeeldig zijn, maar dat ze in de praktijk daadwerkelijk optreden.⁴²

Zoals gezegd treedt ongewenst strategisch gedrag ook op bij andere indicatoren voor beoordeling van scholen.⁴³ De literatuur over het gebruik van incentives, ook in andere semi-publieke sectoren, wijst daarom op het belang van zorgvuldigheid bij het inzetten ervan. Sterke extrinsieke prikkels kunnen bijvoorbeeld ten koste gaan van de intrinsieke motivatie. Incentives kunnen zeer effectief zijn, maar dus ook ongewenste neveneffecten hebben die de beoogde effecten (deels) teniet doen.⁴⁴

⁴¹ Deming, Cohodes, Jennings & Jencks, 2013.

⁴² Ehren, 2014.

⁴³ Neal & Schanzenbach, 2010.

⁴⁴ Gneezy, Meier & Rey-Biel, 2011.

Scholen kunnen leerwinst en toegevoegde waarde inzetten voor onderwijsverbetering. Als de instrumenten daarvoor worden gebruikt, hebben scholen de ruimte om te leren van de gegevens en ze zich eigen te maken. Daarvoor zijn twee dingen nodig: ondersteuning bij het ontwikkelen van expertise binnen scholen én beschikbaarheid van de gegevens voor onderzoek.

3 **Aanbeveling 1: geef de school eigenaarschap over de instrumenten leerwinst en toegevoegde waarde**

Scholen kunnen beter gebruikmaken van de beschikbare opbrengstgegevens voor kwaliteitsverbetering. Leerwinst en toegevoegde waarde zijn in de ogen van de raad geen instrumenten om schoolleiders en leraren op af te rekenen, maar om ze te ondersteunen in hun dagelijkse praktijk. Mits goed gebruikt geven de waarden informatie over hoe het ervoor staat met de leerlingen en waar mogelijkheden liggen om hun prestaties te verbeteren. Als schoolleiders en leraren zich de instrumenten eigen kunnen maken, dan kunnen ze die inzetten ten behoeve van de ontwikkeling van de leerlingen, maar ook om hun eigen werk aantrekkelijker te maken. Ze moeten hiervoor wel ruimte en ondersteuning krijgen.

3.1 **Vergroot de expertise op scholen over opbrengsten en onderzoek**

Het meten van leerwinst en toegevoegde waarde en het gebruik van deze gegevens krijgt binnen een schoolcontext pas betekenis wanneer er ook een gesloten kwaliteitscyclus bestaat. Als opbrengstgericht werken als onderdeel van kwaliteitsverbetering nog geen gemeengoed is, dan zijn leerwinst en toegevoegde waarde niet meer dan losse instrumenten. Het risico bestaat dan dat het verzamelen van opbrengstgegevens een doel op zich wordt en daarmee betekenisloos is. Daarom is het noodzakelijk dat scholen zich eerst richten op een werkwijze voor een juiste inbedding van dergelijke gegevens.

Zorg dat kennis over statistiek en praktijkgericht onderzoek beschikbaar is

Een veelgenoemde reden voor het niet optimaal gebruikmaken van opbrengstgegevens is dat scholen datamanagement lastig vinden en onvoldoende expertise erover in huis hebben. Het op waarde schatten van de beschikbare gegevens vereist bijvoorbeeld dat schoolleiders en bestuurders over een zekere mate van statistische kennis beschikken.⁴⁵ Het belang van 'data-wijsheid' geldt ook voor leraren. Om opbrengstgericht te kunnen werken is het nodig dat leerkrachten de leerwinstgegevens van hun leerlingen begrijpen. Door hierover ook met elkaar in gesprek te gaan, kan de doorgaande leerlijn verbeterd worden.

⁴⁵ Lang, 2010.

Ook voor de volgende stappen in de cyclus van het opbrengstgericht werken is begrip van opbrengstgegevens belangrijk: het vertalen van de gegevens naar concrete doelen voor (groepen) leerlingen en vervolgens het vertalen van deze doelen naar maatregelen en concreet handelen in de klas – vooral het aanbieden van een gedifferentieerde aanpak.⁴⁶ Om deze vaardigheden te ontwikkelen is een verdere professionalisering van schoolleiders en leraren op deze gebieden essentieel. In een eerder advies deed de raad suggesties om de onderzoeks- en ontwikkelingsvaardigheden van leraren als beroepsgroep te stimuleren. Een daarvan was het bieden van tijd en ruimte aan een aantal leraren binnen de school om zich te specialiseren in praktijkgericht onderzoek. Samenwerkingsverbanden tussen scholen en universiteiten of hogescholen kunnen hierbij helpen.⁴⁷

Benut datateams op scholen: ervaringen zijn positief

Om scholen bij bovengenoemde knelpunten te begeleiden, vinden experimenten met zogenoemde datateams plaats. Binnen deze datateams werken schoolleiders en een aantal docenten onder begeleiding van een wetenschappelijk onderzoeker aan oplossingen voor concrete vragen – met behulp van beschikbare data. Hierbij wordt gebruikgemaakt van de evaluatieve cyclus waarbinnen hypothesen opgesteld en getoetst worden. Op deze manier investeren de leden van het datateam in hun eigen professionalisering, terwijl zij tegelijkertijd aan onderwijsverbetering werken. De internationale onderzoeksliteratuur naar de effectiviteit van de datateammethode rapporteert positieve uitkomsten.⁴⁸ Ook de voorlopige nationale resultaten naar het functioneren van deze teams geven een positief beeld.⁴⁹ De raad pleit er dan ook voor om een dergelijke methode op grotere schaal beschikbaar te stellen, zodat alle onderwijsinstellingen de mogelijkheid wordt geboden om onder begeleiding doelgericht te werken aan het vergroten van de vaardigheden op het gebied van opbrengstgericht werken.

Ook een cultuurverandering nodig

Scholing van schoolleiders en docenten alléén is niet genoeg om opbrengstgegevens in te kunnen zetten voor onderwijsverbetering. Ook veranderingen in gedrag en houding van onderwijsgeevenden zijn nodig. Schoolleiders en leerkrachten beseffen steeds meer dat zij zelf op onderzoek uit zullen moeten gaan om alle beschikbare informatie te benutten en dat er ruimte moet zijn voor reflectie op het eigen functioneren en voor het geven en ontvangen van feedback. De school wordt daarmee een sterkere lerende en professionele organisatie. In de praktijk is dit niet zo makkelijk te realiseren. Er is niet altijd voldoende vermogen om te veranderen of te vernieuwen, en initiatieven daarvoor zijn bijvoorbeeld van tijdelijke aard of niet verankerd in de organisatie. Schoolleiders hebben een taak in het bespreekbaar maken van resultaten in het team en in het aangaan van de discussie over onderwijsverbetering. Zij kunnen er tevens voor zorgen dat professionalisering op dit vlak tot stand komt en verankerd wordt in het team en de verschillende (vak)secties.

