

Kees Both

## **OVER STANDAARDEN EN TOETSEN** **kritische stemmen uit het buitenland**

### **Vooraf**

Het onderwijsbeleid in ons land en de discussie daar omheen zijn vaak voor een deel een echo van wat ook elders speelt. Er zijn internationale trends zichtbaar en internationale vergelijkende studies van leerprestaties hebben vaak veel invloed op de keuzes die gemaakt worden, ook wat betreft de instrumenten voor kwaliteitszorg (Tillmann, 2001 – zie de literatuurlijst aan het eind van deze brochure). De tendens van het formuleren van ‘leerstandaarden’ en het invoeren van een verplicht systeem van toetsen dat daarmee verbonden is, is zo’n internationale trend.

Merkwaardig genoeg hoor en lees je bij ons heel weinig over kritiek op trends in landen die daarmee al veel langer bezig zijn, zoals de Verenigde Staten en het Verenigd Koninkrijk. Zo hebben, om bij het laatste te blijven, in 2001 Wales en Noord-Ierland het publiceren van toetsresultaten in de krant afgeschaft, vanwege de negatieve bijwerkingen die dit heeft. We hebben hierover in Nederland nergens iets gehoord en gelezen.

In de Verenigde Staten verslechtert het laatste jaar de situatie voor scholen die recht willen doen aan het hele kind. Een citaat uit een mail aan een van ons van Patricia Carini, een in kringen van huidige ‘progressive educators’ in de USA bekende pedagoge (zie Himley/ Carini, 2000; Carini, 1996), maakt iets daarvan voelbaar:

*‘..... an era of regressiveness in education across the country on a scale never experienced in my lifetime. How hard it is for teachers to make room for children, their play and learning, when even 5 year olds are denied recess and testing has assumed maniacal intensity. So, of course, this is the time when what Prospect (het netwerk waarin Carini actief is, KB) can offer is most needed.’*

Zij maakt ook duidelijk wat er volgens haar in deze situatie te doen staat:

*‘What I know is that we who have other visions must both persist and resist -- and continually commit and recommit ourselves to the hard, recursive work of keeping the door ajar (= de deur op een kier laten) so these visions of a humane education do not slip altogether from view and from practice’.*

Dit gedeelte van de brochure geeft enkele fragmenten uit discussie en onderzoek in het buitenland. Het is niet meer dan het geven van inkijkjes, om te laten zien hoe kritisch nagedacht wordt over ontwikkelingen die hier in bespreking zijn en soms al op stapel staan. De keuze van bronnen is eenzijdig, in de zin dat het om kritische stemmen gaat, die hier zeer zelden gehoord worden. Het gaat daarbij niet om onbekende, enigszins sektarische lieden, die maar wat roepen en om marginale tijdschriften. Mensen als Elliott Eisner en Deborah Meier, Linda Darling-Hammond en Selma Wassermann, Klaus-Jürgen Tillmann en John Elliott, William Spady en Patricia Carini zijn bekende pedagogen en het gaat om gerenommeerde tijdschriften en uitgeverijen. De genoemde publicaties die in dit gedeelte gebruikt worden grijpen zelf weer terug op onderzoek – Alfie Kohn bijvoorbeeld documenteert wat hij schrijft zeer zorgvuldig door verwijzingen naar literatuur. Er wordt ook enkele keren verwezen naar de bijdrage van Roger Standaert, die achterin deze brochure (in de bijlagen) is opgenomen.

De beknoptheid van deze bijdrage leidt ertoe dat er sterk geselecteerd moest worden en een dergelijke selectie is altijd subjectief. Maar wie dat wil kan natuurlijk altijd zelf de genoemde bronnen raadplegen.

Om misverstanden te voorkomen: het gaat hier niet om kritiek op toetsen als zodanig, dat zou onzin zijn (Boes, 2002). Hetzelfde geldt voor leerstandaarden als zodanig. Linda Darling-Hammond (Darling-Hammond, 1997) wijst erop dat kwaliteitsstandaarden werkzaam zijn binnen een professionele cultuur, bijvoorbeeld van artsen, architecten, psychologen en pedagogen. Zij omvatten

criteria voor ‘goed werk’ (‘standards of practice’), een beroepsethos. Zulke standaarden zijn ook basis voor opleiding, diploma’s, etc. Dit ‘professionaliseringsmotief’ moet scherp onderscheiden worden van het ‘beheersingsmotief’, dat in dit hoofdstuk bekritiseerd wordt.

Wie overigens denkt dat het bij ons niet zo’n vaart zal lopen is wel erg naïef. Je hoeft nog niet uit te gaan van complottheorieën of behept zijn met een grote angst voor ‘big brother is watching you’, om toch te zien dat hier sterke krachten aan het werk zijn en grote (ook commerciële) belangen in het spel zijn. Dergelijke systemen hebben – eenmaal in gang gezet - een ingebouwde neiging zich te verfijnen en alles en iedereen in te sluiten (‘systeemdwang’).

Met het oog op sceptici geven wij eerst aandacht aan recente waarnemingen van ‘Hollandse toestanden’ door een Duitse pedagoog. Het kan verkeren: bij ons wordt wel gesproken over ‘de Engelse ziekte’ als het om het evaluatiesysteem in Engeland gaat, maar die ziekte lijkt besmettelijk te zijn.

### **“Hollandse toestanden”**

Klaus-Jürgen Tillmann, hoogleraar in Bielefeld, beschrijft de tendens naar internationale vergelijkende studies van de prestaties van leerlingen op terreinen als wiskunde, natuurwetenschappen, lezen en andere facetten van taalonderwijs, vreemde talen. De uitkomsten daarvan leiden vaak tot grote opwinding en – als een land slecht scoort – soms ingrijpende beleidsmaatregelen om tot verbetering te komen. Dergelijke vergelijkende studies zullen – zo is de gangbare redenering – kunnen bijdragen aan de kwaliteitsverbetering in scholen. Maar dan moet er wel een koppeling gemaakt worden met de evaluatiegegevens van afzonderlijke scholen en zelfs klassen binnen deze scholen. Maar willen we dat wel? Want wat zijn de (onbedoelde) neveneffecten ervan? Wat zijn de valkuilen?

