

Worden varkens zwaarder door ze te wegen?

Notities n.a.v. *The Death and Life of the Great American School System. How Testing and Choice Are Undermining Education* van Diane Ravitch

Eric Verbiest - Emeritus lector Schoolontwikkeling en schoolmanagement, Fontys Hogescholen, Nederland; gastprofessor Onderwijsinnovatie, Universiteit Antwerpen
Publicatiedatum: 15/3/2012 in: School en begeleiding: Personeel en organisatie (Wolters Plantyn, Vlaanderen)

1. SAMENVATTING

Er wordt steeds meer getest, getoetst en gemeten in het onderwijs. Maar de vraag is natuurlijk wat, hoe en waarom er wordt gemeten. En vooral wat er met die resultaten allemaal wordt gedaan. Diane Ravitch stelt in haar boek *The Death and Life of the Great American School System. How Testing and Choice Are Undermining Education* dit toenemende toetsbeleid en enkele andere vernieuwingen in het Amerikaanse onderwijs aan de kaak. De auteur van dit artikel bespreekt daarvan de belangrijkste zaken, maakt de vergelijking met de huidige trends in Nederland en Vlaanderen en licht ten slotte toe wat we hieruit kunnen leren. De rode draad door dit hele verhaal heen is de vraag: wordt het onderwijs beter van die toenemende nadruk op toetsen en meten?

2. INLEIDING

Beleidsmakers, ook in Vlaanderen en Nederland, verwijzen graag naar internationale vergelijkingen van leerlingprestaties om het eigen beleid te rechtvaardigen. De huidige Vlaamse minister van onderwijs Smet heeft de ambitie om het Vlaamse onderwijsbeleid nog beter af te stemmen op de beschikbare internationale vergelijkende gegevens (Penninckx, Vanhoof & Van Petegem, 2011). En na – in de ogen van de Nederlandse onderwijsminister van Bijsterveld – tegenvallende PISA-resultaten van Nederlandse leerlingen, kondigde ze plannen aan voor een verplichte eindtoets voor elke basisschool, voor een verplichte tussentijdse toets halverwege het voortgezet onderwijs (een verplicht eindexamen bestaat al) en voor een grotere beperking tot de kernvakken (Nederlands, Engels, wiskunde en de bètavakken) in de eerste fase van het voortgezet onderwijs. 'Een hoger ambitieniveau is onontkoombaar', aldus de minister in een interview met *De Volkskrant* (7 december 2010), 'Je ziet, ook nu weer aan PISA, dat de landen in het Verre Oosten ons voorbijstreven. We zijn in Europa, ook in Nederland, te gemakzuchtig geworden. Ik wil niet achteroverleunen.'

3. WORDEN VARKENS ZWAARDER DOOR WEGING?

3.1. Onderwijsvernieuwing in Nederland en Vlaanderen

Reeds in 2004 wees Vandenberghen op het opvallende en moeilijk te begrijpen verschijnsel dat landen met een verschillende cultuur, een verschillende traditie van onderwijsvernieuwing en een verschillende graad van economische ontwikkeling, alle grijpen naar quasi dezelfde middelen om het onderwijs te vernieuwen. Vragen of bepaalde ontwikkelingen passen bij de nationale onderwijscontext, worden nauwelijks gesteld.

New Public Management

Zo zien we de laatste decennia in het onderwijsbeleid van allerlei landen de invloed van een neoliberale besturingsfilosofie, het zogeheten *New Public Management*. Centrale waarden in dit *New Public Management* zijn resultaatgerichtheid, verhoogde efficiëntie, effectiviteit en kwaliteit van dienstverlening en klantgerichtheid. De overheid scheidt het beleid (het 'wat') en de uitvoering (het 'hoe') overeenkomstig de slogan dat de overheid niet moet roeien maar moet sturen. Belangrijke kenmerken van deze sturingsfilosofie zijn onder meer: nadruk op outputgerichte sturing ten koste van sturing op input en processen; nadruk op het meten van resultaten; introductie in de publieke sector van marktprikkels en meer invloed voor de

klanten van diensten, via vraagsturing (Pollit, 1993; Trommel, 2006).