De raad is daarnaast van mening dat de (verdere) ontwikkeling naar een lerende organisatie niet alleen moet plaatsvinden binnen de schoolmuren, maar gevoed moet worden door de uitwisseling van kennis en ervaring. Te denken valt aan een versterking van de contacten met scholen binnen een schoolbestuur of aan regionale samenwerkingsverbanden op het gebied van opbrengstgericht werken (bijvoorbeeld intervisiegroepen). Daarbij zou het geven van feedback op elkaars aanpak en resultaten een vast onderdeel van gesprek moeten worden.

46 De meeste leraren beschikken bijvoorbeeld wel over basisvaardigheden maar veel leraren hebben moeite met complexere vaardigheden zoals het afstemmen van het onderwijs op verschillen tussen leerlingen; Inspectie van het Onderwijs, 2013a.

47 Onderwijsraad, 2011a.

48 Boudett, Cidy & Murnane, 2007; Mandinach & Honey, 2008.

49 Schildkamp, 2011; Schildkamp, 2012

Gegevens over leerwinst en toegevoegde waarde leveren daarvoor waardevolle input. Het zou goed zijn om ook direct belanghebbenden zoals ouders (bij de basisschool en de middelbare school) of bedrijven (bij het middelbaar beroepsonderwijs) te betrekken bij het bespreken van de opbrengsten van het onderwijs.

3.2 Stel gegevens beschikbaar voor onderzoek

Onderzoekers kunnen, samen met scholen, bijdragen aan beter onderwijs. Gezamenlijk kunnen ze instrumenten als leerwinst en toegevoegde waarde verder ontwikkelen.

Onderzoek draagt bij aan het denken over en het ontwikkelen van leerwinst en toegevoegde waarde
Juist omdat er niet één manier is om leerwinst en toegevoegde waarde te berekenen en omdat er nog methodologische problemen spelen, is inbreng van onderzoekers belangrijk in het denken over (en ontwikkelen van) leerwinst en toegevoegde waarde. Om de potentie van de instrumenten goed te benutten is het ook van belang om niet één vaste definitie als uitgangspunt te nemen, maar om ruimte te houden voor ontwikkeling van het concept en het gebruik ervan. In een open wetenschappelijke discussie kunnen veronderstellingen die ten grondslag liggen aan gehanteerde maten voor toegevoegde waarde worden getoetst en kan het gebruik worden onderzocht. Op deze wijze kan ook de verdere ontwikkeling van gebruiksmogelijkheden worden gestimuleerd en kunnen nieuwe toepassingen ontstaan, die de ontwikkeling van het onderwijs bevorderen. Het hoeft hierbij niet alleen om schoolprestaties te gaan, maar kan bijvoorbeeld ook betrekking hebben op analyses van onderwijsmethodes of van het functioneren van het stelsel als geheel. Het benutten van de kracht van een dergelijke open discussie en open innovatie vergt dat onderzoekers over transparante informatie kunnen beschikken.

Scholen en leraren moeten zelf betrokken zijn bij dergelijk onderzoek. Dit zal overigens ook het draagvlak voor het gebruik van leerwinst en toegevoegde waarde onder hen vergroten. De raad heeft eerder al in een ander verband gepleit voor de vorming van uoc's (universitaire onderwijscentra). Hierin zoeken scholen en onderzoekers samen naar antwoorden op vragen over onderwijs die op schoolniveau leven. Deze samenwerking bevordert de professionalisering van scholen en leraren met betrekking tot het gebruik van opbrengstgegevens voor kwaliteitsverbetering en de totstandkoming van een onderzoekende cultuur in de school.⁵⁰ En die samenwerking bevordert tevens de bereidheid van onderzoekers om aan te sluiten bij praktijkvragen en hun resultaten zodanig te presenteren dat ze bijdragen aan onderwijsverbetering.

⁵⁰ Onderwijsraad, 2011a.

Indicatoren van toegevoegde waarde zullen (net als andere opbrengstindicatoren) altijd beperkingen houden. Belangrijke aspecten van onderwijskwaliteit zijn er niet in opgenomen. Scholen doen er goed aan om ook andere factoren te betrekken bij het vaststellen van de kwaliteit van hun onderwijs. Ook in het toezicht moet ruimte zijn voor een brede kijk op onderwijs én aandacht voor mogelijke tekortkomingen van indicatoren. Tot slot zijn bestaande opbrengstindicatoren te verbeteren.

4 **Aanbeveling 2: houd bij oordeel over de school rekening met tekortkoming van opbrengstindicatoren**

De wettelijke opdracht in het basisonderwijs, het voortgezet onderwijs en het middelbaar beroepsonderwijs houdt méér in dan bijdragen aan leerwinst in taal en rekenen. Een belangrijk deel van de kwaliteit van scholen bestaat bijvoorbeeld uit het verzorgen van brede vorming in verschillende vakken en uit een gedegen voorbereiding op een vervolgopleiding. Het toezicht moet gericht zijn op deze brede kwaliteit van het onderwijs en het bevorderen van verbetering ervan. Dit type toezicht steunt niet alleen op opbrengstindicatoren.

Ook scholen zelf doen er goed aan om bij de bepaling van de kwaliteit van hun onderwijs verder te kijken dan naar interne rendements-, toets- en examencijfers. Zij kunnen bijvoorbeeld het succes van hun leerlingen in het vervolgonderwijs (en in de maatschappij) beter volgen.

4.1 **Laat in het toezicht de kwaliteitszorg van scholen zwaarder wegen**

De raad bepleit om opbrengstindicatoren een minder zware rol te laten spelen bij het toezicht op de onderwijskwaliteit. De kwaliteitszorg door scholen verdient een grotere plaats. Scholen moeten (bijvoorbeeld) zichtbaar maken wat hun leeraanbod is, welke pedagogisch-didactische werkwijzen ze hanteren, wat ze daarmee bereiken en hoe ze werken aan een systematische verbetering van hun onderwijs. In het toezicht moet een goede balans tot stand komen tussen opbrengst- en procesindicatoren. Instellingen moeten de ruimte krijgen om toe te lichten welke omstandigheden hebben geleid tot de resultaten.⁵¹ Ook als het gaat om opbrengsten waarvoor geen goede indicatoren beschikbaar zijn, is het toezicht aangewezen op procesindicatoren en op de toelichting van de instelling daarop. Dit geldt bijvoorbeeld voor sociale opbrengsten en vakken die niet centraal getoetst worden, niet alleen in het basisonderwijs en in het middelbaar beroepsonderwijs, maar ook in de onderbouw van het voortgezet onderwijs.

⁵¹ Onderwijsraad, 2013a.

