

Het betere is de vijand van het goede. Een reactie op de commentaren over het *Trouw* rapportcijfer

J. Dronkers, Universiteit van Amsterdam

In deze reactie ga ik in op de commentaren van Veenstra, Dijkstra, Peschar en Snijders en van Van Putten en Van der Kamp, elders in dit nummer afgedrukt. Ik probeer in een algemeen betoog hun bezwaren te bespreken, waarbij ik niet steeds naar hun commentaren verwijs. Hoewel ik hieronder tegen onderdelen van hun commentaar bezwaren zal inbrengen, waardeer ik hun collegiaal commentaar erg.

1 Twee bezwaren tegen het openbaarmaking van inspectiegegevens

Op 5 september vorig jaar besloot de rechter dat de onderwijsinspectie verplicht was de door haar verzamelde gegevens over het intern rendement openbaar te maken. Onmiddellijk kondigde *Trouw* deze rendementsgegevens in enige vorm te zullen publiceren. Aan een dergelijk publicatie kleven twee bezwaren, die door *Trouw* in begeleidende artikelen in september werden besproken (Agerbeek, 1997) en waarop ik zelf vooraf ook heb gewezen (Dronkers, 1997).

Allereerst zouden scholen alleen geordend worden op een smalle indicator voor rendement, bijvoorbeeld het gemiddelde eindexamencijfer (zoals in Frankrijk en Engeland gebeurd). Een krant heeft geen ander instrument dan een bestaande indicator te kiezen en neemt dan uiteraard de belangrijkste. Maar met eindexamencijfers als smalle indicator dreigt men selectie tijdens de school en aan de poort te belonen. Dat ook *Trouw* een dergelijke smalle indicator voor rendement zou gaan gebrui-

ken, blijkt nog uit de tabellen in de *Trouw* publicatie waarin alle scholen geordend zijn naar het percentage onvertraagd geslaagden (percentage slaggers voor eindexamen minus percentage zittenblijvers in voorexamen klas).

Ten tweede zou geen rekening worden gehouden met de verschillen in cultureel kapitaal van leerlingen van verschillende scholen. In dat geval wordt een school met veel leerlingen met weinig cultureel kapitaal (bijvoorbeeld door het opleidingsniveau van hun ouders), die er toch in slaagt zijn leerlingen gemiddelde schoolprestaties te laten halen, als even goed beoordeeld als een school met veel leerlingen met veel cultureel kapitaal, die dezelfde gemiddelde schoolprestaties behalen. Toch is de eerstgenoemde school duidelijk beter dan de tweede school, omdat zijn 'toegevoegde waarde' groter is. Als men geen rekening houdt met deze verschillen tussen leerlingen van verschillende scholen, beloont men selectie aan de schoolpoort en vergroot onnodig het verschil tussen cultureel 'arme' en 'rijke' scholen. Om die reden had *Trouw* het enig beschikbare achtergrondgegeven voor het voortgezet onderwijs bij OC&W opgevraagd en in de tabellen in de *Trouw* publicatie opgenomen.: het percentage cumi-leerlingen.

Deze twee bezwaren zijn volstrekt niet origineel, maar zijn door Engelse, Schotse en Franse collega's geregeld naar voren gebracht tegen publicatie van eindexamenresultaten van middelbare scholen in die landen. Het Franse ministerie van Onderwijs is in zijn publicatie van schoolresultaten ook gedeeltelijk aan deze bezwaren tegemoet gekomen, in tegenstelling tot de Engelse autoriteiten. De opvatting dat het bij de kwaliteit van scholen om 'toegevoegde waarde' gaat, is ook gemeengoed onder Nederlandse onderwijskundigen (Dekkers, 1997, p. 299). Het ontwikkelen van een goede maat voor schoolsucces, beter dan alleen eindexamenresultaten, is in de onderwijswetenschappen een bekend thema: de leerjarenladder die Veenstra et al. toepassen is daarvan een voorbeeld. Kortom, beide bezwaren behoren tot de geaccepteerde inzichten in de onderwijswetenschappen.

Het rapportcijfer dat ik voor *Trouw* heb berekend, probeert deze twee bezwaren te onderwerpen. Met dat rapportcijfer heb ik het totaal

intern rendement van een school, rekeninghoudend met de leerlingenmerken, willen van- gen. Tijd voor een nationaal debat over de vast- stelling van kwaliteitsverschillen van scholen was er niet. Lang voordat dat debat van de grond zou kunnen komen, zou *Trouw* rende- mentscijfers publiceren. Het ministerie van OC&W en de inspectie hadden een dergelijk debat jarenlang ontlopen, en waren dus daarop door hun eigen nalatigheid absoluut niet voor- bereid. De gang van zaken rond de kwaliteits- kaart, die nu door de inspectie ontwikkeld wordt, illustreert deze onvoorbereidheid treffend.