De ontwikkeling in Nederland gaat sterk in deze richting en wordt daarom aan een nadere beschouwing onderworpen. ‘Nederland is in Duitsland vooral bekend vanwege het hoge aandeel aan ‘bijzondere’ (“freie”) scholen en geldt dan ook als de “schöne Welt” van de schoolautonomie. Andere tendensen zijn veel minder bekend, met name vormen van centrale toetsing, verbonden met het schooltoezicht (inspectie) en de gevolgen daarvan voor het aanscherpen van de kwaliteitscontrole.’ Hij beschrijft dan drie elementen van dit systeem (daarbij aansluitend bij een publicatie van de Nederlandse inspectie), bij ons uiteraard bekend:

1. Centrale toetsen en ‘ranking’: de (nog niet verplichte) CITO-eindtoets basisonderwijs, welke ook gebruikt wordt als instrument voor het bepalen van de kwaliteit van de school, de verplichte centrale toetsen aan het eind van de eerste fase en tweede fase van het VO, die voor de helft bijdragen aan de eindbeoordeling voor elk vak (naast het schoolonderzoek), met ‘uitslagen’ in de krant.
2. Het schooltoezicht, met name IST, inclusief de eis van het bepalen van de ‘opbrengst’ van het onderwijs door middel van erkende, methodeonafhankelijke toetsen, c.q. leerlingvolgsystemen. Ook de IST-gegevens zijn openbaar. De uitkomsten kunnen aanleiding geven tot ‘verbeteringsplannen’.
3. Doelgericht werken aan schoolverbetering, op grond van onder andere aanbevelingen van de inspectie. ‘Als de scholen drie jaar achtereen “opbrengsten” vertonen, die meer dan een halve standaardafwijking onder de gemiddelde groepsnormen liggen, dan volgt in alle gevallen een verdiepte, uitvoerig inspectie, gericht op het beter analyseren van de problemen en het met de school komen tot zo nauwkeurig mogelijke afspraken voor verbeteringsmaatregelen’.

‘Dit model wordt door zijn Nederlandse voorstanders gezien als een goed gelukte koppeling van het objectief meten van schoolprestaties, het waarborgen van kwaliteit op schoolniveau en het afleggen van rekenschap aan overheid, ouders, e.a. Van mijn – Duitse - perspectief uit ziet het er eerder uit als een technocratisch controlemodel, waarvan de gevolgen en neveneffecten kritisch bezien moeten worden’ (zie voor een vergelijkbare visie Standaert, 2001).

Omdat de kritiek die hij dan uit ook elders in deze brochure wordt beschreven, wordt deze hier heel kort samengevat:

1. De koppeling van periodieke centrale toetsing met de evaluatie van individuele leervorderingen en –beoordeling (middels leerlingvolgsystemen) dwingt leraren om hun onderwijs af te stemmen op

de toets, waarschijnlijk zelfs tot specifieke trainingen voor de toets. Het is een moderne variant van de 'Nürnberger Trichter' – het 'overbrengen' en 'ingieten' van kennis en vaardigheden.

2. De toetsmakers krijgen een ongelooflijk sterke machtspositie.
3. 'Ranking' van scholen – in hun grote verscheidenheid – aan de hand van een enkel statistisch kengetal – leidt er via de ouders toe, dat sociale verschillen zich verscherpen. 'Goede scholen' krijgen kinderen met meer kansen, 'slechte' zullen verder 'achterblijven'.

Zijn advies: 'Een stevige discussie is nodig over de vraag of we dit (in Duitsland,) wel willen'.

In het volgende zullen negatieve effecten eerst kort beschreven worden. Vervolgens worden enkele thema's verder uitgediept.

### **Negatieve effecten kort beschreven**

Hier worden enkele voorbeelden gegeven van omschrijvingen van negatieve effecten van centrale toetsing, (Kohn (2000 en 2001; Meier, 2000; Ohanian 1999 en 2001). Een enkele keer nemen we de vrijheid om daaraan wat toe te voegen.

1. Leraren gaan op den duur de testcores meer vertrouwen dan het eigen professionele oordeel over kinderen, gebaseerd op langdurige observaties en interacties. Dit laatste garandeert op zichzelf nog geen kwalitatief goed oordeel, maar het eerste nog veel minder. Het herinnert aan de kritiek destijds van de Utrechtse onderwijsonderzoeker Karel Stokking op het SAVU-leerlingvolgsysteem, n.l. dat leraren daardoor 'gedeprofessionaliseerd' worden.

In dit verband horen we ook regelmatig het argument dat sterkere druk van buiten de school – met externe evaluatiemethoden – nodig is, omdat de kwaliteit van de leraren ver onder de maat is. Als je deze weg inslaat wordt dit echter een 'selffulfilling prophecy' (KB).

2. Goede – creatieve – leraren en directeuren raken gedemotiveerd en verlaten het onderwijs. In The Times Educational Supplement stonden in het jaar 2000 enkele interviews met leraren die het onderwijs verlaten hadden en inmiddels in het bedrijfsleven werkten. Teneur: 'Wij merkten voortdurend dat de school benaderd werd als een bedrijf en het leren van de kinderen als product. Daarvoor zijn wij niet het onderwijs ingegaan en het stuitte ons steeds meer tegen de borst. We zijn nu maar in echte bedrijven gaan werken.'

3. Leraren gaan steeds minder reflecteren op hun werk en leren steeds minder daarvan en van collega's. In plaats hiervan gaan ze zich afschermen en verdedigen en worden ze tevens in feite elkaars concurrenten.

John Elliott, hoogleraar aan de University of East Anglia (Norwich, England) beschrijft dit psychologische proces (Elliott, 2002) en heeft als vooropmerking dat pakweg vijftien jaar geleden leraren in hun werk wellicht te weinig kritisch benaderd werden en zonder meer vertrouwd werden. Nu is het omgekeerde het geval: wantrouwen overheerst. Onderwijzen is echter onmogelijk in een sfeer van wantrouwen. De externe controle door de inspectie in Engeland is zijns inziens te sterk geworden. In het verlengde van de toenemende 'transparantie' van het werk van leraren voor het externe publiek (het 'glazen huis' waarbinnen zij hun werk moeten doen, de zeer grote mate van kwetsbaarheid van leraren in hun beroep, vanuit hun engagement voor de kinderen) verdwijnt het vertrouwen. Leraren voelen dat en ontwikkelen op hun beurt wantrouwen tegenover de integriteit van de inspectie. Dat leidt gemakkelijk tot een ziekelijke angst voor de inspectie en de vertwijfeling die zo bij leraren ontstaat is contraproductief met het oog op verbetering van het de scholen op langere termijn. Leraren verliezen hun motivatie en belangstelling voor hun werk, vermijden risico's ('playing safe'). Er is geen ontwikkeling, maar juist verstarring, terwijl in onze onstabiele, dynamische maatschappij zin om te experimenteren en te vernieuwen juist zo belangrijk zijn (vgl. Troman, 2000).

4. Het is evident dat – gezien de grote druk en de ernstige consequenties van laag scoren - fraude uitgelokt wordt.

5. Leraren ontwikkelen in veel gevallen een negatieve houding tegenover kinderen die laag scoren, ook omdat zij erop worden afgerekend.

6. Het gesprek in de school over kinderen en onderwijs wordt ernstig versmald: het gaat niet meer over wat wij als school belangrijk vinden en wat kinderen (vaak verrassend, onverwachts) ontdekt hebben, maar steeds meer over toetsen en hoe die het best gemaakt kunnen worden.