Lage resultaten niet per definitie te wijten aan onderwijskwaliteit

In het bestaande toezicht op de basiskwaliteit leggen opbrengstindicatoren veel gewicht in de schaal. Een basisschool waarvan de opbrengstindicator (score bij de eindtoets) onvoldoende is, wordt in de regel als zwak beoordeeld. Beoordeling van proceskenmerken speelt alleen een rol bij de bepaling van het eindoordeel zwak of zeer zwak. Omgekeerd kan een school waarvan de opbrengstindicator voldoende is, alleen bij uitzondering als zwak worden gekwalificeerd (als de Inspectie vier of meer procesindicatoren als onvoldoende beoordeelt).⁵²

Gegeven de beperkingen van de opbrengstindicatoren is het cruciaal dat er ruimte is voor een inhoudelijke dialoog tussen de Inspectie en de school. Als een school op basis van dergelijke indicatoren tegenvallende resultaten lijkt te hebben, wil dat nog niet altijd zeggen dat de onderwijskwaliteit onvoldoende is. In de inhoudelijke dialoog kan onderscheid worden gemaakt tussen verschillende procesfactoren (bijvoorbeeld: hoe zorgt de school voor een goede kwaliteit van de lessen of hoe werkt de school aan systematische kwaliteitsverbetering).

Opbrengstindicatoren kunnen brede onderwijskwaliteit niet omvatten

Zelfs voor verbeterde opbrengstindicatoren geldt dat zij niet voldoende de onderwijskwaliteit omvatten (en nooit zullen omvatten) om daar de school alleen op te beoordelen. Dit blijkt onder meer uit interviews met interne begeleiders en schooldirecteuren in het basisonderwijs, in het kader van de pilot over leerwinst en toegevoegde waarde. 28% van de pilotscholen vindt dat de gerapporteerde toegevoegde waarde geen eerlijke vergelijking biedt van de (door de scholen ervaren) toegevoegde waarde ten opzichte van andere scholen. Een vergelijkbaar deel van de scholen (31%) vindt het geen goed idee om de Inspectie een rapport over de toegevoegde waarde van de school te laten opstellen.⁵³

Te zwaar gewicht van opbrengsten bij toezicht zet diagnostisch gebruik van gegevens onder druk

Een zwaar gewicht van opbrengstindicatoren bij de beoordeling van scholen heeft bovendien het risico dat de waarde van opbrengstgegevens voor diagnostisch gebruik binnen de school onder druk komt te staan. Er wordt immers strategisch gedrag van scholen mee uitgelokt (zie paragraaf 2.5). Een grote nadruk op toegevoegde waarde geeft scholen een incentive om zich vooral op de indicatoren voor toegevoegde waarde te richten, wat mogelijk ten koste kan gaan van niet-gemeten aspecten van kwaliteit. Dit wordt nog sterker wanneer de overheid de hoogte van de bekostiging koppelt aan dergelijke indicatoren, zoals bijvoorbeeld is voorgesteld voor het middelbaar beroepsonderwijs.

Dialoog past goed bij de kwaliteitsbevorderende taak die de Inspectie ook heeft

De Inspectie ziet niet alleen toe op de basiskwaliteit van scholen (waar normen bij passen), maar heeft ook een belangrijke kwaliteitsbevorderende taak (waarbij het instrument verantwoordt).⁵⁴ Bij deze laatste taak ligt ook de nadruk op de inhoudelijke dialoog. De wijze waarop scholen omgaan met gegevens over de ontwikkeling van hun leerlingen zou onderwerp van gesprek moeten zijn in deze verantwoordingsdialoog. In de ogen van de raad is dit een belangrijk onderdeel van de invulling van de kwaliteitsbevorderende rol van de Inspectie. De ervaringen die de Inspectie hierbij opdoet, kunnen weer benut worden voor de verdere verbetering van de gebruikte indicatoren.

⁵² Inspectie van het Onderwijs, 2012c.

⁵³ Janssens, Rekers-Mombarg & Lacor, 2014. 54

⁵⁴ Onderwijsraad, 2013a.

Verbeter bestaande opbrengstindicatoren door opleiding van ouders op te nemen

Het feit dat opbrengstindicatoren minder zwaar moeten wegen in het inspectieoordeel over de school, neemt niet weg dat bestaande indicatoren vatbaar zijn voor verbetering. De raad vindt dat het opleidingsniveau van ouders er preciezer in verdisconteerd moet worden. Zo ontstaat een genuanceerder beeld van de leeropbrengsten van scholen met verschillende leerlingenpopulaties, zonder dat een nieuw beoordelingssysteem nodig is. Ook de aanwezigheid van zorgleerlingen in de school zou standaard bij het bepalen van de opbrengsten van een school betrokken moeten worden. De Inspectie houdt nu slechts in beperkte mate rekening met de achtergrond van leerlingen. Dit levert soms een vertekend beeld op van de inspanningen die scholen leveren om de opbrengsten te realiseren en ook van de ruimte die bestaat voor verbetering. In dit opzicht is een verfijning van de informatie over leeropbrengsten (en daarmee van de bestaande opbrengstindicatoren) gewenst, met daarbij nogmaals de kanttekening dat het noodzakelijk is om met een bredere blik naar verschillende kwaliteitsaspecten van onderwijs te kijken.

Opleidingsniveau van ouders speelt slechts een kleine rol in inspectieoordeel

In het primair onderwijs wordt in de beslisregels van de Inspectie voor de beoordeling van leeropbrengsten alleen rekening gehouden met leerlinggewichten. Leerlingen krijgen extra 'gewicht' als hun ouders laagopgeleid zijn. Hoe zwaarder het totaal aan leerlinggewichten, hoe meer middelen de school krijgt. Bij de bepaling van het gewicht wordt alleen onderscheid gemaakt tussen een laag opleidingsniveau en een opleidingsniveau daarboven. Er bestaat geen onderscheid tussen leerlingen van ouders met een middelbaar niveau en die van ouders met een hoger opleidingsniveau. Dit kan betekenen dat de opbrengsten van scholen met vooral leerlingen van hoger opgeleide ouders overschat worden ten opzichte van de opbrengsten van scholen met vooral middelbaar opgeleide ouders. Meer gedetailleerde informatie over opleidingsniveaus van ouders kan een genuanceerder beeld van de leeropbrengsten opleveren.

De Inspectie heeft in dit kader samen met de PO-raad een verkenning uitgevoerd. De conclusie daarvan was dat de gevolgen voor de opbrengstenbeoordeling van scholen te beperkt zouden zijn om over te gaan tot aanpassing van de beoordelingssystematiek. Voor 4% van scholen zou de beoordeling veranderen van voldoende naar (zeer) zwak of vice versa. Dit resultaat zou niet opwegen tegen de hogere administratieve lasten en de verminderde transparantie voor scholen.⁵⁵ Deze verkenning heeft echter alleen gekeken naar wijzigingen in de beoordeling van scholen met betrekking tot het oordeel zwak en niet zwak. Buiten beschouwing bleven verschuivingen binnen de hele range van opbrengstuitkomsten.⁵⁶

De Universiteit Maastricht heeft wel gebruikgemaakt van een verfijnde indeling van het opleidingsniveau van ouders bij onderzoek naar de opbrengsten van scholen in basis- en voortgezet onderwijs (in het kader van *Limburg – proefregio opbrengstgericht werken*). Daaruit bleek dat bij het voorspellen van de gemiddelde scores op de Cito-eindtoets de nieuwe indeling heel andere uitkomsten gaf dan de indeling die alleen op gewichtenleerlingen is gebaseerd. Dat is belangrijk, omdat de beoordeling van basisscholen voor een groot deel gebaseerd is op de resultaten van de Cito-eindtoets. Andere gemiddelde Cito-scores geven dus aanleiding tot een andere classificatie van een aantal scholen (bijvoorbeeld zwak in plaats van voldoende,

⁵⁵ Ministerie van Onderwijs, Cultuur en Wetenschap, 2013a.