Daardoor moesten er bij de vaststelling van dit rapportcijfer lastige en ook arbitraire beslis- singen genomen worden, die in de commenta- ren terecht gesignaleerd worden. Maar in deze situatie is het beter lastige en arbitraire beslis- singen te nemen om zo deze twee bezwaren alvast gedeeltelijk te ondervangen, dan zich afzijdig te houden in zijn ivoren toren. Weten- schappers hebben naast hun wetenschappelijke verantwoordelijkheid ook een maatschappelij- ke verantwoordelijkheid. Dat laatste betekent ook dat wetenschappers bereid moeten zijn geaccepteerde wetenschappelijke inzichten toe te passen op maatschappelijke vraagstukken, juist als dat maatschappelijk nog niet geaccep- teerd is. Als wij met die toepassing wachten tot die wetenschappelijke toepassing maatschap- pelijk geaccepteerd is en door het beleid wordt toegestaan, zijn wij technocraten geworden. Als wetenschappers hun inzichten niet willen toepassen uit angst voor het maatschappelijk rumoer dat voortvloeit uit die omstreden toe- passing, laten wij onze maatschappelijke ver- antwoordelijkheid aan derden over: die moeten maar zorgen dat onze wetenschappelijke inzichten worden toegepast. Bij een dergelijke afzijdigheid van wetenschappers zou na publi- catie van ruwe rendementscijfers van scholen voor *Pedagogische Studiën* een commentaar kunnen verschijnen, waarin de schrijver zou klagen dat 'men' niets met wetenschappelijke inzichten doet (de huidige situatie van onze Engelse en Schotse collega's). Het berekenen van het rapportcijfer is daarom geen sociaal- wetenschappelijk onderzoek, maar maatschap- pelijke dienstverlening met behulp van weten- schappelijke inzichten. Van Putten en Van der Kamp zien dit aspect over het hoofd in hun

commentaar. Uiteraard kan men betwisten of ik het goede evenwicht tussen maatschappelij- ke dienstverlening en toepassing van weten- schappelijke inzichten gevonden heb. Terug- kijkend zou ik een aantal zaken anders gedaan hebben (de berekening gelijk erbij afdrukken in plaats van 14 dagen later; nog duidelijker maken dat het rapportcijfer niet rechtstreeks uit de eerste vier kolommen was af te leiden; de rapportcijfers wellicht omzetten in plussen en minnen; het nog beter uitleggen dat het percen- tage cumi-leerlingen een goede proxy is van het intellectuele instroomniveau van de leerlin- gen; nog meer de matige betrouwbaarheid van de inspectiegegevens benadrukken), maar die tekortkomingen in de *Trouw* publicatie doen niet af aan het belang van mijn werkwijze: een brede meting van het intern rendement van scholen en het gebruik van 'toegevoegde waar- de'.

2 De kwaliteit van de inspectiegegevens

Bij de berekening van het rapportcijfer voor *Trouw* heb ik inspectiegegevens gebruikt. In het commentaar van Van Putten en Van der Kamp lijkt dat een bijkomstigheid, alsof ik die inspectiegegevens de meest geschikte data voor de berekening van die rapportcijfers vind. Het tegengestelde is echter het geval. Deze gegevens over scholen worden door de inspec- tie verzameld in het kader van de wettelijke taak van de inspectie om toe te zien op de deug- delijkheidseisen (een kernbegrip uit het grond- wetsartikel over onderwijs). Dat zijn de enige gegevens die systematisch over alle scholen verzameld worden met het oog op dat toezicht op de kwaliteit van scholen. Om die reden heeft *Trouw* om openbaarmaking gevraagd en meen- de de rechter dat ook ouders en belastingbeta- lers recht hebben op die officiële gegevens. Dat vervolgens een aantal gebreken aan deze inspectiegegevens bleken te kleven, is geen relevant bezwaar tegen het berekenen van een rapportcijfer, maar dient de inspectie en het ministerie van OC&W aangerekend te worden: blijkbaar vond men daar deugdelijkheid van het onderwijs niet belangrijk genoeg. Uiteraard zijn er wel andere en betere gegevens over scholen en hun toegevoegde waarde bekend.

Het school-effectiviteitsonderzoek werkt daarmee, maar de daarbij gebruikte gegevens (bijvoorbeeld het VOCL'89 cohortonderzoek dat Veenstra et al. gebruiken) en de daarmee vastgestelde 'toegevoegde waarde' mogen van de scholen, het ministerie van OC&W en het CBS niet op schoolnaam gepubliceerd worden. Kortom, het ging in de *Trouw* publicatie niet om de beste wetenschappelijke gegevens om 'toegevoegde waarde' van scholen vast te stellen, maar om de officiële gegevens, verzameld met het oog op de wettelijke taak van de inspectie en die als enige daarvoor openlijk gebruikt mochten worden dankzij de rechter.

Dat scholen en inspectie wellicht deze wettelijke taak bij de invulling niet zo serieus namen, doet niet af aan het belang van deze gegevens. Daarom kan men niet beweren dat de inspectiegegevens voor een ander doel verzameld zouden zijn dan het toezichthouden op scholen en de gegevens daarom niet gebruikt hadden mogen worden. Dat andere doel komt in de wet niet voor en de onderwijsinspectie is nog steeds geen schoolbegeleidingsdienst of een onderzoeksinstituut. Het verweer van de inspectie tegen dit publieke gebruik, dat deze cijfers allen bruikbaar zouden zijn nadat zij door de inspectie geïnterpreteerd zijn in het kader van de lokale context na schoolbezoek, zal diegenen die de vroegere inspectierapporten heeft mogen lezen niet overtuigen. Daarom is het in deze situatie niet de taak van *Trouw* of van mij de kwaliteit van deze inspectiegegevens vooraf te controleren, alsof het bij het leveren van gegevens door scholen aan de inspectie zou gaan om een wetenschappelijke enquête onder schooldirecteuren door een onderzoeksinstituut. Bovendien heeft de staatssecretaris kort na de gerechtelijke uitspraak in september 1997 alle scholen gewaarschuwd dat de inspectie gedwongen was haar gegevens aan *Trouw* te geven en dat scholen er verstandig aan zouden doen eventuele fouten in hun opgave te verbeteren. Slechts één school heeft van deze gelegenheid gebruik gemaakt. Ook is in de landelijke pers én in vakbladen in de periode tussen de rechtelijke uitspraak en de publicatie (bijna 2 maanden) aandacht besteed aan de op handen zijnde openbaarmaking. Het bezwaar van Veenstra et al. dat scholen niet wisten dat hun opgave openbaar zou worden, is dus onjuist. Dat scholen deze openbaarmaking

vooraf niet serieus namen, valt *Trouw* of mij niet te verwijten. Geen enkele schoolleider of onderwijsorganisatie heeft *Trouw* of mij vooraf benaderd om nader overleg over de publicatie van de inspectiegegevens.