7. In Amerikaanse scholen heeft het systeem van centraal toetsen geleid tot radicale veranderingen in de didactiek, met als resultaat een grotere eenvormigheid, vooral gericht op trainen voor de test. In de weken voor de periodieke toets worden ‘niet relevante’ vakken verwaarloosd en worden ‘inefficiënte’ activiteiten zoals excursies, het doen van proefjes, gesprekken over wat je bezighoudt, voorlezen, lezen voor je plezier, etc. opgeofferd (Kohn, Ohanian 1999 en 2001b).
8. Kinderen worden door het systeem van genormeerde toetsen gesorteerd, in plaats van dat vastgesteld wordt wat ze kunnen, hetgeen een pedagogisch ongewenste rivaliteit versterkt (Kohn; Ohanian).
9. Kinderen ontwikkelen een bepaald en zeer eenzijdig, om niet te zeggen vals beeld van wat intelligent gedrag is, namelijk het behalen van goede testresultaten. Hoge testcores duiden eerder op het wennen aan testsituaties, dan op betere leerprestaties. Goede testresultaten geven daarom heel vaak geen valide beeld van wat kinderen kunnen, ze vormen eerder een reflectie van kinderen die gedrild zijn om deze test te kunnen maken.
10. Linda Darling-Hammond geeft dit beeld: stel je voor dat beslist is dat ziekenhuizen beter moeten presteren en de standaard dus verhoogd moet worden. Alle patiënten dienen dan op een gestandaardiseerde wijze - dat wil zeggen op ongeveer hetzelfde moment - getemperatuurd te worden. Vlak daarvoor deelt het verplegend personeel hoge doses aspirine en koude dranken uit. De resultaten zijn inderdaad opmerkelijk – niemand heeft koorts! De kwaliteit van dit ziekenhuis is excellent. Uiteraard heeft de werkelijkheid achter deze hoge score weinig te maken met goede medische zorg en het nauwkeurig evalueren daarvan, net zoals het onderwijzen voor de toets volstrekt verschilt van goed onderwijs en het nauwkeurig evalueren daarvan.
11. De gangbare, technocratische, wijze van omgaan met leerstandaarden en van het evalueren daarvan leidt tot standaardisering van inhouden en didactische werkvormen (Darling-Hammond, 1997). Een voorbeeld van het laatste is in Engeland de invoeren van het verplichte ‘literacy-hour’, met een vast lesschema.
12. Alfie Kohn citeert een test-expert, die voorstelt om de ‘league-tables’ met toetsgemiddelden van scholen in de krant – als men ze toch beslist wil publiceren – dan maar daar te plaatsen waar ze thuishoren: op de sportpagina’s.
13. Deborah Meier (Meier, 2000) benadrukt in haar kritiek de desastreuze gevolgen voor de betrokkenheid van leraren, ouders en leerlingen en andere mensen en instellingen in de omgeving van de school.

### **Democratische controle?**

Volgens Meier wordt in naam van ‘democratische controle’ in werkelijkheid de democratie in en rond de school uitgehold. Het begint met het door experts voor het hele onderwijssysteem van een land vaststellen of het wel genoeg ‘oplevert’, gezien de investeringen, en in vergelijking met andere hoogontwikkelde landen. Daarvoor heb je landelijke normen nodig, die gespecificeerd worden naar schoolsoort, en verder naar verschillende ontwikkelingsniveaus. Dit wordt voorts omgezet in gestandaardiseerde toetsen, waarmee je individuen en scholen ‘objectief’ kunt vergelijken – regionaal, landelijk, internationaal. Voorts komt er een systeem van sancties voor falende scholen en deze moeten op een ‘onpartijdige’ wijze vastgesteld worden, dus door instellingen op enige afstand van de scholen. Men meent op deze wijze ook meer rechtvaardigheid te kunnen bereiken – gelijke kansen op een minimum aan goed onderwijs voor alle kinderen. Tenslotte komen er aanwijzingen voor ‘effectief’ onderwijs die ‘verbeteringen’ kunnen bewerkstelligen. De stappen in een dergelijk(e) systeem, c.q. denkwijze, vloeien logisch uit elkaar voort. Hieraan liggen aannames ten grondslag over democratie, schoolontwikkeling, kwaliteit en het borgen daarvan, etc.

Meier pleit voor ‘alternatieve’ aannames:

- Pluraliteit van scholen als rijkdom – meer definities van wat een ‘goede school’ is. Waar er een duidelijke hoofdstroom in het onderwijs zichtbaar is, moeten ‘alternatieve’ visies ook een kans krijgen, binnen ruim gedefinieerde algemene voorwaarden. Dit geeft binnen een onderwijssysteem een gezonde spanning.

- Plaatselijke controle mag niet uitgehold worden. Naast de oordelen van experts van buiten de locale situatie is democratische zeggenschap (‘ownership’) van de gemeenschap van direct betrokkenen – leraren, ouders, leerlingen en ‘afnemers’ van de school (vervolg scholen, werkgevers, e.a.) van

wezenlijk belang, ook met betrekking tot doelen, inhouden en werkwijzen. Daarin is het ook een belangrijk voorbeeld van democratisch leven voor de kinderen.

-Bij de evaluatie moeten verschillende soorten bronnen benut worden, waarbij de nadruk ligt op door de school zelf te formuleren en voor alle betrokkenen heldere ('stevige') standaarden en de plaatselijke toetsing daarvan. Allereerst door systematische interne evaluatie – waarbij de nadruk ligt op het laten zien van wat je werkelijk kunt ('performance-evaluatie', ook wel 'authentieke evaluatie' genoemd, zie daarover ook Both, 1995 en Darling-Hammond, c.s., 1995 en 2000), aangevuld door, op uitnodiging van de school, periodiek een 'externe audit' te organiseren door een onafhankelijke evaluatiegroep uit de regio. Evaluatie-activiteiten vanuit de overheid (in de VS de deelstaat) hebben in zo'n situatie het karakter van een aanvullende 'second opinion', binnen flexibel te hanteren landelijke normen.

### **Goed voor sociaal belemmerde kinderen?**

Meier schrijft ook nog over 'maatschappelijke gelijkheid', c.q. sociale rechtvaardigheid. Zij krijgt in het hier gebruikte boek echter repliek van iemand die haar vraagt of ze misschien vergeten is dat het formuleren van leerstandaarden juist bedoeld is om sociaal-economisch bepaalde verschillen aan te pakken. Juist arme gezinnen moeten profiteren van hogere eisen aan de output van het onderwijs. Het onderwijs aan kinderen uit achtergestelde gezinnen is nu slecht tot zeer slecht en dat wordt ook veroorzaakt door lage verwachtingen ten aanzien van deze kinderen, waardoor ze niet voldoende gestimuleerd worden. De reactie is in de trant van: 'Het is toch zo goed bedoeld, en hoe kan dan juist jij, Debbie Meier, daartegen zijn?'

Meier gelooft echter niet dat standaardisering en centrale bureaucratische controle van buiten de plaatselijke gemeenschap hiervoor het goede middel is. 'Het laatste wat we moeten doen is juist het vergroten van de centralisering en standaardisering, dat immers altijd al de scholen voor de armen heeft gedomineerd. Het is juist een deel van het probleem'. Zij wil dat wat in de VS het voorrecht is voor welvarende privé-scholen naar de gewone scholen in achterstandwijken brengen: meer beslissingsvrijheid om deze kinderen en hun gemeenschappen te dienen. Zij wijst ook op het succes dat zij had met de door haar opgebouwde en geleide scholen in East Harlem in New York, de in de VS en daarbuiten beroemde 'Central Park East Elementary and Secondary School' (zie daarover Meier, 1996). Deze school en de daaruit voortgekomen scholen in New York worden gekenmerkt door een sterke betrokkenheid van ouders bij de school, team-teaching van leraren, participatie van leerlingen in de vormgeving van het schoolleven en het onderwijs, kleinschaligheid en gemeenschap (scholen mogen niet te groot worden, of er moeten 'scholen in scholen' gevormd worden), hoge intellectuele en andere standaarden, die helder zijn voor iedereen en waarover ook in het openbaar verantwoording afgelegd wordt (zie boven). Een technocratische verandering van bovenaf dreigt experimenten als in New York/ East Harlem juist onmogelijk en kapot te maken, zoals zij zelf ervaart in de school die zij inmiddels in Boston aan het opbouwen is.