⁵⁶ Inspectie van het Onderwijs, 2012a.

of omgekeerd). Ook in het voortgezet onderwijs blijkt dan dat een aantal scholen dat volgens de Inspectie zwak is, bovengemiddeld scoort als beter rekening wordt gehouden met de achtergrond van leerlingen.⁵⁷ Het is daarom voor een beter beeld van de leeropbrengsten van belang dat bij de interpretatie van opbrengsten een differentiatie naar opleidingsniveau van de ouders wordt gehanteerd. De administratieve lasten die hiermee gepaard gaan, kunnen beperkt blijven door deze informatie eenmalig te registreren buiten de scholen om, bijvoorbeeld via het consultatiebureau. De raad heeft dit eerder bepleit in het advies *Vooruitgang boeken met achterstandsmiddelen*.⁵⁸

Tot slot wijst de raad er nadrukkelijk op dat scholen op dit moment een incentive hebben om kwetsbare leerlingen niet aan te nemen. Deze leerlingen hebben mogelijk een negatief effect op het rendement, het prestatieniveau of de cijfers over het bestrijden van voortijdig schoolverlaten zoals de Inspectie ze beoordeelt. Dit geldt zowel voor het primair en voortgezet onderwijs als voor mbo-instellingen.

4.3 Laat scholen het succes van hun leerlingen in het vervolgonderwijs volgen

Scholen doen er zelf ook goed aan om bij de bepaling van de kwaliteit van hun onderwijs verder te kijken dan interne rendements-, toets- en examencijfers. Het succes van hun leerlingen in het vervolgonderwijs en in de maatschappij kan ook een indicator zijn voor onderwijskwaliteit.

Basisscholen kunnen actiever de schoolloopbaan van hun leerlingen in het voortgezet onderwijs volgen. Hoewel de verantwoordelijkheid voor het niveau waarop leerlingen in het voortgezet onderwijs terechtkomen en het succes dat ze daar hebben niet geheel aan de basisschool zijn toe te schrijven, heeft deze wel een bijdrage geleverd aan hun ontwikkeling. Daarom kan het voor de afstemming van het onderwijsaanbod en het zicht op de onderwijskwaliteit relevant zijn voor basisscholen om het schoolsucces van hun leerlingen in het voortgezet onderwijs te volgen.

Ook het voortgezet onderwijs zou meer oog moeten hebben voor het succes van leerlingen in het vervolgonderwijs, zodat ook zij de voorbereiding erop kunnen verbeteren.

Het middelbaar beroepsonderwijs heeft volgens de WEB een drievoudige kwalificatieopdracht, namelijk uitoefening van een beroep, deelname aan de maatschappij (burgerschap) en doorstroom naar vervolgonderwijs. Ook hier kunnen indicatoren voor toegevoegde waarde op basis van leerwinst slechts een beperkt deel van de opbrengsten in beeld brengen. Dit verklaart voor een deel waarom de nu gehanteerde 'toegevoegde-waarde-indicator' in het middelbaar beroepsonderwijs niet op leerwinst is gebaseerd. Deze toegevoegde-waarde-indicator kijkt op dit moment alleen naar het diplomarendement van deelnemers op instellingsniveau, waarbij rekening wordt gehouden met de vooropleiding.⁵⁹ Gezien de kwalificatieopdracht van het middelbaar beroepsonderwijs is dit een te beperkte benadering van de onderwijsopbrengsten. In een brief over een proef met het gebruik van deze indicator wordt bovendien gesproken over de mogelijkheid om deze op termijn te gebruiken als criterium voor

⁵⁷ Schils, 2012.

⁵⁸ Onderwijsraad, 2013.

⁵⁹ Ministerie van Onderwijs, Cultuur en Wetenschap, 2013b.

de bekostiging.⁶⁰ In de ogen van de raad gaat dit gepaard met grote risico's op ongewenst strategisch gedrag.

Om een breder beeld van de opbrengsten te krijgen, kunnen mbo-instellingen baat hebben bij opbrengstgegevens op meerdere terreinen, zoals het succes van gediplomeerden in het vervolgonderwijs en op de arbeidsmarkt. In dit verband heeft de raad gepleit voor knippenlichtnormen die aangeven wanneer er bijvoorbeeld zorg is over het arbeidsmarktperspectief van afgestudeerden, of wanneer veel afgestudeerden achteraf spijt hebben van de gekozen opleiding.⁶¹

⁶⁰ Ministerie van Onderwijs, Cultuur en Wetenschap, 2013b.

⁶¹ Onderwijsraad, 2012.

Afkortingen

ict	informatie- en communicatietechnologie
mbo	middelbaar beroepsonderwijs
OCW	Onderwijs, Cultuur en Wetenschap
po	primair onderwijs
uoc	universitair onderwijscentrum
vo	voortgezet onderwijs
WEB	Wet educatie en beroepsonderwijs

Literatuur

- Allen, R. & Burgess, S. (2013). Evaluating the provision of school performance information for school choice. *Economics of Education Review*, 34, 175-190.
- Bonhomme, S., Jolivet, G. & Leuven, E. (2012). *Job Characteristics and Labor Turnover: Assessing the Role of Preferences and Opportunities in Teacher Mobility*. Geraadpleegd op 25 maart 2014 via http://www.cemfi.es/~bonhomme/profs_01July2013.pdf.
- Boudett, K.P., Cidy, E.A. & Murnane, R.J. (2007). *Data wise. A step-by-step guide to using assessment results to improve teaching and learning*. Cambridge: Harvard Education Press.
- Carrell, S.E. & Hoekstra, M.L. (2010). Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids. *American Economic Journal: Applied Economics*, 2(1), 211-228.
- Clotfelter, C.T., Ladd, H.F. & Vigdor, J.L. (2006). Teacher Sorting, Teacher Shopping, and the Assessment of Teacher Effectiveness. *Journal of Human Resources*, 41(4), 778-820.
- Deming, D., Cohodes, S., Jennings, J. & Jencks, C. (2013). *School Accountability, Postsecondary Attainment and Earnings*. Geraadpleegd op 21 maart 2014 via <http://www.nber.org/papers/W19444>.
- Dijkstra, A.B. (2012). *Sociale opbrengsten van onderwijs*. Amsterdam: Vossiuspers.
- Dobbie, W. & Fryer, R.G. (2011). Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children's Zone. *American Economic Journal: Applied Economics*, 3(3), 158-187.
- Ehlert, M., Koedel, C., Parsons, E. & Podgursky, M. (2014). Choosing the right growth measure. *Education Next*, 14(2), 66-71.
- Ehren, M.C.M. (2014). *Startdossier toegevoegde waarde*. Studie in opdracht van de Onderwijsraad. London/Enschede: Institute of Education/Universiteit Twente.
- Figlio, D. & Loeb, S. (2011). School Accountability. In E.A. Hanushek, S. Machin & L. Woessmann (eds.), *Handbook of the Economics of Education, Volume 3* (383-421). Amsterdam: Elsevier.
- Gemeente Amsterdam DMO (2010). *Resultaten Amsterdams basisonderwijs 2010*. Amsterdam: Gemeente Amsterdam DMO.
- Gneezy, U., Meier, S. & Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, 25(4), 191-210.
- Haertel, E.H. (2013). *Reliability and validity of inferences about teachers based on student test scores*. Princeton: ETS.
- Inspectie van het Onderwijs (2012a). *Differentiatie van het opleidingsniveau van ouders bij het beoordelen van opbrengsten van basisscholen*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2012b). *Toezichtkader Beroepsonderwijs en volwasseneneducatie 2012: Uitgangspunten, werkwijze, waarderingskader*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2012c). *Toezichtkader PO/VO 2012*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2013a). *De staat van het onderwijs. Onderwijsverslag 2011/2012*. Utrecht: Inspectie van het Onderwijs.
- Inspectie van het Onderwijs (2013b). *Over de volle breedte*. Utrecht: Inspectie van het Onderwijs.
- Jackson, C.K. (2013). Can higher-achieving peers explain the benefits to attending selective schools? Evidence for Trinidad and Tobago. *Journal of Public Economics*, 108, 63-77.
- Janssens, F.J.G., Rekers-Mombarg, L. & Lacor, E. (2014). *Leerwinst en toegevoegde waarde in het primair onderwijs*. Enschede/Groningen/Rotterdam: Universiteit Twente/GION/CED-Groep.
- Kane, T.J. & Staiger, D.O. (2002). The promise and pitfalls of using imprecise school accountability measures. *Journal of Economic Perspectives*, 16(4), 91-114.