Gegeven deze geringe belangstelling is het vooraf overleggen van het berekende rapportcijfer door *Trouw* of mij aan schooldirecties een naïef idee. Dit houdt onvoldoende rekening met de dynamiek tussen de eisen van journalistiek en wetenschap, zoals Veenstra et al. terecht opmerken. Maar zij zien niet de onvermijdelijkheid van de door *Trouw* en mij gevolgde werkwijze. Bij de *Trouw* publicatie gaat het om ruim 700 scholen in plaats van een klein aantal landen bij het OECD-onderzoek naar onderwijsindicatoren, waarnaar Veenstra et al. verwijzen. Een groot aantal scholen en bijna alle onderwijsorganisaties hadden zich direct of indirect jarenlang met hand en tand verzet tegen elke openbaarmaking, in tegenstelling tot die OECD-landen die vrijwillig meededen aan deze vergelijking. Kortom, de Ausgangssituatie in beide gevallen verschilde drastisch. Het vooraf overleggen met schoolleiders of onderwijsorganisaties was daarom geen haalbare kaart. Dat zoveel mensen na de *Trouw* publicatie plotseling bekeerd zijn tot publiceren van schoolresultaten, zie ik eerder als een positief resultaat van de door ons gevolgde werkwijze dan als een indicator van hun medewerking vooraf.

Gegeven dit wettelijk karakter van de inspectiecijfers en de geboden mogelijkheden voor correctie was het ook niet onze taak om onwaarschijnlijkheden in deze gegevens op te sporen. Als de inspectie bij een school in Harreveld alle leerlingen van een jeugdgevangenis bij elkaar optelt, waardoor het intern rendement van die school heel laag wordt, maakt de inspectie een fout. Zij had de twee groepen moeten onderscheiden, juist om het intern rendement van die school goed te kunnen beoordelen. Het is immers niet zeker dat dat lage rendement alleen veroorzaakt wordt door de leerlingen uit de jeugdgevangenis. Dat zelfde geldt ook voor scholen waarbij jeugdige asielzoekers door inspectie én schoolleiding gewoon bij elkaar werden opgeteld, zoals bij Da Vinci College in Leiden gebeurde. Pas nadat deze school buiten de inspectie om deze leerlingen had uitgesplitst, heb ik uitgerekend

welk cijfer die school zonder asielzoekers zou hebben. Deze herberekening was geen correctie van het eerder gegeven rapportcijfer, maar alleen een illustratie tijdens een debat op de Twaalfde Onderwijssociologische Conferentie van mijn problemen met de inspectiegegevens. Dat het gymnasium van het Marnix College te Ede heel anders scoort dan van de VWO afdeling van dat zelfde college, behoeft geen toevalligheid te zijn: ook als leerlingen bij elkaar in de klas zitten en dezelfde lessen volgen, kunnen zij in de praktijk door school en leerkrachten heel anders behandeld worden (verborgen curriculum). Ook dat is dus geen reden om inspectiegegevens zomaar als onbetrouwbaar terzijde te schuiven.

Van Putten en Van der Kamp hebben bij de toelichting bij de berekening van de rapportcijfers in de *Trouw* over het hoofd gezien dat scholen met veel missende waarden ('ernstige hiaten') geen rapportcijfer hebben gekregen, ja zelfs niet in de lijst voorkomen (bij sommige scholen tot hun woede). Zij concluderen ten onrechte uit het door mij aan hen ter beschikking gestelde databestand dat alle daarin voorkomende scholen ook in de *Trouw* publicatie een rapportcijfer hebben gekregen. Maar alleen scholen die voldoende geldige waarden hadden, komen dus in de *Trouw* publicatie voor. Hun begrijpelijke bezwaren op dit punt snijden dus geen hout.

De inspectie heeft in de loop der jaren, zonder twijfel onder druk van OC&W, een 'beleidsrijke' invulling gegeven aan sommige begrippen. Zo werd een overgang naar een ander onderwijstype binnen een scholengemeenschap waarbij sprake was van een niet-succesvolle overgang (bijvoorbeeld van 3VWO naar 4HAVO) niet als zittenblijven beschouwd, terwijl dezelfde overgang buiten een scholengemeenschap wel als ongediplomeerde uitstroom werd bestempeld. Het zal duidelijk zijn dat dergelijke beleidsrijke invullingen brede scholengemeenschappen kunstmatig in een beter daglicht plaatsen dan categoriale scholen. Maar het is niet aan *Trouw* of aan mij om deze beleidsrijke invullingen te corrigeren bij publicatie van de inspectiegegevens. Toezicht op de kwaliteit van OC&W en de inspectie is een taak van het parlement. Overigens is het interessant op te merken dat deze beleidsrijke invullingen reeds geruime tijd

gebruikelijk waren bij de inspectie zonder dat categoriale scholen de moeite namen daartegen te protesteren of zonder dat in het parlement daarnaar gevraagd werd.