De situatie in Nederland

Hoewel Nederland in vergelijking met veel andere landen een sterk gedecentraliseerd onderwijsbeleid heeft, kunnen we in het overheidsbeleid van de afgelopen decennia de invloed van het *New Public Management* goed herkennen. Enerzijds kregen de onderwijsinstellingen steeds meer autonomie op bepaalde terreinen, met name op personeel en financieel terrein. Anderzijds werd de greep van de overheid op de inhoud van het onderwijs sterker, bijvoorbeeld door de invoering van kerndoelen en door een stringenter inspectietoezicht, en wordt de nadruk op hogere opbrengsten op met name taal- en rekenprestaties en het toetsen ervan steeds sterker, bijvoorbeeld door de voorgenomen invoering van referentieniveaus, een verplichte eindtoets in het basisonderwijs en het sneller sluiten van zeer zwakke scholen. Via de inspectie worden de resultaten gemeten en beoordeeld en kan op de resultaten worden gestuurd. Tegelijkertijd impliceert de toegenomen verantwoordelijkheid en de toegenomen autonomie op het personele en financiële gebied dat de scholen meer en meer in een marktpositie ten opzichte van elkaar komen. Ook de aangekondigde prestatiebeloning en het toekennen van een predicaat 'excellente school' zullen de competitie tussen scholen aanwakkeren en dus de resultaten verbeteren, zo verwachten althans beleidsmakers in Nederland.

De Vlaamse situatie

Ook Vlaamse scholen kennen een grote mate van autonomie. In tegenstelling tot verschillende buurlanden kent Vlaanderen echter geen centrale examens. Wel neemt in de discussie over en in het beleid ten aanzien van de kwaliteitsbewaking van onderwijs het belang toe van externe controle en outputmeting. Het belang van eindtermen en ontwikkelingsdoelen neemt als sturend mechanisme toe en de rol van de inspectie is gegroeid. Zo werden tot 2008 de context, input en procesvariabelen voornamelijk gebruikt om de outputvariabelen te relativeren. In het huidige inspectiebeleid zijn deze outputvariabelen niet langer relatief maar absoluut. Ook is er tegenwoordig sprake van actieve (in plaats van passieve) openbaarheid (via een website) van inspectierapporten (Penninckx, Vanhoof & Van Petegem, 2011).

3.2. Het Amerikaanse schoolsysteem

Hogere opbrengsten op kernvakken, via centrale toetsen verantwoording afleggen over die opbrengsten en introduceren van marktprikkels in het onderwijssysteem zijn ook typerend voor het onderwijsbeleid zoals dat in het afgelopen decennium te zien was in de Verenigde Staten. Het is daarom de moeite waard kennis te nemen van een goed gedocumenteerde en toegankelijk geschreven evaluatie van dit beleid. Die is te vinden in een boek van Diane Ravitch: *The Death and Life of the Great American School System. How Testing and Choice Are Undermining Education*.

Diane Ravitch

Ravitch is onderwijshistorica en onderzoekt onderwijsbeleid. Hoewel ze democraat is, was ze enkele jaren (1991-1993) in het republikeinse kabinet van president Bush sr. onderminister op het onderwijsdepartement (*Assistant Secretary of Education*), waar ze leiding gaf aan het ontwikkelen van academische standaarden voor verschillende vakken. Onder het presidentschap van de democraat Clinton was ze jarenlang lid van de *National Assessment Governing Board*. Dit orgaan gaf leiding aan het federale toetsprogramma.

Onderwijsvernieuwing in de VS

Zowel de regering-Bush als de regering-Clinton wilden in de publieke sector – en dus ook in het onderwijs – de bureaucratie terugdringen en ondernemerschap bevorderen, onder meer

via deregulering en privatisering. Zoals blijkt uit haar loopbaan, was Ravitch, en met haar vele anderen, aanvankelijk enthousiast over de potentiële resultaten van toetsen, verantwoording afleggen en marktprikkels in het onderwijs. Streven naar hogere leeropbrengsten bij leerlingen, nagaan via centrale toetsen welk niveau leerlingen en scholen bereiken en scholen stimuleren via marktprikkels om het beter te doen, lijken elementen van een beleid waar geen weldenkend mens tegen kan zijn. Maar de uitwerking van dit beleid in de praktijk bleek minder positief. Het boek van Ravitch – dat past bij de toenemende kritiek op sturingsmechanismen in de publieke sector die, hoewel goed bedoeld, kunnen omslaan in perversiteit (bv. RMO, 2011) – is dan ook het resultaat van een persoonlijke intellectuele crisis van Ravitch.

Ravitch bespreekt tal van thema's, zoals het overplanten van (min of meer geslaagde) onderwijsvernieuwingen naar andere scholen en de invloed op het onderwijsbeleid van het *philanthrocapitalism* zoals de *Bill & Melinda Gates Foundation*. Op enkele thema's uit het boek gaan we dieper in.