- Ladd, H.F. & Walsh, R.P. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, 21(1), 1-17.
- Lang, K. (2010). Measurement matters: perspectives on education policy from an economist and school board member. *Journal of Economic Perspectives*, 24(3), 167-182.
- Mandinach, E.B. & Honey, M. (2008). *Data-driven school improvement. Linking data and learning*. New York: Teachers College Press.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2013a). *Differentiatie van het opleidingsniveau van ouders bij het beoordelen van opbrengsten van basisscholen*. Brief van Staatssecretaris van OCW aan Voorzitter Tweede Kamer, 7 januari 2013. Kamerstukken II, 2013-2014, 32193, 35. Kenmerk: 451663.
- Ministerie van Onderwijs, Cultuur en Wetenschap (2013b). *Kwaliteitsafspraken mbo*. Brief van Minister van OCW aan Voorzitter Tweede Kamer, 16 december 2013. Kamerstukken II, 2013-2014, 31524, 189. Kenmerk: 572683.
- Neal, D. & Schanzenbach, D.W. (2010). Left Behind by Design: Proficiency Counts and Test-based Accountability. *Review of Economics and Statistics*, 92(2), 263-283.
- Onderwijsraad (2011a). *Ruim baan voor stapsgewijze verbeteringen*. Den Haag: Onderwijsraad.
- Onderwijsraad (2011b). *Toetsing in het primair onderwijs*. Den Haag: Onderwijsraad.
- Onderwijsraad (2012). *Zicht op een macrodoelmatig opleidingsaanbod*. Den Haag: Onderwijsraad.
- Onderwijsraad (2013a). *Publieke belangen dienen. Naar bestuurlijk evenwicht tussen overheid en onderwijsinstellingen*. Den Haag: Onderwijsraad.
- Onderwijsraad (2013b). *Vooruitgang boeken met achterstandsmiddelen*. Den Haag: Onderwijsraad.
- Organisation for Economic Cooperation and Development (2008). *Measuring Improvements in learning Outcomes*. Paris: OECD.
- Sass, T.R., Semykina, A. & Harris, D.N. (2013). Value-Added Models and the Measurement of Teacher Productivity. *Economics of Education Review*, 38, 9-23.
- Schildkamp, K. (2012). Werken met datateams op scholen. *Examens*, 9(2), 5-9.
- Schildkamp, K. (2011). Onderbuikgevoelens kloppen vaak niet. Werken met datateams aan onderwijsverbetering. *Meso Focus*, 83, 75-85.
- Schils, T. (2012). *Onderwijsverbetering door beter gebruik van beschikbare gegevens. Limburg proef-regio?* Presentatie voor de Inspectie van het Onderwijs. September 2012.
- Timmermans, A.C. (2012). *Value added in educational accountability: Possible, fair and useful?*. Proefschrift. Groningen: Rijksuniversiteit Groningen.
- Van der Sluis, M.E., Reezigt, G.J. & Borghans, L. (2013). Quantifying Stakeholder Values of VET Provision in the Netherlands. *Vocations and Learning*, 7(1), 1-19.
- Vreeburg, B. (2013). *Opbrengstbeoordeling van scholen: gebruik van toegevoegde waarde en leerwinst (presentatie tijdens expertmeeting 16 juli 2013)*. Utrecht: Inspectie van het Onderwijs.

Geraadpleegde deskundigen

Individuele gesprekken

De heer A. Béguin	Cito
De heer R. Feskens	Cito
De heer A. Jonk	Inspectie van het Onderwijs
Mevrouw F. Hendricks	School aan Zet
De heer F. Janssens	Inspectie van het Onderwijs/ Universiteit Twente
Mevrouw M. van der Lubbe	PO-Raad
Mevrouw H. van Oostrom	VO-Raad
Mevrouw L. Rekers-Mombarg	GION
De heer W. Schut	PO-Raad/ Lid CvB SCO Delft
Mevrouw A. Timmermans	GION
De heer P. Veelenturf	MBO-Raad
Mevrouw C. Versteeg	CED-Groep
Mevrouw J. Visser	Cito
De heer G. Wammes	MBO-Raad
Mevrouw M. Wijshake	Basisschool Het Mozaïek Delft
Mevrouw D. Zeldenrijk	School aan Zet

Schriftelijke bijdragen

Besturenraad
Suzanne Beek (KPC-groep)

Panelbijeenkomst Inspectie van het Onderwijs, 5 september 2013

De heer S. Baijens	Inspectie van het Onderwijs
Mevrouw I. Drewes	Inspectie van het Onderwijs
De heer B. Lichtenberg	Inspectie van het Onderwijs
De heer K. Teunis	Inspectie van het Onderwijs
Mevrouw E. Timminga	Inspectie van het Onderwijs
De heer B. Vreeburg	Inspectie van het Onderwijs

Panelbijeenkomst primair onderwijs, 6 september 2013

Mevrouw J. Aarts	Sint Bavo, Rijsbergen
Mevrouw T. van den Berg	Bos en Lommerschool, Amsterdam
De heer P. van Dijk	P.C. Jenaplan Basisschool De Peppels, Boxmeer
Mevrouw C. van der Grinten	De Vrije Ruimte, Den Haag
De heer B. Hakkeling	CBS De Ark, Vlaardingen
De heer T. de Jong	Bos en Lommerschool, Amsterdam
Mevrouw I. Schaefers	Bos en Lommerschool, Amsterdam