In de voorgaande paragrafen is al een en ander opgemerkt over de negatieve effecten voor met name de sociaal-economisch achtergestelde kinderen. Het hangt ook samen met de meestal gebruikte landelijk genormeerde toetsen. Het gaat hier immers om een relatieve maat – de score geeft aan welke plek ('percentielscore') een kind inneemt in het totaal van de populatie die deze test maakt. Als de kinderen bijvoorbeeld veel beter scoren op de rekentoets, dan zullen er toch 'slechte rekenaars' blijven, n.l. de kinderen in de laagst scorende sector (20%), hoezeer je ook je best doet als kind, leraar of school. Het is altijd een systeem van winnaars en verliezers, iemand (of een school) die goed scoort drukt anderen naar beneden.

De negatieve effecten van gestandaardiseerde toetsen voor met name de kinderen uit arme gezinnen worden zeer gedetailleerd beschreven in het boek van McNeil (2000). De conclusie luidt: '*Standaardiseren reduceert de kwaliteit en kwantiteit van wat in scholen onderwezen en geleerd wordt.* Dit onmiddellijke negatieve effect van standaardisering is de overheersende conclusie van een studie van scholen, waar de dwang van gestandaardiseerde controles de breedte en kwaliteit van de onderwijsinhouden reduceerde, de rol van de leraren verzwakte en actief leren door de leerlingen vrijwel onmogelijk maakte'.

De effecten op lange termijn zijn nog schadelijker: '*Op lange termijn bezien schept de standaardisering onrechtvaardige ongelijkheid en verbreedt de kloof tussen de onderwijskwaliteit*

*voor jongeren uit arme gezinnen en minderheden enerzijds en die van meer bevoorrechte leerlingen anderzijds ..... hoe langer de gestandaardiseerde controles werkzaam zijn, des te breder wordt deze kloof als het systeem van toetsen en het voorbereiden op de toets in scholen voor kinderen van minderheden in de plaats komt van het curriculum dat meer bevoorrechte leerlingen aangeboden krijgen. Deze nieuwe structuren van discriminatie worden voortgebracht door controles die begonnen in de scholen die in dit onderzoek onder de loep genomen werden en in de loop der jaren het dominante model geworden zijn voor het onderwijs in een van de grootste en qua bevolking meest pluriforme staten van de USA: Texas..... Het documenteert de groeiende macht van deze externe controles om de scholing van alle kinderen om zeep te helpen, in het bijzonder van kinderen met een Afrikaans-Amerikaanse en Latino achtergrond'.*

Het onderzoek richtte zich aanvankelijk op een heel andere vraagstelling, namelijk als zoektocht naar organisatiemodellen die gericht waren op structurele ondersteuning van authentiek en betrokken leren en onderwijzen. Dergelijke werkwijzen - in 'Magneetscholen', die vooral leerlingen uit arme - , hoofdzakelijk zwarte - en latino-gezinnen' ten dienste stonden - zouden nauwkeurig gedocumenteerd worden, dat was tenminste de bedoeling. Maar gedurende deze studie werden de betreffende scholen onderworpen aan een verandering, die inhield dat zij onder het regime van een gecentraliseerd systeem van voorgeschreven regels en gestandaardiseerde procedures van kwaliteitscontrole ('monitoring') gebracht werden. En de specifieke kwaliteiten van ook en zelfs deze veelbelovende scholen zag je onder je ogen verdwijnen.

De ontdekkingen die in dit onderzoek zo gedaan werden, zonder dat ze aanvankelijk gezocht waren, zijn op deze manier des te overtuigender.

Aan het eind van het boek schrijft zij:

*'Wij, als mensen in het algemeen en onderwijskundigen in het bijzonder, hebben nooit zoveel geweten over hoe kinderen zich ontwikkelen als nu. We hebben nooit zoveel geweten over cognitie en de werking van de hersenen en de effecten van aandacht en zorg en gemeenschap op de groei en ontwikkeling van kinderen. We hebben nog nooit zoveel als nu kunnen weten over en genieten van onze rijke culturele erfenis in de kunsten, de literatuur, geschiedenis, de natuurwetenschappen, wiskunde, astronomie, sport, de talen van deze wereld de geografie van plaatsen ver weg, het functioneren van onze lichaamscellen. Nooit eerder hebben we onze kinderen zoveel kunnen bieden. Hoe is het dan mogelijk dat zich systemen ontwikkelen, die de mogelijkheden voor het leren van kinderen ernstig versmallen, die oude vormen van discriminatie en achterstelling voortzetten en nieuwe vormen ontwikkelen om onze armste kinderen af te schermen van deze rijke culturele erfenis? Hoe kunnen we structuren als normaal accepteren, die de invloed van ouders en de overige gemeenschap om de school heen op de doelen van het onderwijs minimaliseren? We moeten juist die manieren van onderwijzen en het organiseren van scholen omarmen, die leraren stimuleren om het beste van wat zij weten en kunnen in te zetten in de klas. We moeten curricula en leeromgevingen scheppen die alle leervermogens van kinderen erkennen en bevestigen, met daarbij de erkenning van de diversiteit van kinderen als een inherent sterk punt van onze pedagogische inspanningen. En we moeten de wetenschappelijke expertise bevorderen die, in de woorden van Maxine Greene, "een meervormig en meerperspectivisch lezen van de wereld aanmoedigt" en die voortdurend tracht niet slechts onderwijsprocessen te interpreteren, maar uiteindelijk ook de kosten wil dragen.'* (p. 271)