3.3. No child left behind

De NCLB-wet

Ravitch schenkt uitgebreid aandacht aan het onderwijsbeleid in de VS van het afgelopen decennium, dat beheerst wordt door de zogeheten *No Child Left Behind Act* (NCLB) van president Bush jr. Deze wet beoogt de kwaliteit van het onderwijs te verbeteren, met name de basisvaardigheden van de leerlingen uit het primair en secundair onderwijs. Met behulp van gestandaardiseerde toetsen dienen de staten van de VS de voortgang van de onderwijsresultaten (op de *public schools*) vast te stellen. Deze voortgang wordt uitgedrukt in de *Adequate Yearly Progress* (AYP). In het schooljaar 2013-2014 moeten nagenoeg alle leerlingen een voldoende niveau halen voor taal en rekenen. Als een school onvoldoende vooruitgang toont, wordt deze een *school in need of improvement* en intervenueert het onderwijsdepartement van de desbetreffende staat in de school (zie verder Verbiest, 2011).

Kritiek op het NCLB-beleid

Ravitch maakt duidelijk dat, hoewel ze aanvankelijk dit beleid steunde, ze in toenemende mate gedesilluseerd werd. In het NCLB-beleid kwam steeds meer nadruk te liggen op het afleggen van verantwoording, op toetsing, op privatisering, op prestatiebeloning en op competitie tussen scholen. De belangrijkste remedies die NCLB voorschreef – met name vrije schoolkeuze en lessen na schooltijd – bleken niet te werken: slechts een klein aantal leerlingen maakte gebruik van deze mogelijkheden. Ouders protesteerden soms tegen het excessief toetsen van de leerlingen. Meer en meer scholen halen, naarmate 2014 nadert, niet de AYP en worden dan ook als falende scholen gezien. In 2007-2008 betrof dit al meer dan een derde van alle *public schools*. En het gevolg is dat duizenden scholen het risico lopen om gesloten of in een of andere vorm geprivatiseerd te worden. Tegelijk blijkt niet dat NCLB het onderwijs heeft verbeterd. De toetsresultaten van leerlingen op federale toetsen blijven vaak fors achter op de resultaten op de toetsen van de afzonderlijke staten – die immers hun eigen toetsen mogen kiezen. Leerlingen zijn beter geworden in het strategisch beantwoorden van meerkeuzevragen, maar hebben moeite om open vragen over een gelezen tekst te beantwoorden. Ook is de prestatiekloof tussen verschillende etnische groepen niet verkleind. In het curriculum wordt meer tijd besteed aan de vakken die getoetst worden en aan voorbereiding op de toets, ten koste van de tijd die besteed kan worden aan andere vakken.

3.4. Vrije schoolkeuze

Een ander belangrijk thema dat Ravitch behandelt, betreft de vrije schoolkeuze (leerlingen van *public schools* dienen naar scholen in hun buurt te gaan). In de loop van het afgelopen decennium stimuleerde de overheid vrije schoolkeuze, bijvoorbeeld via *voucher schools*

(privéscholen waarbij leerlingen met *vouchers* naar de school van hun keuze kunnen gaan; meestal dekten de *vouchers* slechts een deel van de kosten) en *charter schools* (ingericht door al of niet op winst gerichte organisaties of personen die van de staat een vergunning en geld uit de openbare middelen kregen). De basisstrategie was het marktmodel. Men geloofde dat door de competitie met *voucher* of *charter schools* de *public schools* zouden verbeteren.

Kritiek op de voucher schools en charter schools

Onderzoek naar de effecten van *vouchers* en van *charter schools* laat een tegenstrijdige indruk achter. Voorstanders menen positieve effecten te kunnen aantonen, tegenstanders stellen dat er geen bewijs is voor een significante verbetering van de leerlingresultaten als gevolg van *vouchers* of *charters schools*. Een aantal *charter schools* haalt uitstekende resultaten. Soms is dat te danken aan een selectie van de leerlingen. Daarnaast krijgen *charter schools* vaak additionele middelen van sponsors waardoor bijvoorbeeld kleinere klassen of extracurriculaire activiteiten mogelijk zijn. Maar hoewel overtuigend bewijs ontbreekt dat *charter schools* als zodanig betere resultaten bereiken dan *public schools*, blijkt, ook door maatregelen die genomen zijn door de regering-Obama, het aantal *charter schools* in de VS flink te groeien. Ravitch vreest dan ook dat hierdoor het systeem van *public schools* erodeert, waardoor de ongelijkheid in het onderwijssysteem toeneemt. Bovendien ondervinden ook privéscholen veel concurrentie van de (gesubsidieerde) charterscholen. De komst van *charter schools* droeg bij aan het sluiten van onder meer veel katholieke scholen die juist succesvol waren in het onderwijs aan gedepriveerde kinderen.

