

Examens en toetsen in accountability systemen

□ Melanie Ehren

De laatste jaren worden examens en toetsresultaten steeds vaker gebruikt om scholen te beoordelen, als onderdeel van accountability systemen in het onderwijs. In een goed functionerend (*test-based*) accountability systeem moet zorgvuldig en terughoudend worden omgegaan met targets voor toetsresultaten en daaraan gerelateerde sancties.

Doel van examens en toetsen

Examens en toetsresultaten worden in de eerste plaats gebruikt om leerlingen te beoordelen en om aanvullend onderwijs of begeleiding te geven over onderwerpen die zij nog niet zo goed beheersen. Examens en toetsresultaten hebben daarnaast ook een kwalificerende en selecterende functie wanneer leerlingen hun opleiding afsluiten met een examen en (bij goed resultaat) met een diploma toegang krijgen tot vervolgonderwijs. Van meer recente datum is het gebruik van examens en toetsresultaten als middel ter beoordeling van de kwaliteit van scholen.

Accountability systemen

Door middel van accountability systemen leggen scholen verantwoording af over de aspecten van het onderwijs die de samenleving van belang vindt. Scholen worden geëvalueerd aan de hand van vaststaande doelen en targets; ze krijgen sancties opgelegd bij het niet behalen van de targets of beloningen toegekend wanneer ze beter presteren dan de minimale target. In Nederland bestaat het accountability systeem vooral uit het inspectietoezicht waarbij scholen worden beoordeeld op de output die zij leveren: de leerprestaties van leerlingen. Daarnaast wordt de inrichting van hun onderwijsproces geëvalueerd in

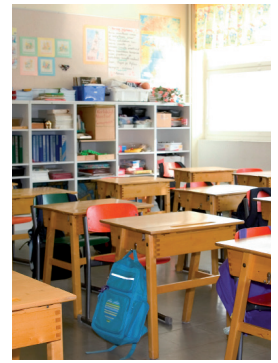
relatie tot wet en regelgeving of tot professionele normen over goed onderwijs zoals de kwaliteit van instructie en begeleiding van leerlingen. De laatste jaren is bij de beoordeling en verantwoording van scholen steeds meer nadruk gelegd op examens en toetsresultaten.

De onderwijsinspectie introduceerde ruim een jaar geleden het risicogestuurde toezicht. Voor het toezicht gebruikt de inspectie een aantal indicatoren, waaronder de resultaten van leerlingen op toetsen tijdens en aan het eind van de schoolloopbaan. Deze gegevens in combinatie met andere indicatoren, zoals uitval van leerlingen en ziekteverzuim van personeel, vormen de basis voor een inschatting van de kwaliteit van een school. Deze ontwikkeling past in de trend om scholen vooral op resultaten te beoordelen en vrijheid te geven in de inrichting van de (onderwijs)processen.

Ervaringen in de Verenigde Staten van Amerika

Het Amerikaanse (test-based) accountability systeem kent al langer een benadering van afrekenen op en verantwoorden voor resultaten. Sinds de invoering van de *No Child Left Behind Act* in 2001 zijn staten verplicht om een accountability systeem te implementeren waarbij alle leerlingen jaarlijks worden getoetst en waarbij deze toetsresultaten worden gebruikt om inzicht te geven in de prestaties van de school. Staten zijn daarbij verplicht om targets te formuleren waaraan scholen moeten voldoen; scholen die hier niet aan voldoen krijgen sancties opgelegd zoals boetes of vervanging van het managementteam van de school. Scholen die beter presteren worden in sommige staten (financieel) beloond. In sommige staten worden ook interventieteams ingezet van experts die scholen helpen om beter te presteren. De targets die scholen moeten halen kunnen bestaan uit een minimumscore op een toets (*cut point*), een jaarlijkse groeitarget waarbij steeds een hogere score moet worden gehaald, of het verkleinen van het verschil in

Door middel van accountability systemen leggen scholen verantwoording af



Toezicht voor goed onderwijs – Inspectie van het onderwijs

Op basis van de jaarlijkse risicoanalyse schat de inspectie in of er bij een school risico's zijn voor de onderwijskundige en financiële kwaliteit en voor de naleving van wet- en regelgeving. De inspectie voert de risicoanalyse uit op basis van de opbrengsten, de jaarstukken en eventuele signalen over de school.

Opbrengstgegevens vormen een belangrijk onderdeel van de risicoanalyse. In het basisonderwijs gaat het in eerste instantie over de eindtoetsen. In het voortgezet onderwijs kijkt de inspectie vooral naar de resultaten van het schoolexamen en het centraal examen, maar bijvoorbeeld ook naar het percentage leerlingen dat zonder te blijven zitten het diploma haalt. In de analyse houdt de inspectie rekening met de leerlingenpopulatie op de school.