### **Greep willen hebben op het onderwijs en de grenzen daarvan**

Elliott Eisner, hoogleraar onderwijskunde aan de Stanford-universiteit, graaft dieper en zoekt naar de grondhouding die de basis vormt van het hele systeem van leerstandaarden en de controle daarvan. (Eisner, 2001). Hij benoemt deze grondhouding als het rationeel willen beheersen van onze omgeving, inclusief de ontwikkeling van mensen en de processen tussen mensen. Het is de opvatting dat kennen = kunnen voorspellen/ voorzien = kunnen beheersen. Zo wil men ook de uitkomst van onderwijsprocessen kunnen voorspellen en de bijdrage die onderwijs levert aan de ontwikkeling van kinderen. Je formuleert zo nauwkeurig mogelijk de beoogde uitkomsten, zoveel mogelijk in meetbare vorm op korte en langere termijn (van lessen, leerjaren, enkele jaren, het basisonderwijs), stelt de beginsituatie van kinderen vast en gaat na afloop na of de beoogde uitkomsten bereikt zijn, c.q. wat het onderwijs aan de ontwikkeling van kinderen bijgedragen heeft: de 'toegevoegde waarde'. 'Kwaliteit'

moet uitgedrukt worden in kwantiteit, vanwege de grotere 'objectiviteit', meten is weten. Er wordt gestreefd naar zo universeel mogelijk geldende, 'wetenschappelijk vastgestelde' onderwijsmethoden ('what works'), die de beoogde uitkomsten zouden garanderen en die dan ook algemeen ingevoerd moeten worden. Als 'wetenschappelijk' geldt in dit verband alleen een 'harde' wetenschappelijke benadering als legitiem – alleen meetbare effecten tellen. En zo lijkt de cirkel zich te sluiten:

a. er wordt van uitgegaan dat 'weten' is kunnen voorspellen en kunnen voorspellen is beheersen; bovendien kan je binnen deze benadering alleen zeker weten door te meten (kwantificering) en op grond van deze beide aannames wordt de beoogde 'opbrengst' zo nauwkeurig mogelijk geformuleerd;

b. met hulp van een methode die de voorspelde uitkomsten zo goed mogelijk garandeert wordt aan deze doelen gewerkt;

c. de 'opbrengst' wordt gemeten (diagnose);

d. voor die kinderen die onvoldoende gescoord hebben wordt de methodiek eventueel aangescherpt. Deze beheersingsfilosofie leidt gemakkelijk tot allerlei reducties:

-Doelen die niet of moeilijk meetbaar / meetbaar te maken zijn vallen raken uit het zicht. Er is dan bijvoorbeeld onvoldoende geld om voor complexe vaardigheden passende evaluatie-instrumenten te ontwikkelen, of men vindt de beoogde doelen al gauw 'te vaag'.

-De reductie van onderwijsmethoden tot wat eenvoudig meetbare doelen dient.

-De nadruk op 'hapklare', 'overzienbare' modules, met korte leercycli, ten koste van langer lopende leerprocessen die per definitie niet te modulariseren zijn.

### **'Effectief' niet per definitie 'goed'**

John Elliott, hoogleraar aan de University of East Anglia (Norwich, Engeland), analyseert de discussie rond de 'effectieve school', inclusief het onderzoek daar omheen (Elliott, 1998, hoofdstuk 5). Zijn bevindingen bevestigen en versterken die van Eisner. Voorstanders van de 'effectieve school' verwijzen bij kritische vragen graag naar 'onderzoek', dat hun gelijk zou bewijzen ('what works'). Zelden komt ter sprake dat deze 'gegevens' verworven worden binnen een type denken over onderwijs zoals Eisner dat beschrijft en dat Elliott typeert als 'social engineering' - het ontwikkelen van een systeem van controle, via psychologische technieken, dwang, belonen en straffen. Dit systeem heeft eenzelfde structuur als Russische poppen – een subsysteem dat gecontroleerd wordt door een hoger subsysteem en zelf weer de controle uitoefent op subsystemen binnen het eigen bereik. 'Kwaliteit' wordt primair gedefinieerd als technische efficiëntie, maar daarachter gaat een onderwijsvisie schuil, waarin waarden als sociale controle en aanpassing een belangrijke rol spelen. Bureaucratisch ingestelde bestuurders vinden hier een ideaal instrumentarium om beleid in te voeren en te controleren – naar we mogen aannemen vanuit goede bedoelingen. Voor dit type onderwijsprojecten is dan ook vrij eenvoudig geld te krijgen. Leraren worden binnen dit geheel vooral gezien als technici, die de aanwijzingen zo goed mogelijk uitvoeren en niet als 'reflectieve practici'. Elliott wijst erop dat een 'effectieve school' niet per definitie identiek is aan een 'goede school'. Het hangt er maar vanaf wat je onder het laatste verstaat. De 'effectieve school-beweging' doet zich voor als 'research-based', maar vertegenwoordigt primair een bepaalde visie op het onderwijs. Binnen een andere onderwijscontext, vanuit een andere visie vormgegeven, zijn via onderzoek andere 'effecten' van onderwijs te ontdekken, die zich binnen 'effectieve scholen' niet of nauwelijks voordoen (Kleinespel, 1998).

### **Extrinsieke motivatie**

Een van de neveneffecten van de hiervoor beschreven trend is volgens Eisner ook het verschijnsel van de calculerende leerling. Leerlingen krijgen voortdurende de boodschap te horen dat het behalen van hoge testcores het belangrijkste is. Zij vinden allerlei manieren om met zo min mogelijk moeite zo hoog mogelijke scores te behalen. Niet de inhoud, waarin je je kunt verdiepen, waarin je je kunt verliezen is van belang, maar de toets. 'Ballast' moet dus overboord, met als resultaat: verschraving in doelen en inhouden.

Het merkwaardige is dat het onderwijs steeds minder gericht lijkt te zijn op het goed functioneren buiten een schoolcontext in het gewone leven (om dat laatste gaat het toch?), maar steeds meer op het beter functioneren op school zelf: goede testcores halen, om later betere testcores te kunnen behalen. Het schoolklimaat wordt bovendien steeds meer gericht op extrinsieke motivatie.

Eisner beweert niet dat 'beheersingleren' en 'directe instructie' met een meetbare feedback op zichzelf verwerpelijk zijn, integendeel. Zij kunnen heel functioneel zijn, ook binnen 'zelfgestuurd' leren. Hij waarschuwt echter – op grond van wat hij in talloze scholen waarneemt – voor een verre gaande inhoudelijke en didactische verschraving van het onderwijs. Het merkwaardige is dat momenteel in de onderwijswetenschappen tegelijkertijd sterker benadrukt wordt dat leren allereerst is, c.q. moet zijn: persoonlijk betekenis geven aan wat je leert, mede in interactie met anderen. Waar zijn de verrassing, de verbeeldingskracht, het improviseren gebleven?

Is onderwijs niet steeds ook: interactie, met alle onvoorspelbaarheid daarvan? (zie ook: Eisner, 1992 en 1998). Maak je met de sterke nadruk op het willen beheersen essentiële dimensies van het onderwijs niet kapot?

### **Alternatieve indicatoren 'goede school'**

Stel je eens voor – zegt Eisner – dat we voor de komende vijf jaar een moratorium op landelijk afgenomen gestandaardiseerde toetsen zouden afkondigen. Je kunt dan dus bij het beoordelen van scholen niet meer afgaan op hogere test scores. Waar zou je (als ouders, inspectie, etc.) wel op kunnen letten? Eisner geeft dan een hele lijst van kwaliteitsindicatoren, waaruit we er enkele noemen:

-De soorten problemen en activiteiten waarin leerlingen zich engageren. Worden ze gestimuleerd tot verwondering, tot het stellen van vragen over wat ze bestudeerd hebben? Let wel: hun eigen vragen en niet zozeer de vragen die wij hen stellen? De belangrijkste intellectuele prestatie die je kunt leveren is immers niet zozeer het oplossen van problemen, maar het formuleren ervan, de vragen die je kunt stellen.

-Worden de leerlingen geconfronteerd met meer en verschillende perspectieven? Kunnen ze daarmee omgaan, zonder in de war te raken?