3.3. Accountability

Verantwoording over bereikte resultaten

Ook het onderwerp *accountability* (verantwoording afleggen) komt uitgebreid in het boek aan de orde. Beleidsmakers willen dat scholen verantwoording afleggen over de bereikte resultaten. Ook in het beleid van NCLB ligt de nadruk op toetsen en verantwoording afleggen.

Kritiek op het geïsoleerd gebruik van toetsen

Ravitch wijst hier op enkele problemen met het gebruik van toetsen. Toetsen zijn geen zeer precieze instrumenten, maar kennen altijd foutenmarges. Toetsen dienen daarom niet geïsoleerd gebruikt te worden, maar in combinatie met andere beoordelingswijzen, zoals oordelen van leerkrachten, terwijl ook rekening moet worden gehouden met zaken zoals de klasgrootte, het opleidingsniveau van leerkrachten, het curriculum, enz. Er zijn allerlei redenen – buiten de prestaties van leerlingen – waarom toetsscores in de ene school beter kunnen zijn dan in de andere school: beter onderwijs in eerdere leerjaren, een ander samenstelling van de leerlingpopulatie, meer steun van de ouders, een homogener leerling populatie, enz. Toetsen meten dus niet enkel de prestaties van leerlingen, maar ook factoren zoals motivatie en ouderbetrokkenheid.

Manipulatie van toetsen

Problemen zijn er verder als toetsscores worden gebruikt voor andere doelen dan waarvoor de toetsen bedoeld zijn, zoals de beoordeling van leerkrachten. Gegeven het belang dat gehecht wordt aan toetsscores, is het niet verwonderlijk dat er op veel manieren geprobeerd wordt om de resultaten te manipuleren. Soms door regelrecht bedrog, vaker door maatregelen zoals het beperken van het aantal laagpresterende leerlingen in de school of het uitsluiten van deze leerlingen van de toets. De meest voorkomende vorm van toetsmanipulatie is het voorbereiden van de leerlingen op de toets (*teaching to the test*), zoals veel aandacht schenken aan de onderwerpen die in de toets zullen komen en leren om in multiplechoicevragen onwaarschijnlijke alternatieven te herkennen. Deze toetstraining gaat ten koste van een breed curriculum waarin ook andere onderwerpen van belang worden

geacht dan alleen bepaalde basisvaardigheden. Ook staten manipuleren de toetsscores door bijvoorbeeld de criteria te verlagen: leerlingen kunnen daardoor met een lagere score dan vroeger toch slagen. Ravitch geeft voorbeelden van staten die hoge scores van geslaagde leerlingen melden (bv. in Texas 85 %), terwijl volgens federale normen veel minder leerlingen (29 %) zouden slagen. Bovendien is er vooral sprake van 'bestraffende verantwoording': scholen worden niet zozeer geholpen om beter te worden, maar worden vooral bestraft als de resultaten achterblijven.

3.4. De kwaliteit van de leerkrachten

Goed onderwijs = goede leerkrachten

Ravitch bespreekt verder enkele aannames – populair bij de aanhangers van NCLB – die inhouden dat de kwaliteit van het onderwijs nagenoeg uitsluitend afhankelijk is van de kwaliteit van leerkrachten en dat verbetering van het onderwijs meer effectieve leerkrachten vraagt. Markmechanismen zoals het gemakkelijker kunnen benoemen en ontslaan van leerkrachten en prestatiebeloning kunnen bijdragen aan meer effectieve leerkrachten. Ook wordt in deze gedachtengang minder waarde gehecht aan de professionaliteit van leerkrachten. Wat telt zijn leerkrachten, ongeacht hun opleiding, die de toetsresultaten verhogen. In de context is er dan ook aandacht voor *value-added assessment* – VAA (toegevoegde waarde meten). Gebruikmakend van VAA-modellen zouden de bijdragen van scholen en individuele leerkrachten aan de leerwinst van leerlingen vastgesteld kunnen worden.

Kritiek op deze gedachtengang

Ravitch herhaalt haar argument dat toetsen voor het meten van leerlingresultaten niet accuraat genoeg zijn om er consequenties aan te kunnen verbinden over de kwaliteit van een specifieke leerkracht. Er spelen veel andere factoren, zoals de thuissituatie van leerlingen. De meeste VAA-methoden controleren niet voor deze niet-schoolfactoren. Verder zijn de toetsresultaten van leerlingen over de jaren heen instabiel. In een onderzoek bleek minder dan de helft van de leerkrachten in rekenen en taal die voor hun vaste aanstelling tot de top 20 % qua leerlingresultaten behoorden, na het krijgen van die aanstelling nog tot die top 20 % te behoren. Ook het koppelen van toetsresultaten aan prestatiebeloning of ontslag van leerkrachten roept veel vragen op. Er is weinig bewijs dat prestatiebeloning tot iets anders leidt dan tot *teaching to the test* met negatieve gevolgen zoals versmalling van het curriculum.