Meer informatie over de werkwijze van de inspectie is te vinden op de website: www.onderwijsinspectie.nl.

prestaties van leerlingen die goed en zwak presteren (verkleinen van de *achievement gap*).

De combinatie van dergelijke targets en zware sancties bij het niet behalen van deze targets blijkt in de Verenigde Staten tot schadelijke neveneffecten te leiden, met name in scholen met veel achterstandsleerlingen die grote moeite hebben om de targets te halen. Eén van de wetenschappers, Dan Koretz van Harvard University, die veel onderzoek doet naar neveneffecten van *test-based accountability* verklaart het als volgt:

"The problem is that you can raise scores the hard way by teaching more effectively and getting the students to work harder, or you can take shortcuts and start figuring out ways to "game" the system." [New York Times, 7 oktober 2007]

Strategisch gedrag van scholen

Scholen gebruiken dus strategisch gedrag om targets te halen en sancties te voorkomen. Er zijn verschillende manieren waarop ze dat kunnen doen; variërend van frauduleus tot twijfelachtig. Frauderen gebeurt door antwoorden van leerlingen op toetsen te wijzigen, door het voorzeggen van antwoorden en het neerleggen van hulpmaterialen in het lokaal. Het oefenen van oude vergelijkbare toetsen is redelijk gangbaar. In de Verenigde Staten is veel te doen over allerlei vormen van *'teaching to the test'* waarbij

veel instructietijd in de klas wordt opgeofferd om leerlingen voor te bereiden op de toets. De vele uren waarin leerlingen oefenen met oude toetsopgaven, strategieën krijgen aangeleerd om bepaalde typen vragen goed te beantwoorden en alleen nog onderwijs krijgen in onderwerpen die standaard in de toets voorkomen, blijken uiteindelijk een negatief effect te hebben op leerprestaties van leerlingen. Ze scoren weliswaar goed op de toets die in het accountability systeem wordt gebruikt, maar veel minder goed op toetsen waaraan geen targets zijn verbonden en waar andere onderwerpen of vaardigheden worden gemeten.

Andere vormen van strategisch gedrag zijn het selecteren van leerlingen voor deelname aan de toets (*reshaping the test pool*), waarbij leerlingen die naar verwachting laag scoren ziek worden gemeld of worden geplaatst in een categorie leerlingen die niet hoeft deel te nemen aan de toets (bijvoorbeeld de categorie leerlingen die naar het speciaal onderwijs wordt verwezen). Selectie vindt in sommige gevallen ook al plaats 'aan de poort' wanneer leerlingen die naar verwachting laag gaan presteren niet tot de school worden toegelaten.

Targets die bestaan uit minimale toetsscores blijken tot slot vooral te leiden tot strategisch gedrag dat met *'educational triage'* wordt aangeduid. Scholen richten hun aandacht en (extra) instructie en begeleidingstijd



Scholen gebruiken soms strategisch gedrag om hun targets te halen

vooral op leerlingen die rondom de grensscore presteren, met als doel om deze leerlingen een toetscore op of boven de minimale targetscore te laten halen. De leerlingen die ver onder de minimale targetscore presteren worden als 'hopeloos' beschouwd en ook de leerlingen die al goed presteren krijgen geen extra aandacht. Deze vormen van accountability stimuleren scholen vooral om een zo hoog mogelijke overall score te halen in plaats van elke leerling de tijd, aandacht en begeleiding te geven die deze nodig heeft.

Positieve effecten van accountability systemen

Het gebruik van toetsresultaten in accountability systemen kan natuurlijk ook positieve effecten hebben. De lessen uit de Verenigde Staten kunnen helpen om het Nederlandse accountability systeem met risicogestuurd toezicht zo in te richten dat ongewenste neveneffecten worden voorkomen en positieve effecten worden gestimuleerd. Gewenste effecten zijn dan in de eerste plaats het verbeteren van het onderwijs en de leerprestaties van alle leerlingen. Toetsresultaten kunnen hierbij een positieve rol spelen wanneer leerkrachten en schoolleiders deze gebruiken om na te gaan welke leerlingen, in welke klassen en in welke leerstofgebieden extra ondersteuning en begeleiding nodig hebben. Het risicogestuurde toezicht kan scholen en schoolbesturen stimuleren om (ook) te handelen op basis van leerprestaties en toetsresultaten van leerlingen in plaats van alleen op basis van activiteiten en randvoorwaardelijke doelen.