-Kennismaken met verschillende manieren van kennen van de werkelijkheid, die elkaar niet uitsluiten, maar elkaar juist aanvullen: wetenschap, kunst, filosofie, e.a.?

-Worden er relaties gelegd tussen wat op school geleerd wordt en de wereld buiten de school, kunnen de leerlingen toepassen wat zij geleerd hebben? 'De werkelijk belangrijke afhankelijke variabelen in het onderwijs zijn niet gelokaliseerd in klaslokalen, noch in schoolgebouwen, maar buiten de school. Onze evaluatiepraktijken komen zelfs nog niet in de buurt daarvan. De echte uitkomst van het onderwijs is daar te vinden waar leerlingen kunnen doen wat ze zelf willen doen'.

Aan het slot van zijn artikel spreekt Eisner de hoop uit dat er 'een meer verlicht' begrip zal zijn van wat er werkelijk toe doet in de school en wat er werkelijk gebeurt. Dat zal ons hopelijk beter in staat stellen om het onderwijs te verbeteren.

### **De aantrekkingskracht van getallen**

In de lijn van Eisner doordenkend kan je je afvragen waarom mensen zoveel waarde hechten aan meten en aan getallen. Er wordt door de mensheid heel wat geteld, gemeten en berekend. Voor een deel ligt dat in het verlengde van natuurlijke processen – zo wordt er in ons eigen lijf al heel wat geteld, gemeten en gerekend en vinden op grond van gevonden waarden (bijvoorbeeld van hormoonspiegels) terugkoppelingen plaats. In de samenleving, zeker in onze moderne samenleving, is het een prominent verschijnsel – zie alleen al de aanduidingen op verpakkingen. Dat is uitermate functioneel, maakt zaken controleerbaar. Het is niet overdreven te zeggen dat onze maatschappij 'loopt op getallen'. We gaan er daarbij als regel maar van uit dat de getallen kloppen. Het uitdrukken van kwaliteiten in getallen is overigens – zie Eisner – niet zonder problemen. Niet alles laat zich zo 'vangen'.

Selma Wassermann, professor aan de Simon Fraser University in British Columbia (Canada), stelt ook de vraag waarom mensen zoveel waarde hechten aan getallen. (Wassermann, 2001). Het gaat volgens haar dieper dan 'intersubjectieve controleerbaarheid' zoals in de wetenschap aan de orde is en ook dieper dan de praktische functie in de samenleving. Het heeft vooral een existentiële dimensie: In een snel veranderende wereld voelen mensen zich onzeker, ervaren stress, moeten talloze beslissingen nemen op grond van onvoldoende zekerheden en waarvan zij de gevolgen ook niet kunnen overzien. In een dergelijke als chaotisch ervaren werkelijkheid is daar dan het getal. 'Niets voelt zekerder of geeft meer zekerheid dan een getal. Het doet er daarbij niet toe of de getallen of het meten voldoende



nauwkeurig zijn, we voelen ons eenvoudigweg geruster, als we kunnen geloven dat er iets is dat vastigheid biedt, iets waarop we met zekerheid kunnen rekenen, iets waaraan we ons kunnen vastklemmen als we met grote snelheid onzekere tijden tegemoet gaan.’

Misschien ligt het inderdaad op dit existentiële niveau, waarom ouders en leerkrachten zo graag leerprestaties van kinderen willen uitdrukken in een enkelvoudig getal. Een percentielscore van de CITO-toets ‘geeft houvast’, zo voelt het voor velen tenminste. Onlangs werd van een schoolleider vernomen dat zij elk jaar weer ‘gezeur hadden’ met veel ouders over de adviezen voor het voortgezet onderwijs. ‘We zijn toen begonnen met de CITO-eindtoets en het gezeur was vrijwel over’.

In dergelijke gevallen wordt zelden gevraagd naar de betrouwbaarheid en validiteit van het getal, die zijn soms ronduit bedenkelijk. Wassermann geeft in haar artikel voorbeelden van schijnexactheid in gevallen waar het ging om leven of dood (in de ruimtevaart), van regelrechte fouten in de wereld van natuurwetenschap en techniek, waarbij ook Nobelprijswinnaar Richard Feynman wordt geciteerd, die waarschuwt tegen een blind geloof in getallen.

*‘Ik denk dat het hard nodig is dat leraren geïnformeerde, evaluatieve observaties doen in hun praktijk. Ja, het is waar dat deze onjuist kunnen zijn; ja, het is waar dat ze subjectief zijn; ja, het is waar dat de oordelen van leraren over de leerprestaties van kinderen verkeerd, misleidend of bevooroordeeld kunnen zijn. Maar in plaats van deze oordelen van leraren te vervangen door het veel minder bruikbare alternatief van gestandaardiseerde toetsen kunnen we wellicht onze energie beter richten op het verbeteren van de door de leraar uitgevoerde evaluaties, door ze bruikbaar en ondersteunend te maken voor leerlingen en ouders en voor het totale leerproces.’*

Voorts geeft zij daarvoor aanwijzingen, die, aan het eind van haar artikel worden besloten met zeer concrete ‘standaarden’ voor leren en ontwikkeling van leerlingen in het voortgezet onderwijs, die leraren kunnen gebruiken als bouwstenen voor hun klas. Ze hebben betrekking op de kwaliteit van het denken, communicatie van ideeën, onderzoeksvaardigheden, interpersoonlijke vaardigheden, het ontwikkelen van een gezond zelfbeeld, waardeontwikkeling en zelfevaluatie-vaardigheden en – houdingen. Wassermann zoekt het in een heel andere richting – die van de ‘authentieke evaluatie’, verbonden met portfolio’s (Castelijns, e.a. 2002).

### **Veel kinderen worden onderpresteerders**

Gestandaardiseerde toetsen leunen sterk aan tegen de denk- en werkwijze van het klassieke intelligentiebegrip. Er wordt – in dit ‘psychometrische’ model – uitgegaan van de volgende vooronderstellingen:

- a. Er bestaat bij elk mens zoiets als een redelijk samenhangend en vast te stellen geheel aan denkvaardigheden dat ‘intelligentie’ genoemd wordt;
- b. mensen beschikken in meerdere of mindere mate over intelligentie en dat kan gemeten worden en samengevat in een enkelvoudig getal, het IQ;
- c. Dit IQ is redelijk stabiel en geeft de bovengrens van de intellectuele ontwikkeling van de betreffende persoon aan.