3.5. Wat kunnen we hieruit leren?

Varkens worden niet zwaarder door ze te wegen

In het slothoofdstuk formuleert Ravitch wat er te leren valt van de onderwijshervorming van de laatste jaren. Scholen worden niet beter als er te weinig oog is voor de essentiële doelen van onderwijs, als er slechts nadruk wordt gelegd op rekenen en taal, als slechts gewaardeerd wordt wat via toetsen gemeten kan worden, als men zich uitsluitend op de toetsscores baseert om beslissingen te nemen over het lot van leerlingen, leerkrachten, schoolleiders en scholen, als *charter schools* de meest gemotiveerde leerlingen uit de *public schools* halen en men buurtscholen blijft sluiten.

Een visie over schoolverbetering

Als we iets weten over hoe we onderwijs kunnen vernieuwen, dan is het dat er geen eenvoudig antwoord is – zoals verantwoording vragen of marktprikkels introduceren in het onderwijssysteem. Schoolverbetering vraagt een visie op wat goed onderwijs is. Een focus op basisvaardigheden is belangrijk, maar niet voldoende. Ook de voorbereiding op het leven als zelfstandig denkende burger in een samenleving is essentieel. Om die visie waar te maken, is een hoogwaardig curriculum nodig. Verder dient het beoordelingssysteem aan te sluiten op dit brede curriculum en gepaste toetsvormen te bevatten, zoals essays voor

literatuur, demonstratie van denkvaardigheden in wiskunde, enz. Bovendien dienen de toetsscores niet alleen de maat te zijn voor het beoordelen van de kwaliteit van de school. Inspectieteams dienen ook het onderwijs en het leren te beoordelen. Beoordeling dient niet gericht te zijn op het sluiten van zwakke scholen, maar op het verbeteren ervan. Verbetering van het onderwijs vraagt ook goed opgeleide en ervaren leerkrachten en effectieve instructie. En omdat scholen niet geïsoleerd bestaan, vereist goed onderwijs een gemeenschap die onderwijs waardeert. Het gezin speelt een grote rol in het motiveren van leerlingen. Om dit alles te bereiken, is het toepassen van marktmechanismen niet de juiste benadering.

3.6. Terugkoppeling naar Vlaanderen en Nederland

Wat is de betekenis van dit boek voor het Vlaamse en Nederlandse onderwijs? Onmiddellijk kan worden vastgesteld dat er grote verschillen zijn in het onderwijssysteem en het onderwijsbeleid tussen de VS enerzijds en Vlaanderen en Nederland anderzijds. In de lage landen kennen we niet eenzelfde systeem van gedwongen schoolkeuze (bij *public schools*) als in de VS. We kennen geen *charter schools* of *voucher schools* en hoeven ook veel minder bang te zijn dat door onderlinge concurrentie het openbaar onderwijs verwordt tot – zoals premier Blair van Groot-Brittannië het ooit noemde – *a poor service for poor people*. Wij kennen hier ook niet de grote bemoeienis van invloedrijke private sponsors, zoals de *Gates Foundation*, die op politiek oncontroleerbare wijze het onderwijs proberen te beïnvloeden. De waarde van Ravitch' boek ligt naar mijn mening vooral in de vragen die het oproept bij vooronderstellingen, ideeën en soms ook praktijken die ook meer en meer in Nederland en, zij het in mindere mate, ook in Vlaanderen in het beleid merkbaar zijn. Het gaat daarbij vaak om onderwerpen waarbij ingrijpende beslissingen en consequenties in het geding zijn. Voor leerlingen gaat het bijvoorbeeld om al of niet blijven zitten of al of niet naar een bepaald type vervolgonderwijs kunnen gaan. Voor leerkrachten en schoolleiders gaat het onder meer om de faam van de school, inkomen en carrièremogelijkheden.

3.6.1. De toetscultuur

Het toenemende belang van toetsen

Om te beginnen kan worden gedacht aan het toenemende belang dat men hecht aan het toetsen van de leerlingresultaten. De toegenomen nadruk op het toetsen van leerlingresultaten blijkt in Nederland onder meer uit de pilot *Leerwinst en toegevoegde waarde* die vanaf januari 2012 tot september 2013 zal lopen. Doel van de pilot is om de leerwinst en toegevoegde waarde van scholen beter in beeld te brengen en zo het opbrengstgericht werken te stimuleren (OCW, 2011b). Ook in Vlaanderen wint, zoals gezegd, externe controle en outputmeting aan belang.