Een ernstig gebrek aan de gepubliceerde inspectiegegevens is dat zij het rendement van één jaar betreffen. Het is dus mogelijk dat een school in een ander jaar een hoger of lager rapportcijfer krijgt. Meerdere jaren inspectiegegevens werden echter niet aan *Trouw* beschikbaar gesteld. In de discussie rond de rapportcijfers wordt dit probleem echter overdreven, omdat het rapportcijfer niet gebaseerd is op één eindexamenklas, maar op alle klassen (met uitzondering van de brugklas) in dat jaar. Als de eindexamenklas toevallig een 'dipje' had, waarom had een andere klas dan niet even toevallig een 'upje'? Als dat 'dipje' in meerdere klassen voorkomt zonder compensatie door een 'upje' (bijvoorbeeld door een moeizaam fusieproces), dan is er geen sprake meer van toeval maar geeft een lager rapportcijfer terecht een geringer intern rendement weer. En het is in die situatie nog maar de vraag of die 'dipjes' het volgend schooljaar zomaar verdwenen zullen zijn of door zoveel 'upjes' gecompenseerd zullen worden.

De inspectiegegevens hadden alleen betrekking op het intern rendement van de scholen: in-, door- en uitstroom, inclusief eindexamencijfers. Meer was er niet en heeft de inspectie nooit systematisch behoeven te verzamelen. Dus geen gegevens over het extern rendement van scholen (het succes van leerlingen in de vervolgopleiding of op de arbeidsmarkt). Dus geen gegevens over de culturele participatie van (ex-)leerlingen. Dus geen gegevens over de mate van sociale integratie van (ex-)leerlingen in de samenleving. Als de scholen, de onderwijsorganisaties en de politiek dit echt beschouwen als belangrijke andere dimensies van kwaliteit van scholen, moet zo spoedig mogelijk met de verzameling daarvan begonnen worden, al dan niet door de inspectie. Maar Van Putten en Van der Kamp hebben ongelijk als zij zeggen dat ik maar een klein deel van de beschikbare inspectiegegevens gebruikt heb. Het ontbreken van deze andere indicatoren van kwaliteit van scholen hebben scholen, onderwijsorganisaties en politici bij hun kritiek op de

Trouw publicatie genegeerd, terwijl zij zelf direct of indirect verantwoordelijk zijn voor dat ontbreken van die andere indicatoren. Het lijkt mij, gegeven mijn eerdere publicaties, vanzelfsprekend dat ook ik van oordeel ben dat kwaliteit van scholen meerdere dimensies heeft. Maar politiek en onderwijsorganisaties hebben het blijkbaar nooit de moeite waard gevonden over die belangrijke andere kwaliteit dimensies gegevens te laten verzamelen. Als inspectiegegevens bij openbaarmaking maar één dimensie van kwaliteit blijken te bestrijken, moet men dat niet de openbaarmakers of de inspectie verwijten, maar is dat de schuld van politiek en onderwijsorganisaties. Door deze openbaarmaking wordt iedereen alvast gedwongen serieus na te denken over die dimensies van kwaliteit van scholen. Maar voorlopig hebben politiek en onderwijsorganisaties door hun eigen beleid niet meer te bieden dan het intern rendement als kwaliteitsmeting. De andere dimensies zijn blijkbaar alleen voor feestredes, krokodillentranen en politieke retoriek.

3 De vaststelling van het totaal intern rendement

Uit het voorafgaande volgt dat ik geïnteresseerd was in een totaal meting van het intern rendement van scholen. Een dergelijke totaal meting voorkwam dat ik mij zou moeten beperken tot eindexamencijfers en bood mij bovendien de mogelijkheid dit gemakkelijk te corrigeren voor de kenmerken van de leerlingen. Ik heb een goed totaal intern rendement arbitrair gedefinieerd als een school met hoge eindexamencijfers, veel geslaagden, weinig zittenblijvers en weinig uitstroom en onder een laag totaal intern rendement heb ik het tegengestelde hiervan verstaan. Ik heb voor deze operationalisatie gekozen, omdat ik veronderstel dat ouders dat graag weten: op welke school blijven er weinig leerlingen zitten, stromen er weinig ongediplomeerd uit, wordt er veel geslaagd met hoge eindexamencijfers. In tegenstelling tot wat Van Putten en Van der Kamp menen, was ik niet bezig met een onderzoek naar verschillende dimensies binnen het intern rendement van scholen. Niet de factoranalyse was mijn uitgangspunt, maar mijn a priori gedefinieerd totaal intern rendement. Ik paste factor-

analyse (PCA) in eerste instantie alleen toe om te zien of een dergelijke dimensie überhaupt bestond. Dat bleek bij alle onderwijstypen het geval, of als eerste of als tweede factor. Dat betekent dat de door mij arbitrair gedefinieerde intern rendementsmaat een reële dimensie in de data is. Dat het niet de enige factor of in enkele gevallen zelfs niet de belangrijkste is, is in de geschetste context niet relevant. Wel ben ik met Van der Putten en Van der Kamp eens dat het de moeite waard is die andere dimensies aan een nader onderzoek te onderwerpen.

Doordat de onderwijsorganisaties en de politiek elke publieke discussie daarover jarenlang tegen hebben gehouden, zijn er geen duidelijke maatstaven voor het voldoende of onvoldoende functioneren van scholen op het punt van hun intern rendement. Ook zijn er nergens indicaties wat men belangrijker vindt bij het intern rendement van een school: zittenblijven, ongediplomeerd schoolverlaten, slagen of hoge eindexamencijfers. Daardoor zijn er geen duidelijke indicaties wat bij een totaal meting van het intern rendement het zwaarste zou behoren te wegen. Daarom heb ik in tweede instantie met behulp van dezelfde factoranalyse (PCA) de gewichten van de verschillende indicatoren berekend, zodanig dat scholen op de door mij gekozen intern rendementsmaatstaf zo goed mogelijk onderscheiden zouden kunnen worden. Daardoor krijgen de verschillende indicatoren soms uiteenlopende gewichten. Maar de berekende gewichten van de indicatoren zijn op dit moment de beste om scholen te ordenen ten aanzien van hun totale interne rendement. Uiteraard had ik ook op grond van mijn particuliere inzichten alle indicatoren van intern rendement een gewicht kunnen toekennen. Maar dat zou nog veel meer arbitrair zijn dan de thans gebruikte gewichten, die alvast scholen zo veel mogelijk statistisch onderscheiden ten aanzien van hun totale interne rendement.