Dit psychometrische model wordt steeds sterker bekritiseerd (zie ook Gipps, 1998 en Gould, 1996). Allereerst bestaat er niet zoiets als ‘de’ intelligentie van een persoon, maar wordt intelligentie steeds meer gezien als een bundel van veelvormige mogelijkheden, die bij alle mensen in potentie aanwezig zijn, maar bij verschillende mensen in verschillende mate aanwezig zijn, c.q. ontwikkeld kunnen worden, met bij verschillende mensen verschillende accenten. Howard Gardner spreekt over ‘meervoudige intelligenties’ en beschreef er inmiddels acht. Naast het IQ hebben we nu bijvoorbeeld ook al het EQ, als maat van de emotionele intelligentie. Bovendien blijkt de stabiliteit, c.q. voorspelbaarheid van het leerplafond veel minder groot te zijn. Zie het werk van Feuerstein, die weigert van een vooraf vastgelegd plafond aan mogelijkheden uit te gaan en bijvoorbeeld met kinderen met het syndroom van Down verbazingwekkende resultaten bereikte. Robert Sternberg (Yale University) spreekt op grond van zijn ervaring in het onderzoek van selectieprocessen over drie typen intelligentie, die belangrijk zijn voor iemands leven, inclusief het beroepsleven: analytische intelligentie, creatieve intelligentie en praktische intelligentie. Hij gelooft dat alle mensen de mogelijkheid hebben deze intelligentietypen te ontwikkelen, hun leven lang, en moedigt leraren aan om ze alle te stimuleren. Het IQ betreft alleen het eerste type en bevat daarom een zeer beperkte uitspraak over latere successen in school, beroep en overige levenssferen (Perkins, 1998). David

Perkins, net zoals Gardner van de Harvard Universiteit, stelt duidelijk dat intelligentie onderwezen en geleerd kan worden – ‘the good use of the mind; the artful deployment of our faculties of thinking’ (Perkins, 1995).

De huidige praktijk van toetsen pakt vaak slecht uit voor kinderen, met name bij kinderen in sociaal-economische achterstandssituaties. Binnen het project aan de universiteit van Cambridge (Engeland) wordt onderzocht of nieuwe vooronderstellingen over intelligentie en cognitieve ontwikkeling een verschil kunnen maken. We volgen de redenering die aan dit project, met de titel ‘Learning Without Limits, ten grondslag ligt (Hart, 1998; Dixon, e.a. 2002).

In Engeland probeert men in het kader van de toetsen bij het National Curriculum en het publiceren daarvan in de krant de vergelijking eerlijker te maken voor scholen met veel ‘achterstandskinderen’. Zij scoren bijna altijd laag, gezien hun leefsituatie en scholen met veel van deze kinderen bengelen onderaan de lijst. Het op een verantwoorde wijze corrigeren van de ruwe toetsscores met het oog op sociaal-economische factoren blijkt een ondoenlijke opgave (Standaert, 2001). Daarom worden, om een eerlijker vergelijking te krijgen, aan het begin van de schooltijd, in basis- en voortgezet onderwijs, de cognitieve vaardigheden getoetst. Deze toets moet de basis zijn voor het voorspellen van het uiteindelijk te behalen niveau aan het eind van de betreffende schoolperiode (over x jaar dus), met het verwerken van de te verwachten natuurlijke ontwikkeling van het kind en de toegevoegde waarde van de school. Er wordt van uitgegaan dat je het ‘eindniveau’ kunt voorspellen. Leerlingen worden op grond van de toetsscores – vooral in het voortgezet onderwijs, minder in het primair onderwijs, - gegroepeerd naar ‘ability’, met daarbij aanvullende ondersteuning. Kinderen worden in feite vastgelegd op een vooraf bepaald niveau. De verwachtingen van leraren van wat de kinderen kunnen wordt hierop afgestemd, bewust of (vaker) onbewust. Voor dat deel van de kinderen waarvan je kunt veronderstellen dat ze onder het voorspelde niveau zitten kan dat gunstig zijn, zij kunnen voldoende uitgedaagd worden. Maar steeds meer blijkt dat een groot deel van de kinderen in dergelijke situaties te weinig uitgedaagd wordt en meer kan bereiken. Die kinderen worden dus door deze aanpak feitelijk benadeeld en tevens beperkt het in de praktijk leraren om te doen wat in hun vermogen ligt om de leerresultaten te verbeteren. Een citaat van de bioloog Stephen Jay Gould vormt een rode draad van ‘Learning Without Limits’.

*‘We leven slechts één maal. Bijna niets is zo onrechtvaardig als een leven de kans te ontnemen om naar iets te streven of zelfs maar hoop te koesteren, door een limiet van buitenaf op te leggen, maar die ten onrechte intern te situeren.’ (Gould, 1996, p. 66)*

Vanuit een *pluriforme* en *niet-deterministische* (‘transformatieve’) opvatting over intelligentie, c.q. leerpotentie ga je heel anders te werk, doe je meer recht aan de kinderen en hun veelvormige leerpotentie. Het onderwijsgedrag van tien leraren, waarvan men kon verwachten dat zij op een ‘transformatieve’ wijze werken, is een jaar lang geobserveerd en gedocumenteerd, aangevuld met diepteinterviews, etc. En zo werden vermoedelijk maatgevende kenmerken van een ‘transformatieve’ didactiek ontdekt, die in aanvullende studies werden getoetst moeten worden.

Als inzichten zoals in dit project zich doorzetten, dan zal ook blijken dat je met de huidige praktijk van gestandaardiseerde, landelijk genormeerde toetsen niet veel meer kunt aanvangen, omdat de grondslagen ervan ondermijnd zijn.

### **Naar een netwerksamenleving**

Op korte termijn bezien hebben we echter in ons land en elders nog te maken met een sterke lobby voor het invoeren van leerstandaarden, verbonden met een technocratisch controlesysteem. Er zijn echter steeds meer mensen, die dit type van onderwijsbeleid beschouwen als naweeën van een maatschappijvorm die aan het verdwijnen is, n.l. de industriële en bureaucratische samenleving (Spady, 2001; Elliott, 1998; in ons land Van Aalst, 2000). Metafoor voor deze samenleving is de lopende band, met massaproductie, een uniforme kwaliteitscontrole, een ver doorgevoerde taakverdeling en rationalisatie (standaardisering) zoals ook Eisner die beschrijft. Ook in de school zijn deze kenmerken terug te vinden. De scholen kunnen, zo dacht en denkt men soms nog steeds, met dit model even efficiënt leerlingen met gespecificeerde kenmerken produceren als Henry Ford T-Fordjes produceerde. Zie de vormen van onderwijsplanning, het leerstofjaarklassensysteem (‘grading’), de

gefragmentariseerde aanpak van de leerstof, de problemen met 'langzame leerlingen' binnen dit systeem. Veel scholen lopen tegen de grenzen van deze aanpak aan, ook omdat de samenleving, inclusief de kinderen, leraren en ouders, aan het veranderen is, langzaam maar zeker. Ook de in de vorige paragraaf beschreven dynamische, niet-deterministische visie op leren en ontwikkeling past in een nieuw maatschappijbeeld, dat getypeerd kan worden als een *netwerkmaatschappij*. Hiërarchische structuren worden in toenemende mate ter discussie gesteld en de samenleving kan steeds minder van bovenaf bestuurd worden. Dit staat haaks op de in dit hoofdstuk beschreven trend. Daar geldt: hoe verder van de school verwijderd, hoe hoger in de hiërarchie, hoe enthousiaster men is over centrale toetsing en omgekeerd: hoe meer kennis van het onderwijs men heeft, dat wil zeggen hoe dichterbij de schoolpraktijk staat, hoe kritischer en genuanceerder men daarover denkt (Kohn, 2000). Leren in de netwerkmaatschappij is bovendien veel minder lineair, er is meer aandacht voor verbeeldingskracht en creativiteit.