Kwaliteit van de toetsen

Om te beginnen kunnen vragen worden gesteld over de kwaliteit van de toetsen. Zijn de toetsen valide genoeg? Geeft, in Nederland, de beperking tot basisvakken (Nederlands, Engels, wiskunde en de bètavakken) een voldoende rijk beeld van de kennis en vaardigheden van de leerlingen om er belangrijke beslissingen aan te verbinden? En als de toetsen slechts uit meerkeuzevragen bestaan, kan men daarmee dan complexe vaardigheden zoals het beheersen van diverse oplossingsstrategieën voor rekenproblemen of de beheersing van een taal meten? Een versmalling van de toetsen tot meerkeuzevragen en het uitsluiten van het toetsen van complexe problemen kan de competenties die gemeten worden, ernstig beperken (zie bv. ook Nussbaum, 2010).

Wat wordt er precies gemeten?

Metten de toetsen inderdaad wat ze willen meten? Het is in dit verband zinvol om onderscheid te maken tussen het meten van prestaties of resultaten van leerlingen, de leerwinst van leerlingen en de door de school en leerkrachten toegevoegde waarde. Het louter meten van leerlingresultaten (zoals bij de PISA-metingen gebeurt) zegt nog niet veel

over de plaats waar die kennis en vaardigheden verworven zijn. Leerwinst van een leerling (het verschil in kennis en vaardigheden tussen twee tijdstippen) laat ook niet zien waardoor die winst bereikt is. Zo kent Nederland een sterk effect van het ouderlijk milieu op onderwijsprestaties (OECD, 2004; Dronkers, 2011). Zowel bij het meten van leerlingresultaten als bij het meten van leerwinst kan gemakkelijk en minstens deels ten onrechte de suggestie worden gewekt dat de prestaties het resultaat zijn van het gevolgde onderwijs, terwijl andere factoren ook een rol spelen. En voor wat betreft het meten van de toegevoegde waarde moet worden vastgesteld dat dit nog tal van problemen oplevert. Ook de minister erkent dit en stelt in de eerder genoemde brief (OCW, 2011b) dat het niet mogelijk is om voor alle externe invloeden te corrigeren en dat de correcties die wel worden doorgevoerd, gebaseerd zijn op gemiddelden. De toegevoegde waarde is daarom slechts een benadering van de werkelijke invloed die scholen hebben op de prestaties van hun leerlingen (zie verder bv. Baker et al., 2010; Amrein, 2008).

Neveneffecten van de toetscultuur

Ravitch wijst in haar boek op een aantal problematische neveneffecten van een toenemende toetscultuur. Deze komen vooral neer op een verarming van het curriculum en het onderwijs. Zo kan om te beginnen de vraag worden gesteld hoe we kunnen vermijden dat toetsing van een beperkt aantal onderdelen uit het curriculum, hoe belangrijk die ook zijn, ertoe leidt dat andere onderdelen als minder belangrijk worden gezien door de leerkrachten. Hoe vermijden we dat leerkrachten bijvoorbeeld meer tijd aan de te toetsen vakken gaan besteden, ten koste van het overige curriculum? Hoe kunnen we vermijden dat het voorbereiden van leerlingen op toetsen ontaardt in een toetstraining die niet alleen in tijd ten koste gaat van het overige curriculum, maar ook ertoe kan leiden dat het zo goed mogelijk leren beantwoorden van specifieke (meerkeuze)toetsvragen belangrijker wordt dan het ontwikkelen van diepgaande kennis en kritisch denken? Een ander effect van een groeiende toetscultuur, zeker als de toetsscores (bijna) uitsluitend worden gezien als het resultaat van wat de school en leerkrachten doen, is dat het alle verantwoordelijkheid voor die scores bij de school en de leerkrachten legt. Dat roept de vraag op hoe we kunnen vermijden dat leerlingen en ouders ontslagen worden van hun verantwoordelijkheid in het leerproces van de leerlingen. Positiever geformuleerd, hoe kan de samenleving bevorderen dat leerlingen en ouders hun verantwoordelijkheid opnemen en gemotiveerd zijn om leerlingen tot goede leerresultaten te brengen?

Het gebruik van toetsresultaten in het kader van schoolverbetering

Een laatste vraag die we in dit verband willen noemen, betreft het gebruik van toetsresultaten in de context van schoolverbetering. Ravitch maakt het misverstand duidelijk dat het zichtbaar maken van resultaten van scholen op zichzelf er al toe zou leiden dat zwakkere scholen geprikkeld worden om zich te verbeteren. Verbetering vraagt echter veel meer dan het benoemen van de resultaten. Dat roept dan ook de vraag op hoe we kunnen bevorderen dat het systeem van verantwoording afleggen tegelijk ook betekent dat men als school kan leren van fouten en minder goede resultaten. Dergelijke systemen van intelligente verantwoording vragen vooral veel formatieve evaluaties (in plaats van uitsluitend summatieve evaluaties) die gebruikt kunnen worden door de school zelf om het onderwijs te verbeteren (zie verder bv. Hopkins, 2007; Elmore, 2008; Verbiest, 2009).