Deze gewichten zijn afzonderlijk per onderwijstype berekend. Dat heeft niets met een data-driven aanpak te maken, zoals Van Putten en Van der Kamp ten onrechte stellen. Het erkennen van verschillen tussen onderwijstypen en dus ook van verschillen in gewicht die de verschillende indicatoren voor het intern rendement hebben, is allereerst inhoudelijk gefundeerd. De wens van Van Putten en Van

der Kamp om een uniforme meting voor verschillende onderwijstypen heeft eenzelfde gebrek als die mij wordt verweten, namelijk dat niet-theoretisch gefundeerde gewichten de uiteindelijke score bepalen.

Engels en Nederlands zijn vakken die door bijna alle leerlingen gevolgd worden; school-specifieke selectiviteit kan hierbij nauwelijks een rol spelen. Als wij echter alleen deze vakken in de berekening hadden betrokken, zou men ons van een talen-bias beschuldigd hebben. Daarom hebben wij het meest voorkomende bèta-vak daaraan toegevoegd: Wiskunde A. Hierbij kan school-specifieke selectiviteit inderdaad optreden, maar ik vind dat gevaar geringer dan een talen-bias. Maar in de toekomst zal ook de verdeling van leerlingen over de verschillende examenvakken bij de bepaling van het intern rendement betrokken moeten worden. Dat zelfde geldt voor de verdeling van MAVO en VBO leerlingen over enerzijds A en B niveau en anderzijds C en D niveau. Het betrekken van verdelingen van leerlingen tussen vakken en over niveaus zal overigens de inzichtelijkheid van het rapportcijfer niet vergroten. Hier is een lastige afweging tussen een zo compleet mogelijke meting en een zo toegankelijk mogelijke meting. Het commentaar van Van Putten en Van der Kamp lijkt op dit lastige punt niet te kunnen kiezen, omdat zij in hun commentaar met tegenstrijdige wensen komen.

Een betwistbaar punt in mijn meting van totaal intern rendement is de opname van zittenblijven als indicator daarvan. Men kan van mening zijn dat zittenblijven geen onderdeel uitmaakt van het intern rendement: het gaat er immers om dat leerlingen het diploma halen en niet hoelang zij daarover doen. Maar dit argument geldt niet als men daarbij verwijst naar scholen die veel zwakke leerlingen zouden toelaten en deze leerlingen via veel zittenblijven toch tot de eindstreep weten te brengen. In de door mij voorgestane aanpak wordt deze school via de berekening van de 'toegevoegde waarde' reeds gecompenseerd voor deze zwakke leerlingen en dus voor het hogere percentage zittenblijvers.

4 De correctie voor leerling kenmerken

Om rekening te houden met de cultureel kapitaal van de leerlingen is nu nog alleen het percentage cumi-leerlingen per school beschikbaar. Daarom heb ik het percentage cumi-leerlingen gebruikt als indicator van die cultureel kapitaal. Dat percentage bleek ook voor het VWO en HAVO een redelijke indicator als gevolg van de 'vlucht' van leerlingen met hooggeschoolde ouders uit scholen met veel leerlingen met laaggeschoolde ouders. Hooggeschoolde ouders geven immers de voorkeur aan smalle scholengemeenschappen van VWO en HAVO met weinig cumi-leerlingen. Omdat die vlucht veel gemakkelijker is voor het VWO en HAVO dan voor MAVO en VBO blijkt het percentage cumi-leerlingen binnen het VWO en HAVO zelfs een sterkere indicator voor verschillen tussen scholen in de cultureel kapitaal van hun leerlingen dan binnen MAVO en VBO. Dat is geen idee van mij, maar de neerslag van het feitelijk gedrag van hooggeschoolde ouders bij de schoolkeuze voor hun kinderen. Met behulp van gegevens van het CBS over schoolloopbanen in het voortgezet onderwijs is dat voor iedereen vast te stellen, zoals Veenstra et al. ook laten zien in hun analyse.

Bewust heb ik afgezien van een correctie van het rapportcijfer voor andere factoren. Het percentage cumi-leerlingen dekt ruimschoots de grote stadsproblematiek, ja heeft daarvoor in vergelijking met de Noordelijke provincies verschillen in leerlingkenmerken enigszins overgecorrigeerd (in het Noorden zijn relatief veel laaggeschoolde autochtonen en weinig allochtonen). Ook voor provincie heb ik bewust niet gecompenseerd; wij hebben in Nederland een nationaal onderwijsstelsel waardoor vergelijking op een nationale grondslag het meest correct is. Dat een nationale vergelijking negatief uitpakt voor de Randstad, kan niemand verrassen die de literatuur bijhoudt.

Uiteraard was het beter geweest als de rapportcijfers niet aan een school als bestuurlijke eenheid toegekend werden, maar aan de afzonderlijke locaties of vestigingen binnen die bestuurlijke eenheden. De inspectiegegevens lieten dat door hun ontoereikendheid niet toe. Overigens werpt het benadrukken van het belang van de afzonderlijke locaties en vesti-

gingen binnen de bestuurlijke eenheid van de school een schril licht op de werkelijkheids-waarde van veel fusies in het voortgezet onderwijs.