Accountability systemen die naar verwachting deze

positieve effecten stimuleren, beschikken over een meting van de toegevoegde waarde van de school. Daarbij laat een meting van het aanvangsniveau van de leerling in combinatie met een eindmeting zien wat de school daadwerkelijk bijdraagt aan de leerprestaties van leerlingen. De output van de scholen kan worden beoordeeld op deze bijdrage aan de leerprestaties. Een maat voor onderwijskwaliteit en leerprestaties van een school op basis van toetsresultaten moet betrouwbaar en valide zijn, en ongevoelig voor eventuele manipulatie en strategisch gedrag van scholen. In dat verband is bijvoorbeeld speciale aandacht nodig voor kleine scholen met minder dan vijftien leerlingen in een klas, waarbij de gemiddelde score op de eindtoets wordt beïnvloed door de resultaten van een enkele leerling die hoog of laag presteert. Ook de vraag of leerprestaties in taal en rekenen alleen een voldoende indicatie geven van de algehele kwaliteit van een school is relevant; veel scholen (en ouders) vinden ook andere cognitieve of sociaal-emotionele (leer)prestaties van leerlingen van belang en verwachten dat de school daar een bijdrage aan levert.

Voorwaarden voor het goed functioneren van een accountability systeem

Het onderwijsnummer waarmee leerlingen tijdens hun hele schoolloopbaan kunnen worden gevolgd, biedt goede perspectieven om een maat te ontwikkelen waarmee de bijdrage van de school aan leerprocessen in kaart kan worden gebracht. Er kan met behulp van het onderwijsnummer bijvoorbeeld

Onderwijsnummer

Het persoonsgebonden nummer of onderwijsnummer is een uniek nummer voor elke leerling die in Nederland door de overheid bekostigd onderwijs volgt. Dit nummer is gelijk aan het sofinummer en maakt het mogelijk leerlingen beter te volgen. Het levert betrouwbare gegevens op over bijvoorbeeld naleving van de leerplicht of voortijdig schoolverlaten. Voor het ontwikkelen van onderwijsbeleid is dat belangrijke informatie. Het onderwijsnummer dient ook als instrument voor het controleren van de bekostiging. Scholen zijn verplicht dit onderwijsnummer te gebruiken in hun administratie.

Meer informatie over het onderwijsnummer is te vinden op de website www.minocw.nl/onderwijsnummer.

Met het onderwijsnummer kunnen leerlingen tijdens hun hele schoolloopbaan worden gevolgd



worden gecorrigeerd voor onderwijs dat leerlingen op een andere school hebben gevolgd. Ook het functioneren van leerlingen in het vervolgonderwijs kan worden gemeten om zo een beeld te krijgen van hoe goed een school haar leerlingen op het vervolgonderwijs voorbereidt.

Een andere voorwaarde voor positieve effecten van het gebruik van toetsresultaten in accountability systemen is dat scholen over informatie beschikken over de wijze waarop zij hun onderwijsproces moeten bijsturen om tot betere prestaties te komen. Alleen een beoordeling van de leerprestaties van een school in termen van onvoldoende, voldoende of goed geeft weinig zicht op veranderingen die moeten worden doorgevoerd. De beoordeling van het onderwijsproces in de school door de onderwijsinspectie gaf hier in het verleden goede handreikingen voor. Deze informatie zal nu zelf door de school moeten worden gegenereerd, bijvoorbeeld door middel van zelfevaluaties of kwaliteitszorg waarin aandacht wordt besteed aan het verbeteren van onderwijsprocessen in relatie tot leerprestaties. De cultuur in, en de capaciteit van de school om verbeteringen op een constructieve en open manier te bespreken en door te voeren vormt daarbij een belangrijke voorwaarde.

De laatste, wellicht meest belangrijke voorwaarde voor een goed functionerend (*test-based*) accountability systeem is om zorgvuldig en terughoudend om te gaan met targets voor toetsresultaten en daaraan gerelateerde sancties. Onderwijskundig onderzoek heeft nog weinig kant-en-klare protocollen opgeleverd voor het realiseren van hoge leerprestaties in allerlei verschillende omstandigheden, met diverse leerlingen en met soms beperkte middelen. Leerprestaties van leerlingen worden bovendien door veel andere factoren verklaard waardoor het ook oneerlijk en onredelijk is om scholen hard af te rekenen op minimale toetsscores. Het Nederlandse accountability systeem waarbij een oordeel over de kwaliteit van een school tot stand komt via verschillende 'checks and balances' in het toezicht, zoals aanvullende bronnen en bij risico's een gesprek met het schoolbestuur, is dan ook een groot goed dat we minimaal moeten behouden.

□ Mw. dr. M.C.M. Ehren is werkzaam aan Universiteit Twente, Faculteit Gedragswetenschappen, Afdeling Onderwijsorganisatie en Management. Email: m.c.m.ehren@utwente.nl.

Ook de EVC
procedures van

EVC-centra **Bouw & Infra**

Zelfstandig onderdeel van



staan onder
toezicht van de

Examen
kamer