Panelbijeenkomst voortgezet onderwijs, 13 september 2013

Mevrouw A. Buijs	Herbert Vissers College, Nieuw Vennep
De heer P.J.J. Hendrikse	Ons Middelbaar Onderwijs, Tilburg
De heer W. Huiberts	Clusius College, Castricum
De heer F.E.G. Kessels	BC Broekhin, Roermond
De heer M. Schoon	Scholengemeenschap Reigersbos, Amsterdam

Panelbijeenkomst mbo, 13 september 2013

De heer C. Govers	Onderwijsgroep Tilburg
Mevrouw F.W. Hengeveld	AOC De Groene Welle, Zwolle
De heer J. de Jong	ROC van Amsterdam
De heer M. Labij	ROC Midden Nederland, Utrecht
De heer R. Meijerink	ROC van Twente, Hengelo
De heer F. Sonsma	ROC Friese Poort, Drachten
De heer A. Vlaardingebroek	Noorderpoort College, Groningen

Bijlage 1

Adviesvraag



> Retouradres Postbus 16375 2500 BJ Den Haag

Aan de voorzitter van de Onderwijsraad
Mevrouw Prof. dr. G.T.M. ten Dam
Nassaulaan 6
2514 JS DEN HAAG

Kennis
IPC 5200

Rijnstraat 50
Den Haag
Postbus 16375
2500 BJ Den Haag
www.rijksoverheid.nl

Contactpersoon
J.T. Schokker

T +31 6 52 36 74 26
t.schokker@minocw.nl

Onze referentie
577805

Uw brief van

Datum **06 JAN. 2014**

Betreft Adviesaanvraag 'Toegevoegde waarde en leerwinst'

Geachte mevrouw Ten Dam,

Het systematisch in beeld brengen van leervorderingen op basis van toetsen biedt veel kansen om de kwaliteit van het onderwijs te verbeteren. In alle onderwijssectoren wordt ervaring opgedaan met het meten van leerwinst en toegevoegde waarde. Ook in de politiek is er veel aandacht voor het gebruik van toetsen in het onderwijs. Het onderwerp stond onlangs tijdens de behandeling van de begroting van onderwijs hoog op de agenda. Hoewel er veel verwacht wordt van het systematisch in kaart brengen van leerwinst en toegevoegde waarde zijn er ook zorgen. Gaat het onderwijs zich niet teveel richten op de kennis die getoetst wordt? Ook in de Tweede Kamer leeft de vraag hoe onderwijskwaliteit het beste gemeten kan worden.¹ Tegelijk leven ook in de Tweede Kamer hoge verwachtingen van het in kaart brengen van de leerwinst en toegevoegde waarde.

Omschrijving begrippen

In deze adviesaanvraag hanteren we de volgende definities van de begrippen 'leerwinst' en 'toegevoegde waarde'. Leerwinst wordt gedefinieerd als de toename van vaardigheden of kennis van individuele leerlingen of groepen van leerlingen, gedurende (een bepaald deel van) de leerweg. De toegevoegde waarde geeft aan welke bijdrage de school of onderwijsinstelling levert aan de ontwikkeling (of leerwinst) van alle leerlingen. Met behulp van de toegevoegde waarde kan worden nagegaan of de onderwijsopbrengst boven of onder het niveau ligt dat op basis van het beginniveau verwacht mocht worden.

Drie doelen van informatie over toegevoegde waarde en leerwinst

Inzicht in de toegevoegde waarde en leerwinst dient verschillende doelen. Het eerste doel is het bieden van informatie voor scholen om aan de eigen kwaliteitsverbetering te werken. Het in beeld brengen van leerwinst en

¹ De aangenomen motie Klaver c.s. verzoekt de regering om met het onderwijsveld tot nuttige criteria te komen om de kwaliteit van het onderwijs te meten. Tweede Kamer, vergaderjaar 2013–2014, 33 750 VIII, nr. 39

toegevoegde waarde biedt scholen en besturen beter inzicht in de eisen van de leerlingen en de bijdrage die de school hieraan levert. Het verschaft scholen waardevolle informatie voor de eigen evaluatie van de kwaliteit van het onderwijs en aanknopingspunten van verbetering van het onderwijs. Inzicht in de toegevoegde waarde en leerwinst van scholen maakt het mogelijk het eigen lerend vermogen van onderwijsinstellingen te bevorderen.

Onze referentie
577805

Het tweede doel is het bieden van informatie voor belanghebbenden, met name ouders, als een basis voor verantwoording voor het gevoerde beleid door schoolbesturen. Besturen moeten tegenspraak kans geven en daarin hoort ook transparante informatie over de bereikte opbrengsten. Ook kan informatie over leerwinst en toegevoegde waarde keuzeprocessen van ouders en leerlingen ondersteunen.

Het derde doel is tenslotte het gebruik van informatie voor verantwoording naar de overheid. Toegevoegde waarde en leerwinst zijn nuttige begrippen in de beoordeling van de opbrengsten van een school of onderwijsinstelling. Het gaat dan niet alleen over de vraag of een school of onderwijsinstelling voldoet aan de basisnormen. De Inspectie van het Onderwijs heeft immers niet alleen een handhavende maar ook kwaliteit bevorderende taak. Toegevoegde waarde en leerwinst kunnen benut worden om scholen en onderwijsinstellingen constructieve feedback te geven.

Adviesaanvraag toegevoegde waarde en leerwinst

Vanuit deze context stel ik voor de volgende adviesaanvraag aan de Onderwijsraad voor te leggen:

Op welke wijze kunnen gegevens over leerwinst en toegevoegde waarde een bijdrage leveren aan de kwaliteit van het onderwijs uitgaande van de huidige onderwijspraktijk?

In het beantwoorden van deze hoofdvraag onderscheiden we de volgende deelvragen:

Omgaan met spanning tussen verbeteren en verantwoorden

Er bestaat een spanning tussen verbeteren (eerste doel) en verantwoorden ten opzichte van direct belanghebbenden en de overheid (tweede en derde doel). Idee daarbij is dat intrinsiek gemotiveerde leer- en verbeterprocessen gehinderd kunnen worden als die vermengd worden met halen van externe doelen. Vraag aan de Onderwijsraad is hoe de verschillende doelen van toegevoegde waarde en leerwinst gecombineerd kunnen worden. Hoe wordt bereikt dat instrumenten voor toegevoegde waarde en leerwinst zijn ingebed in de onderwijspraktijk en ook als eigendom worden ervaren van leraren? Hoe worden ongewenste strategische gedragseffecten voorkomen en ontmoedigd? Hoe kan voorkomen worden dat de toegevoegde waarde en leerwinst systematiek onbedoeld een rem zet op vernieuwing en innovatie in het onderwijs?