Tenslotte behoeft het geen betoog dat ik liever met een betere indicator van de culturele bagage van leerlingen had gewerkt, met name de Cito-score van de instroom. Maar de analyse van Veenstra et al. heeft laten zien dat deze betere indicator niet veel zou hebben uitmaakt voor het rapportcijfer: de kracht van het gebruik van een betere indicator ligt dus vooral in zijn legitimiteit. Als men echter door het niet-beschikbaar zijn van de beste indicator iedere correctie voor leerlingkenmerken achterwege laat, maakt men het betere de vijand van het goede.

5 Het rapportcijfer

In Engeland en Frankrijk worden lijsten gemaakt waarop scholen geordend worden. Dat betekent in de Nederlandse situatie een ordening van scholen op een lijst van 1 tot 700. Om te voorkomen dat er een debat zou ontstaan over het verschil tussen de 200ste en 250ste plaats heb ik gekozen voor gehele cijfers tussen de 1 en de 10, met de 6 als gemiddelde. Dit staat keurig in de toelichting bij de betrokken *Trouw* publicatie vermeld. Dat lijkt mij nog steeds de meest zuivere aanpak. De suggestie dat net als de leerlingen van Posthumus 75% van de scholen wel voldoende functioneert lijkt mij een nog onoverzichtelijker slag in de lucht. Het vervangen van de 6 door een nul, de zeven (de afstand van een standaarddeviatie) door een +, daarboven door meer plussen, de 5 door een - en daaronder door meer minnen, zal misschien minder hard overkomen, maar lost het probleem van de grens tussen voldoende en onvoldoende functionerende scholen geen moment op. Dat zelfde geldt voor de oplossing om alle scholen een voldoende te geven, omdat alle onderwijsgegenden 200% gemotiveerd zijn. Dat zou men toch ook niet van bijvoorbeeld Fokker zeggen. Want ondanks in grote inzet van de Fokker werknemers maakte het bedrijf toch op fatale wijze verlies en functioneerde niet voldoende. De kern van de voldoende/onvoldoende grens bij scholen is dat de verantwoordelijken nooit, zelfs niet bij benade-

ring, hebben durven zeggen, wanneer een school onvoldoende functioneert. Door een heldere keuze (het gemiddelde heeft het cijfer 6) wordt dat probleem nadrukkelijk aan de orde gesteld. De oplossingen van Van Putten en Van der Kamp lossen dat probleem zeker niet op, maar verhullen het alleen.

6 Multilevel of geaggregeerd?

Veenstra et al. doen in hun commentaar twee interessante analyses op het VOCL'89 cohort: de een waarin zij met geaggregeerde gegevens op schoolniveau werken, evenals als ik dat doe met de inspectiegegevens, de ander waarmee zij met gegevens op individueel leerlingniveau schooleffecten met behulp van multilevel analyse vaststellen. Zij menen dat de uitkomsten van hun multilevel analyse de voorkeur verdienen, boven het werken met geaggregeerde gegevens op schoolniveau. Ik meen dat deze voorkeur om alvast twee redenen onjuist is: 1. Met multilevel analyse wordt een andere vraag beantwoord; 2. Multilevel stelt te zware eisen aan de beschikbaarheid van gegevens.

Met multilevel analyse en gegevens op leerling- en schoolniveau kan men vooral goed vragen beantwoorden met betrekking tot de kansen van leerlingen met uiteenlopende kenmerken in verschillende typen onderwijs en scholen. In dat geval is de analyse-eenheid de individuele leerling en niet de school. Het citaat dat Veenstra et al. uit de studie van Kreft (1993) geven spreekt op dit punt boekdelen: het gaat om de individuele leerling met zijn kenmerken in relatie tot een bepaalde school met bepaalde kenmerken. Maar bij kwaliteitsverschillen tussen scholen gaat het daar niet om, want dan is de school de analyse-eenheid en niet individuele leerlingen. Bij kwaliteitsverschillen tussen scholen gaat het om de vraag welke toegevoegde waarde scholen hebben voor de gemiddelde leerling op scholen van dat onderwijstype. Men is bij deze analyse op geaggregeerd niveau niet geïnteresseerd in bijzondere combinaties van leerlingkenmerken en schoolkenmerken, zoals Kreft die onderzocht (hoge Cito-score op een categoriale MAVO of een MAVO in een brede scholengemeenschap). Kortom, Veenstra et al. gebruiken een

techniek die op een andere vraag antwoord geeft dan op de vraag die in de *Trouw* publicatie aan de orde is: welke school heeft voor de gemiddelde leerling in dat onderwijstype de hoogste toegevoegde waarde? Dat betekent tegelijkertijd dat diegenen, die zeiden dat uit de *Trouw* rapportcijfers niet automatisch volgt wat voor een specifieke leerling de beste school was, gelijk hebben. Dat multilevel analyse in schoolloopbaanonderzoek een superieure techniek is, vloeit voort uit het type vragen dat in dat onderzoek aan de orde is (de individuele loopbaan van de leerling en de betekenis van schoolkenmerken daarvoor) maar niet uit de techniek-zelf. Voor andere vragen die zich op één analyseniveau afspelen, zijn andere technieken beter bruikbaar. Dat vervolgens de antwoorden op twee verschillende vragen (wat is de toegevoegde waarde van een school voor een specifieke leerling; wat is de toegevoegde waarde van een school van de gemiddelde leerling op dat onderwijstype) niet hoog correleren, is niet verbazingwekkend, maar is ook geen reden mijn methode met geaggregeerde gegevens of de uitkomsten van multilevel analyse te verwerpen. Uiteraard is de toegevoegde waarde van een bepaalde school voor een specifieke leerling, vastgesteld met multilevel analyse een resultaat dat ook voor de ouders van die specifieke leerling belangrijk is: want een school kan wel goed zijn voor de gemiddelde leerling, maar is de school ook goed voor mijn specifieke kind? Maar dit voordeel van multilevel maakt mijn resultaat voor de ouder van de gemiddelde leerling niet minder interessant, want zo weinig gemiddelde leerlingen zijn er niet in Nederland.