Over deze zich ontwikkelende netwerkmaatschappij valt veel meer te zeggen (zie met name Van Aalst, 2000). Een dergelijk toekomstperspectief kan ons helpen de huidige problemen zelfbewust, met humor en verbeeldingskracht tegemoet te treden en creatief verzet te plegen (Taylor, 1998). Het is mogelijk netwerken van gelijkgerichte individuen, scholen en organisaties te vormen.

### **Wat valt te doen?**

Alfie Kohn geeft aan het eind van zijn boekje 'The Case Against Standardized Testing' een lijst van adviezen voor actie – voor betere alternatieven en tegen de schadelijke tests. Het laatste en meest vergaande advies is: boycot de toetsen. Hij geeft dan een overzicht van dergelijke acties in de USA, met name door ouders opgezet. Je moet, schrijft hij:

- de steun van ouders zoeken en van andere individuen en groepen in de omgeving van de school;
- politiek onafhankelijk blijven, maar tegelijk politiek bewust zijn en zeker niet politiek naïef!
- een heldere en relatief eenvoudig uit te leggen visie en motivering ('rationale') hebben;
- de pers benutten.

Concreet moet men er naar streven om verschillende functies van toetsen goed te onderscheiden en te ontkoppelen. Toetsen voor nationaal periodiek peilingsonderzoek naar de kwaliteit van het onderwijs en voor internationaal vergelijkend onderzoek moeten niet verbonden worden met het vergelijken van scholen. Deze tests zijn nu vaak aanleiding voor maatregelen voor schoolverbetering, maar zijn daarvoor echter niet bedoeld, functioneren op een ander (beleids-) niveau. Het vergelijken van scholen heeft vrijwel altijd negatieve effecten en kan op den duur verbeteringen van het onderwijs eerder verhinderen dan stimuleren. Scholen moeten op z'n minst genoeg speelruimte hebben om de kwaliteitszorg op een manier vorm te geven die past bij hun opvatting van wat 'een goede school' is. En deze *pedagogische* wijze van evalueren moet dan wel systematisch uitgevoerd worden, inclusief het verantwoorden van de wijze waarop de ontwikkeling van individuele kinderen in het schoolteam systematisch gestimuleerd, besproken en gedocumenteerd wordt.

Om terug te komen op het citaat van Carini in het 'vooraf' van dit hoofdstuk: wat ons te doen staat is volhouden ('persist') tegen de huidige hoofdstroom in en verzet bieden ('resist') en zicht blijven geven op 'ander onderwijs', op scholen die, om met Alfie Kohn te spreken, onze kinderen verdienen, waarop ze aanspraak kunnen maken (Kohn, 1999).

## Gebruikte bronnen

- Aalst, H. van (2000), Leren in de netwerkmaatschappij, in: Mensen-kinderen. Tijdschrift voor en over Jenaplanonderwijs, jrg. 16, nr. 1 (sept.), p.10- 15
- Both, K. (1999), Authentiek leren en authentieke evaluatie, in: Mensen-kinderen, jrg. 14, nr 3, p. 4-7
- Carini, P.F. (1996), Voortbouwen op de sterke punten van kinderen, in: Mensen-kinderen, jrg. 11, nr. 3 (januari), p. 41-47
- Castelijns, J./ B. Kenter/ J. Janssen (2002), De kracht van portfolio's, Amersfoort (CPS)
- Darling-Hammond, L./ J. Ancess/ B. Falk (1995), Authentic Assessment in Action, New York/ London
- Darling-Hammond, L. (1997), The Right to Learn. A Blueprint for Creating Schools that Work, San Francisco
- Darling-Hammond/ J. Ancess/ B. Falk (2000), Authentieke evaluatie in actie (2), in: Mensen-kinderen, jrg. 15, nr 3, p. 12-15
- Dixon, A.. a.o. (2002), *Learning without limits*. Maidenhead (Berks.)
- Eisner, E.W. (1992), Educational Reform and the Ecology of Schooling, in: Teachers College Record, Vol. 93, nr. 4, p. 610-627
- Eisner, E.W. (1998), The Kind of Schools We Need. Personal Essays, Portsmouth (NH)
- Eisner, E.W. (2001), What does it mean to say a school is doing well?, in: Phi Delta Kappan, Vol. 82, nr. 1, p. 367-372
- Elliott, J. (1998), The Curriculum Experiment. Meeting the challenge of social change, Buckingham/ Philadelphia
- Gipps, C.V. (1998), Beyond Testing: Towards a Theory of Educational Assessment, London/ Washington, D.C.
- Gould, S.J. (1996), De mens gemeten. De geschiedenis van de intelligentietest, Amsterdam/ Antwerpen
- Hart, S. (1998), A Sorrytail: Ability, Pedagogy and Educational Reform, in: British Journal of Educational Studies, Vol. 46, nr. 2, p. 153-168
- Himley, M./ P. Carini (2000), From Another Angle. Children's Strengths and School Standards, New York/ London
- Kleinespel. K. (1998), Schulpädagogik als Experiment, Weinheim/ Basel
- Kohn, A. (1999), The Schools Our Children Deserve, Boston/ New York
- Kohn, A. (2000), The Case Against Standardized Tests. Raising the Scores, Ruining the Schools, New York
- Kohn, A. (2001), Fighting the Tests. A practical guide to rescuing our schools, in: Phi Delta Kappan, Vol. 82, nr. 1, p. 349-357
- McNeil, L.M. (2001), Contradictions of School Reform. Educational Costs of Standardized Testing, New York
- Meier, D. (1996), The Power of Their Ideas. Lessons for America from a Small School in Harlem, Boston
- Meier, D./ J.Kozol/ J. Cohen (eds.)(2000), Will Standards Save Public Education?, Boston
- Ohanian, S. (1999), One Size Fits Few. The Folly of Educational Standards, Portsmouth
- Ohanian, S. (2001a), Caught in the Middle. Nonstandard Kids and a Killing Curriculum, Portsmouth
- Ohanian, S. (2001b), De uitdaging van de keuzevrijheid, in: Mensen-kinderen, jrg. 17, nr. 1, p. 7-8

- Spady, W. (2001), Beyond Counterfeit Reforms. Forging an Authentic Future for All Learners, Lanham/ Maryland/ London
- Standaert, R. (2001), The temptation of simplification, A plea for integration of internal and external quality control in: H. Bachmann, e.a. Turning the Perspective. New Outlooks for Education, Enschede
- Taylor, Ch. (1998), De Malaise van de Moderniteit, Kampen, m.n. het laatste hoofdstuk: 'Tegen fragmentatie'
- Tillmann, K.J. (2001), Leistungsvergleichsstudien und Qualitätssicherung, oder: Auf dem Weg zu holländischen Verhältnissen?, in: Neue Sammlung, Vol. 41, Heft 3, S. 359-368
- Wassermann, S. (2001), Quantum Theory, the Uncertainty Principle and the Alchemy of Standardized Testing, in: Phi Delta Kappan, sept. 2001, p. 28-40

### **Websites**

[www.alfiekohn.org](http://www.alfiekohn.org)  
[www.fairtest.org](http://www.fairtest.org)