Praktische belemmeringen wegnemen

Onze referentie
577805

Op dit moment wordt in veel onderwijssectoren ervaring opgedaan met het gebruik van de begrippen toegevoegde waarde en leerwinst. Zo is in het primair onderwijs in 2011 een pilot gestart waarvan de resultaten begin 2014 verschijnen. In het voortgezet onderwijs is recent gestart met een pilot. In het middelbaar beroepsonderwijs wordt de eerste ervaringen opgedaan met toegevoegde waarde en wordt nagedacht over de mogelijkheid dit instrument te betrekken bij de bekostiging. Deze eerste ervaringen illustreren dat elke sector een andere uitgangssituatie heeft. U wordt gevraagd per sector te adviseren hoe praktische belemmeringen kunnen worden overwonnen en risico's hanteerbaar gemaakt kunnen worden. U wordt daarbij uitgenodigd lessen die andere landen hebben opgedaan met het benutten van toegevoegde waarde in het onderwijs te vertalen naar de Nederlandse context.

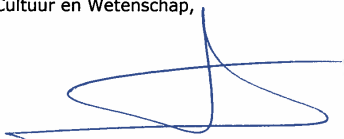
Met vriendelijke groet,

de minister van Onderwijs,
Cultuur en Wetenschap,

de staatssecretaris van Onderwijs
Cultuur en Wetenschap,



dr. Jet Bussemaker



Sander Dekker

Bijlage 2

Wat meten leerwinst en toegevoegde waarde?

Leerwinst meet de toename van vaardigheden of kennis. Dat betekent dus in principe dat vaardigheden en kennis op twee momenten in de tijd moeten worden gemeten. Hoe werkt dat?

De kennis op een bepaald gebied kan gemeten worden door een aantal vragen over dat onderwerp voor te leggen. Tabel 1 geeft een gestileerd voorbeeld, met vier vragen per toets. De kolom schaal geeft een aantal rekenvragen met onderaan de eenvoudigste vraag en bovenaan de moeilijkste vraag. Als het zo zou zijn dat leerlingen die een bepaalde vraag kunnen maken ook alle eenvoudigere vragen goed kunnen beantwoorden, zou het niveau aangeduid kunnen worden op basis van de moeilijkste vraag. In de praktijk is dit natuurlijk niet het geval, maar wordt de kans dat een vraag goed wordt gemaakt wel steeds lager als de moeilijkheidsgraad toeneemt. Door veel meer vragen te gebruiken dan in dit voorbeeld kan daarom gekeken worden bij welke vraag de kans 50% is dat een leerling het antwoord kent. Dat punt wordt dan het niveau van de leerling genoemd.

De meeste leerlingen in groep 3 zullen een niveau aan de onderkant van de schaal hebben. Het heeft daarom weinig zin hen de moeilijkste vragen voor te leggen. Ze maken die bijna nooit goed en het zou frustrerend zijn deze vragen te moeten maken. Naarmate leerlingen beter worden kunnen er moeilijkere vragen worden voorgelegd. Door te zorgen voor overlap tussen de vragen van de verschillende toetsen kunnen de scores van toetsen die verschillen qua moeilijkheidsgraad onderling worden vergeleken.

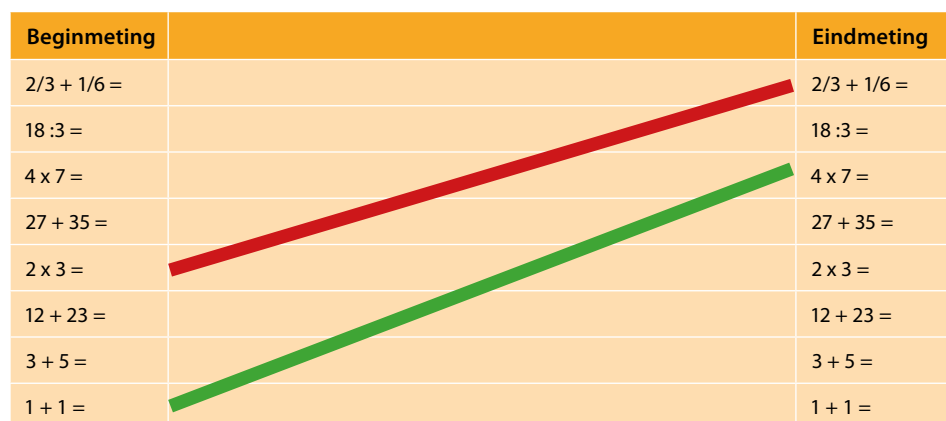
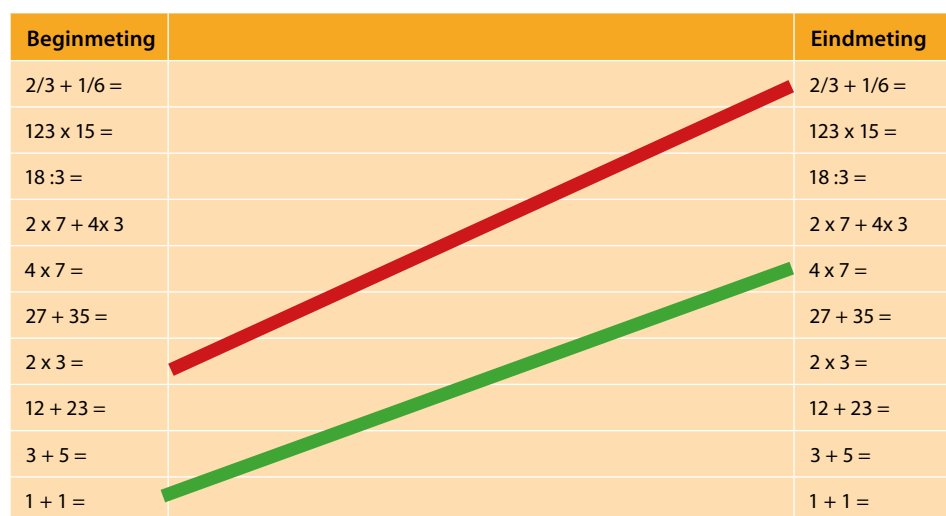
Tabel 1. Voorbeeld van overlappende rekenvragen op verschillende toetsen en bijbehorende schaal als basis voor het berekenen van leerwinst

Toets 1	Toets 2	Toets 3	Toets 4	Schaal
			$2/3 + 1/6 =$	$2/3 + 1/6 =$
			$123 \times 15 =$	$123 \times 15 =$
		$18 : 3 =$	$18 : 3 =$	$18 : 3 =$
		$2 \times 7 + 4 \times 3$	$2 \times 7 + 4 \times 3$	$2 \times 7 + 4 \times 3$
	$4 \times 7 =$	$4 \times 7 =$		$4 \times 7 =$
	$27 + 35 =$	$27 + 35 =$		$27 + 35 =$
$2 \times 3 =$	$2 \times 3 =$			$2 \times 3 =$
$12 + 23 =$	$12 + 23 =$			$12 + 23 =$
$3 + 5 =$				$3 + 5 =$
$1 + 1 =$				$1 + 1 =$