Een tweede bezwaar tegen de door Veenstra et al. voorgestane multilevel methode is dat zij voorlopig niet toepasbaar is bij het vaststellen van mogelijke kwaliteitsverschillen tussen scholen voor specifieke leerlingen. Het is immers ondenkbaar dat gegevens op individueel niveau van voldoende leerlingen van alle scholen beschikbaar gesteld zullen worden om die multilevel analyses mogelijk te maken. Als wetenschappers dan toch staan op multilevel analyses, sluiten zij zichzelf uit van het verdere maatschappelijke debat over de meting van de 'toegevoegde waarde' van scholen. De wetenschappelijke opvattingen over de favoriete

methode hebben dan de bovenhand gekregen, waardoor de maatschappelijke dienstverlening door de wetenschap gehinderd zou worden. In dat geval is het mogelijk betere (een superieur geachte multilevel methode) de vijand van het goede (een min of meer verantwoorde meting van kwaliteit van scholen). Dan zien onderwijswetenschappers af van verdere deelname en belonen door een tekortschietende kwaliteitsmeting scholen indirect voor onterechte selectie vooraf en tijdens het onderwijs.

7 Hoe verder?

De door mij gehanteerde methode kan op vele andere manieren verbeterd worden: een verfijndere meting van intern rendement (zijinstroom, vakkenpakketten), een betere correctie voor leerlingkenmerken, het toevoegen van andere kwaliteitsdimensies (extern rendement, sociaal-culturele vorming, burgerschapsvorming). Maar de kern van mijn aanpak (brede meting van kwaliteit; rekening houden met de kenmerken van de leerlingen) zal daarbij steeds nodig blijven, om te voorkomen dat scholen door een tekortschietende kwaliteitsmeting beloond worden voor onterechte selectie vooraf of tijdens het onderwijs. In tegenstelling tot wat Veenstra et al. in hun slot suggereren dragen hitlijsten van scholen op zich niet bij tot vergroting van ongelijkheid tussen scholen. Dat doen hitlijsten alleen wanneer zij een smalle meting van kwaliteit hanteren (bijvoorbeeld alleen eindexamencijfers of alleen deelname van leerlingen aan traditionele cultuuruitingen) én zij geen rekening houden met de kenmerken van de leerlingen. In dat geval wordt de 'cultureel rijke' en 'selectieve' school geprezen voor een kwaliteit die zij niet hoeft te hebben, en de 'cultureel arme' en 'open-toelating' school gestraft voor een gebrek aan kwaliteit dat zij ook niet hoeft te hebben. Maar als hitlijsten wel toegevoegde waarde weergeven en een brede meting van kwaliteit hanteren, zie ik niet in hoe dergelijke lijsten ongelijkheid in de samenleving vergroten. Integendeel, dergelijke goede hitlijsten laten juist zien dat een goede school niet behoef te betekenen een rijke of selectieve school en dat een arme en open-toelating school grote kwaliteit kan hebben. Maar dergelijke goede hitlijsten komen niet vanzelf tot

stand, omdat zij tegen de bestaande belangen en machtsverhoudingen ingaan. De gang van zaken rond de kwaliteitskaart die de staatssecretaris en de inspectie nu willen ontwikkelen als alternatief op de *Trouw* publicatie met toegevoegde waarde en een brede meting van intern rendement laten die bestaande belangen en machtsverhoudingen goed zien. Daarom horen wetenschappers niet alleen becommentariërend dit proces vanaf de zijlijn gade te slaan, maar daaraan actief deel te nemen als een onderdeel van de maatschappelijke dienstverlening vanuit hun wetenschap.

Literatuur

- Agerbeek, M. (1997). Liegen met statistiek. *Trouw*, 6 september.
- Dekkers, H.P.J.M. (1997). Divisie G; Social Context of Education. Het AERA-Congres 1997, Chicago. *Pedagogische Studiën*, 74, 297-299.
- Dronkers, J. (1997). Het misleidend rendement van scholen. *Index. Feiten en cijfers over onze samenleving*, 4, 1.
- Kreft, I. G. G. (1993). Using multilevel analysis to assess school effectiveness: a study of Dutch secondary schools. *Sociology of Education*, 66, 104-129.

Auteur

J. Dronkers is hoogleraar onderwijskunde (in het bijzonder stratificatie en mobiliteitsvraagstukken) aan de Faculteit der Pedagogische en Onderwijskundige Wetenschappen van de Universiteit van Amsterdam.

Mededelingen

Promoties

Universiteit Twente

Promovendus: Drs. H. de Vos

Titel proefschrift: Educational effects: A simulation-based analysis

Promotor: Prof.dr.ir. J. Scheerens

Co-promotor: Prof.dr. W.J. van der Linden

Datum: 3 april 1998

Rijksuniversiteit Groningen

Promovendus: Drs. B.B. Roberts

Titel proefschrift: Through the keyhole: Dutch child-rearing practices in the 17th and 18th century. Three urban elite families

Promotor: Prof.dr. J.J.H. Dekker

Datum: 26 maart 1998

Rijksuniversiteit Groningen

Promovendus: Drs. A. Camstra

Titel proefschrift: Cross-validation in covariance structure analysis

Promotor: Prof.dr. W. Molenaar

Datum: 7 mei 1998

Universiteit Utrecht

Promovendus: Drs. A. Bok

Titel proefschrift: Taalonderwijs in ontwikkeling. Het ontwerpen van onderwijs op basis van specifieke kenmerken van het taalleer-
vermogen in het kader van ontwikkelings-
onderzoek voor taal

Promotores: Prof.dr. J.S. ten Brinke en

Prof.dr. F. Goffree

Datum: 3 april 1998

Inhoud andere tijdschriften

Comenius

18e jaargang, nr. 1, 1998

Thema: Criminaliteitspreventie en opvoedingsondersteuning

Inleiding op het themadeel, door P. van der Laan, C. van Nijnatten en M. Smit

Het pedagogisch tekort van de strafrechtelijke kinderwet, door I. Weijers

Predictie van criminaliteit en preventief ingrijpen, door J. Junger-Tas

Delinquentie onder jeugdige allochtonen.