3.6.2. De prestatiebeloning

Experimenten met prestatiebeloning in Nederland

Een tweede beleidsontwikkeling – sterk herkenbaar in Nederland maar nauwelijks of niet in Vlaanderen – die naar aanleiding van Ravitch' boek een aantal vragen oproept, betreft de invoering van prestatiebeloning. In Nederland loopt hiervoor een aantal proefprojecten. Een belangrijke overweging van de minister hierbij is dat bijzondere prestaties van leraren moeten worden benoemd en beloond om tot navolging te kunnen leiden (OCW, 2011a). Op

dit moment worden experimenten met prestatiebeloning gestart; in 2016 dient prestatiebeloning definitief te worden ingevoerd. De invoering van prestatiebeloning door de overheid illustreert overigens dat de overheid, die de mond vol heeft van autonome scholen, zich niet beperkt tot het 'wat', maar ook ingrijpt in het 'hoe'. En hoewel in Vlaanderen sommigen pleiten voor een beloningsmechanisme bij een gunstige evaluatie om leerkrachten meer gericht te motiveren, is er nog geen sprake van een dergelijke prestatiebeloning. We spitsen onze vragen daarom toe op het beleid in Nederland. Bij deze voornemens is opnieuw een aantal vragen te formuleren, geïnspireerd door het boek van Ravitch.

Twijfels en bezwaren

Om te beginnen kan de vraag worden gesteld of het wel zo verstandig is zoveel geld (ca. 144 miljoen euro over vier jaar en daarna 250 miljoen euro structureel) op de onderwijsbegroting te bestemmen voor een idee dat tot nu toe wetenschappelijk geen steun heeft. Evenals Ravitch komt ook recent onderzoek naar een langdurig experiment in New York tot de conclusie dat prestatiebeloning geen effect heeft op de resultaten van leerlingen, noch op de motivatie en het gedrag van leerkrachten. De onderzoekers concluderen verder dat, gezien het groeiende bewijs dat prestatiebeloning niet helpt, de achterliggende aannames over prestatiebeloning in het onderwijs aan herziening toe zijn (Marsh et al., 2011). Verder mag men verwachten dat prestatiebeloning op duurzame resultaten van leerlingen en leerkrachten is gebaseerd en niet op toetsresultaten op zich. Tegen die achtergrond kan ook de vraag worden geformuleerd of de voorgenomen experimenten lang genoeg duren om iets onderbouwd te kunnen zeggen over het effect van prestatiebeloning over duurzame resultaten. Het lijkt bovendien redelijk om prestatiebeloning, zo men die wil invoeren, te baseren op de toegevoegde waarde die een school of leerkracht, heeft bereikt. Zoals gezegd is het vaststellen van die toegevoegde waarde niet eenvoudig. Het gaat daarbij ook nog om kansuitspraken, terwijl de beslissingen die erop gebaseerd zijn, ingrijpend zijn. Onderzoekers waarschuwen dan ook tegen het gebruiken van toetsscores, zelfs als die via hoogwaardige methoden zijn verzameld, om beslissingen te onderbouwen voor zaken zoals prestatiebeloning (Baker et al., 2010; Onderwijsraad, 2011).

3.6.3. Onderliggende aannames

Het boek van Ravitch is ook waardevol omdat het stil staat bij de onderliggende aannames waarop enkele van de genoemde beleidsontwikkelingen gebaseerd zijn. Tot slot van deze bijdrage staan we stil bij enkele van deze aannames.

Vooronderstellingen en misverstanden

Een eerste aanname – die Ravitch de wellicht meest naïeve noemt – is hiervoor al uitgebreider besproken, namelijk dat goed onderwijs hetzelfde is als hoge resultaten op gestandaardiseerde toetsen voor basisvakken. Verder wijst Ravitch op het hardnekkige misverstand bij veel onderwijshervormers dat een goed idee of een goede praktijk eenvoudig overgeplant kan worden naar andere scholen, net zoals men een industrieel proces of een nieuwe machine naar een ander bedrijf kan overplanten. Als scholen succesvol zijn, dan spelen daar veel met elkaar interacterende factoren een rol in: de kwaliteiten van het leiderschap en de andere professionals, de cultuur in de school, de leerlingen en hun ouders, de kwaliteit van het curriculum en van de instructie, de materiële mogelijkheden waarover de school beschikt en nog andere. Dat misverstand bleek onder meer in de analyse van de (mislukte) pogingen om het onderwijs in enkele steden, zoals New York en San Diego, te modelleren naar een eerder geslaagde onderwijsverbetering. Ook de ervaringen met de *charter schools* laten zien dat oorspronkelijk goede ideeën en bedoelingen gemakkelijk verworden tot ineffectieve en zelfs contraproductieve praktijken.