Op deze wijze kan de score op toets 1 afgetrokken worden van de score op toets 4 en ontstaat een maat voor leerwinst. Deze winst verschilt echter om twee redenen van de financiële winst zoals die door bedrijven wordt gemaakt. In de eerste plaats is het niet zonder meer zeker dat de kennis die hier gemeten wordt ééndimensionaal is. Een leerling zou goed kunnen zijn in vermenigvuldigen maar niet in delen. Voortuitgang op de schaal zou dan gemiddeld zijn, zonder dat de specifieke lacunes daarbij in beeld komen. In de tweede plaats levert een schaal voor het meten van kennis geen informatie op over de omvang van de stappen die gemaakt worden. De kennisschaal is ordinaal. Winst in de economische betekenis wordt uitgedrukt in geld en kent daardoor een zogenaamde ratioschaal. Iedere eenheid extra winst is even groot. Bij een maat voor kennis kan alleen gezegd worden dat iemand meer weet, maar het is niet mogelijk

om deze hoeveelheid extra kennis een waarde toe te kennen zonder daarbij een extern ijkpunt te hanteren. Dit kan verbeeld worden door in op de schaal enkele vragen weg te halen. De afstand tussen twee punten op de schaal wordt dan kleiner. Figuur 4 vergelijkt de leerwinst van twee leerlingen en laat zien dat op deze manier het in het ene geval kan lijken alsof de betere leerling meer leerwinst boekt, terwijl het in het andere geval lijkt alsof de andere leerling meer leerwinst boekt. Moderne statistische technieken zijn in staat om vragen naar moeilijkheidsgraad op een statistische zinvolle manier te schalen, maar de vraag of we een bepaalde stijging in het kennisniveau aan de onderkant even groot vinden als een bepaalde stijging aan de bovenkant van de schaal, is niet een statistische maar een inhoudelijke vraag over wat we belangrijk vinden in het onderwijs. Dat is geen kwestie van meten, maar een kwestie van keuzes maken door de betrokkenen bij het onderwijs.

Figuur 4. Voorbeeld van ontwikkeling leerwinst bij verschillen in schaal



De stap van leerwinst naar toegevoegde waarde wordt gemaakt door de leerwinst naar schoolniveau te aggregeren en door de component te isoleren waar de school verantwoordelijk voor is. Hierbij doen zich twee moeilijkheden voor. In de eerste plaats is het gebruikelijk om toegevoegde waarde uit te drukken in de gemiddelde leerwinst per leerling. Het voorbeeld van figuur 4 laat echter zien dat een gemiddelde geen goede samenvatting van resultaten hoeft te geven bij ordinale schalen. Een school met veel 'rode' leerlingen zou immers beter scoren dan een school met veel 'groene' leerlingen in geval de eerste schaal wordt gebruikt, terwijl deze slechter zou scoren als de tweede schaal wordt gebruikt. Bij toegevoegde waarde zou daarom naar de verschillende groepen leerlingen afzonderlijk moeten worden gekeken, of er moet een duidelijke keuze worden gemaakt over de waardering van kennisgroei in verschillende delen van de kennisschaal.

Daarnaast pretendeert het concept toegevoegde waarde aan te geven voor welk deel van de leerwinst de school verantwoordelijk is. Doorgaans doet men dat door de gemeten groei te corrigeren voor verklarende factoren, zoals het kennisniveau bij het begin van de periode, het ouderlijk milieu van het kind, en dergelijke. Toegevoegde waarde wordt daarmee teruggebracht tot de vraag of een school het beter doet dan een andere school met vergelijkbare kenmerken. Er is daarbij heel veel keuze met betrekking tot de kenmerken die ter correctie kunnen worden meegenomen, terwijl het lastig is om een scheiding te maken tussen factoren waar de school voor verantwoordelijk is en factoren die buiten de school om een rol spelen. In de wetenschappelijke literatuur is er een groeiende consensus dat deze causaliteitsvraag alleen adequaat beantwoord kan worden door 'at random' sommige leerlingen aan de ene en andere leerlingen aan de andere school toe te wijzen. Fryer en Dobbie gebruiken in een studie naar de leerwinst van bepaalde typen 'charter school' daarom de loting waarmee bepaald is dat sommige leerlingen wel en andere niet worden toegelaten.⁶²

⁶² Dobbie & Fryer, 2011.

Bijlage 3

Verskillende modellen van opbrengstindicatoren

In onderstaande tabel wordt een overzicht gegeven van veelgebruikte modellen om de opbrengst van een school te berekenen. Het groeimodel en het statusmodel zijn twee modellen die gebruikt kunnen worden voor het stellen van targets aan prestaties van scholen.⁶³

Model	Methode	Omschrijving
Norm-referenced*	Normgerelateerde standaard	Het gemiddelde eindresultaat van leerlingen op een school, gecorrigeerd voor leerlingkenmerken/ gemiddelde niveau van vergelijkbare scholen minus een halve standaarddeviatie
	Percentile rank	Het percentage van een referentiegroep leerlingen die lagere scores behaalden dan de betreffende groep leerlingen in de school
	Standaard score	De prestaties van een (groep) leerling(en) ten opzichte van een gemiddelde (groep) leerling(en) (uitgedrukt in z-scores)
	Grade equivalent	De prestaties van een (groep) leerling(en) in vergelijking met de verwachte prestaties van de betreffende leergroep/klas met dezelfde leeftijd
Criterion-referenced		De prestaties van leerlingen in vergelijking met vaststaande standaarden (bijvoorbeeld referentieniveaus). NB: geen vergelijking met prestaties van andere leerlingen
Groeimodel**	Cross-sectioneel	Vergelijking van prestaties groep 8 leerlingen dit jaar met prestaties van groep 8 leerlingen het jaar daarvoor
	Quasi-longitudinaal	Vergelijking van prestaties groep 8 leerlingen dit jaar met prestaties van groep 7 leerlingen het jaar daarvoor
	Longitudinaal	Prestaties van individuele leerlingen worden vergeleken met hun eigen eerdere prestaties
	Relatieve vooruitgang	Werkelijke vooruitgang wordt afgewogen ten opzichte van voorspelde vooruitgang
Statusmodel		De prestaties op één moment afzetten tegen een standaard (bijvoorbeeld minimum prestatie, gemiddeld prestatieniveau van vergelijkbare scholen of een historisch gemiddelde). Hierbij wordt bijvoorbeeld gecorrigeerd op basis van het beginniveau

* Resultaten van leerlingen in een school worden afgezet tegen die van een normpopulatie (vergelijkbare groep scholen).

⁶³ Ehren, 2014; Vreeburg, 2013.

In dit advies geeft de Onderwijsraad antwoord op de vraag in hoeverre leerwinst en toegevoegde waarde geschikte instrumenten zijn voor verbetering van het onderwijs. Scholen hebben steeds meer gegevens over de vorderingen van leerlingen. Tegelijkertijd wordt daar tot dusver weinig gebruik van gemaakt.

De raad komt tot de conclusie dat leerwinst en toegevoegde waarde wél goede instrumenten zijn voor scholen om hun onderwijs te verbeteren, maar níet voor een oordeel over de kwaliteit van een school.