Omvang, achtergronden en preventie, door L. Eldering

Preventie als voorkeursstrategie, door J. Hermanns

Forum

Over de toekomst van het Nederlands als taal voor wetenschappelijke publicaties in de pedagogiek, door B. Levering en P. Smeyers

Pedagogisch Tijdschrift

22e jaargang, nr. 4/5, 1997

Themanummer: Behandeling in de orthopedagogiek

Behandeling in de orthopedagogiek, door F.J.H. Harinck en A.J.J.M. Ruijsenaars

Bouwstenen voor een orthopedagogisch

behandelingsmodel, door F.J.H. Harinck

Behandeling van kinderen en jongeren met

een verstandelijke handicap en ernstige

gedragsproblemen, door P. Swennenhuis en

A. Vermeer

Agressieve jongeren in tehuizen. Complexe

hulpvragen en een cognitief gedragsmatige

benadering, door M. Klomp

Ambulante behandeling van jonge kinderen

met een autistische spectrum-stoornis, door

I.A. van Berckelaer-Onnes

De behandeling van leerstoornissen. Een

'hardnekkige' aanpak van hardnekkige proble-

men, door A.J.J.M. Ruijsenaars

Blijvende verandering? Over de behandeling

van plegers van seksueel misbruik, door

R.A.R. Bullens

Pedagogisch Tijdschrift

22e jaargang, nr. 6, 1997

Twee rechtvaardigingen van kritisch denken en rationaliteit: Siegel en Apel, door J. Zevenbergen

On professional codes and ethical conduct in education, door M. McNamee

Burgerschapsvorming in de 21e eeuw. Een

internationale Delphi-studie, door S. Karsten

Tijdschrift voor Orthopedagogiek
37e jaargang, nr. 1, 1998

Comorbiditeit van depressie en gedragsstoornissen bij leerlingen van de zmk-school, door C.A.M. de Wit en E.M.B. Berg
Boekbespreking Speciaal
Afscheid van de taakanalyse? door A.J.C. Struiksma
Afscheid van de taakanalyse! Antwoord op de reactie van A.J.C. Struiksma, door W. van den Broeck
Naschrift, door A.J.C. Struiksma

Tijdschrift voor Orthopedagogiek
37e jaargang, nr. 2, 1998

Gehandicaptenzorg: context georiënteerde interventie. Oratie A. Vermeer
Zorgen: een zorg voor velen. Kanttekeningen bij de oratie van Prof.dr. A. Vermeer, door J. Kugel
De ontwikkeling en integratie van alleenstaande minderjarige asielzoekers, door M. Smit

Tijdschrift voor Orthopedagogiek
37e jaargang, nr. 3, 1998

Gezondheidsverschillen tussen leerlingen in het reguliere en in het speciaal onderwijs, door J. Spee-van der Wekke, J.F. Meulmeester en T. Radder
Psychosociale jeugdzorg en (dis)kwalificatie: de visies van cliënten en hulpverleners, door E.M. Scholte, M. Colton en F. Casas
De Nijmeegse Ouderlijke Stress Index en uithuisplaatsing: een nadere validiteitsstudie, door A.M.J.M. Donkers, A.J.L.L. De Brock en A.A. Vermulst

Tijdschrift voor Orthopedagogiek
37e jaargang, nr. 4, 1998

Opvoedingshulp: doel, methoden en effecten, door J.M.A.M. Janssens
Commentaar op oratie J.M.A.M. Janssens, door L. Vandemeulebroecke
Diagnostiek, preventie en behandeling van dysgrafie bij slechtiende kinderen, door D.R. van Peer, W. Ferwerda en C. Wassenberg
De identiteit van de orthopedagogiek op de tocht, door F. Veelo-Bouricius

Ontvangen boeken

- Distelbrink, M., *Opvoeding in Surinaams-Creoolse gezinnen in Nederland. Een eigen koers*. Uitgeverij Van Gorcum, Assen, 1998, f 39,50.
- Elen, J., & Staessens, K. (red.), *Onderwijskundige begeleiding in het hoger onderwijs. Uitdagingen en perspectieven*. Garant Uitgevers, Apeldoorn, 1997, f 39,50.
- Haenen, J., & Haitink, A. (red.), *Teamleren op school en in de klas*. Spruyt, Van Mantgem & De Does, Leiden, 1998.
- Löhmer, C., & Standhardt, E., *Themagecentreerde interactie. De kunst zichzelf en groepen te Leiden*. Garant Uitgevers, Apeldoorn, 1998, f 25,95.
- Pels, T., *Opvoeding in Marokkaanse gezinnen in Nederland. De creatie van een nieuw bestaan*. Uitgeverij Van Gorcum, Assen, 1998, f 45,-.
- Verhofstadt-Denève, L., *Adolescentiepsychologie* (5e herziene druk). Garant Uitgevers, Apeldoorn, 1997, f 39,50.