Marktmechanismen

Ravitch wijst er ook op dat een aantal hervormingen, met name zoals beschreven in het NCLB-beleid, gebaseerd is op de (onhoudbare) overtuiging dat schoolverbetering bevorderd kan worden door het bekendmaken van toetsresultaten, door veranderingen in het bestuur en door het blameren van scholen die falen. Bovendien zouden lage resultaten van leerlingen veroorzaakt worden door luie schoolleiders en leerkrachten die men moet dreigen hun job af te pakken of met bonussen moet stimuleren. Veel van deze overtuigingen gaan terug op een sterk geloof in de werking van de markt. De aantrekkingskracht van de markt is het idee dat vrijheid (geen bemoeienis van de overheid) een oplossing op zich is. Maar, aldus Ravitch, de markt is niet geschikt voor de verspreiding van hoogwaardige diensten voor allen, ongeacht waar men woont, wat men kan betalen en welke politieke invloed men heeft. De markt kent winnaars en verliezers. Aan verliezers heeft het onderwijs geen behoefte.

4. VERDERE STUDIE

- Amrein-Beardsley, A. (2008). Methodological Concerns About the Education Value-Added Assessment System. *Educational Researcher*, vol. 37, nr. 2, 65-75.
- Baker, E., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., Ravitch, D. & Shepard, A. (2010). *Problems with the Use of Student Test Scores to Evaluate Teachers*. Briefing paper 278. Economic policy institute.
- Dronkers, J. (2011). De maatschappelijke en wetenschappelijke waarde van internationale data over onderwijsprestaties. *Pedagogische Studiën* 88, 122-131.
- Elmore, R. (2008). Leadership as the practice of improvement. In: Pont, B., Nusche, D. & Hopkins, D. *Improving School Leadership. Volume 2: Case studies on system leadership* (p. 37-68). Parijs: OECD.
- Hopkins, D. (2007). *Every school a great school. Realizing the potential of system leadership*. Maidenhead: Open University Press.
- Marsh, J. (2011). *A Big Apple for Educators. New York City's Experiment with Schoolwide Performance Bonuses. Final Evaluation report*. Santa Monica: Rand Corporation.
- Nussbaum, M. (2010). *Not for profit. Why Democracy Needs the Humanities*. Princeton / Oxford: Princeton University Press.
- OCW (2011a). *Actieplan Leraar 2020 – Een krachtig beroep*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- OCW (2011b). Brief aan de Tweede Kamer (25 november 2011) betreffende de pilot leerwinst en toegevoegde waarde in het primair onderwijs. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschappen.
- Onderwijsraad (2011). *Excellente leraren als inspirerend voorbeeld*. Den Haag: Onderwijsraad.
- Penninckx, M., Vanhoof, J. & Van Petegem, P. (2011). *Evaluatie in het Vlaamse onderwijs. Beleid en praktijk van leerling tot overheid*. Antwerpen / Apeldoorn: Garant.
- Pollit, C. (1993). *Managerialism and the public services*. Oxford: Basil Blackwell.
- Raad voor Maatschappelijke Ontwikkeling (2011). *Tegenkracht Organiseren. Lessen uit de kredietcrisis*. Den Haag: Raad voor Maatschappelijke Ontwikkeling.
- Ravitch, D. (2010): *The Death and Life of the Great American School System. How Testing and Choice Are Undermining Education*. New York: Basic Books.
- Trommel, (2006). NPM en de wedergeboorte van het professionele ideaal. *Beleid en Maatschappij*, (33) 3, 137-147.
- Vandenberghe, R. (2004). Over stuurbaarheid van het onderwijs. Een analyse van 'sturend beleid', resultaten en niet-bedoelde effecten. In: Kelchtermans, G. (red.). *De stuurbaarheid van het onderwijs. Tussen kunnen en willen, mogen en moeten* (p. 89-120). Leuven: Universitaire Pers.
- Verbiest, E. (2009). *Leren leiden. Notities over de professionalisering van schoolleiders in tijden van nieuwe professionaliteit*. Eindhoven: Fontys Hogescholen.
- Verbiest, E. (2011): *Leren innoveren. Een inleiding in de onderwijsinnovatie*. Antwerpen / Apeldoorn: Garant.