

# Taaltoetsen: een hefboom voor schooltaalbeleid?

**Marleen Colpin, Sara Gysen, Robin Heymans, Koen Jaspaert,  
Kris Van den Branden & Machteld Verelst**

In 2006 voerde het Centrum voor Taal en Onderwijs (K.U.Leuven) in opdracht van de Vlaamse minister van Onderwijs een studie uit naar de haalbaarheid en wenselijkheid van centrale taaltoetsen in Vlaanderen. De belangrijkste resultaten van deze studie worden in twee *VONK*-artikelen beschreven. In het vorige *VONK*-nummer concentreerden we ons op het nut van centrale taaltoetsen voor het opvolgen van de taalverwerving van individuele leerlingen. In dit artikel<sup>1</sup> gaan we in op de mogelijke impact van het invoeren van centrale taaltoetsen op het taalbeleid van een school. Opnieuw baseren we ons daarvoor op de nationale en internationale onderzoeksliteratuur.

## SCREENING EN DIAGNOSE OP SCHOOLNIVEAU

De bespreking van de impact van taaltoetsing op een schooltaalbeleid ligt uiteraard sterk in het verlengde van de screening op leerling-niveau. Toch ligt het accent hier anders. Daar waar op leerling-niveau de basisvraag is 'Hoe doet de leerling het?' is bij screening op school-niveau de basisvraag 'Hoe doet de school het?'. Eerder dan in te zoomen op individuele leerlingen zoekt het schoolteam in eerste instantie een antwoord op de vraag: 'Slagen wij erin om met ons onderwijs het vereiste niveau op het vlak van taalvaardigheid bij onze leerlingen te halen?' Ook hier zal het voor het schoolteam belangrijk zijn om (1) een maatstaf te bepalen (wat is het vereiste niveau?), (2) een valide, betrouwbare voorstelling van de bestaande situatie (de taalvaardigheid van de leerlingen) op te bouwen en (3) uit de vergelijking van maatstaf en bestaande situatie de gepaste conclusies te trekken.

## EIND-, BEGIN-, EN TUSSENSCREENING

Wat de maatstaf betreft, ligt het voor Vlaamse scholen voor de hand om de eindtermen en ontwikkelingsdoelen (en leerplannen) als richtsnoer te nemen. De eindtermen en ontwikkelingsdoelen beschrijven op verschillende niveaus van de onderwijsloopbaan wat leerlingen met taal moeten kunnen doen.

Als een school aan het einde van een bepaald niveau (bijvoorbeeld aan het einde van het basis-onderwijs) screent of de leerlingen de eindtermen halen, spreken we van *eindscreening*. Ook aan het begin van een onderwijsniveau kan screening nuttig zijn, bijvoorbeeld om vast te stellen hoe het met het beginniveau van de leerlingen is gesteld. Deze vorm van *beginscreening* kan gebaseerd zijn op de ontwikkelingsdoelen/

*Screening van de prestaties van haar leerlingen kan de school een beeld opleveren van haar effectiviteit.*

eindtermen van het voorgaande niveau (bijvoorbeeld een toets afnemen van leerlingen eerste jaar secundair die qua moeilijkheidsgraad gelinkt is aan de eindtermen basisonderwijs) of aan de ontwikkelingsdoelen/eindtermen van het niveau zelf. In het laatste geval zal men vaak werken met een normgerelateerde maatstaf (bijvoorbeeld de vergelijking met andere leerlingen) omdat het niet eenvoudig is te bepalen welke doelen op welk niveau op een bepaald tussenmoment behaald moeten zijn, of zal men een

relatie tot stand moeten brengen met prestaties op een eindscreening van datzelfde niveau.

***Niet voor alle scholen werken landelijk genormeerde screenings-instrumenten echter even goed.***

Naast eind- en begin-screening kan men uiteraard ook *tussentijds* screenen. Zo kan men bijvoorbeeld tussen de beginscreening in het eerste leerjaar van het lager

onderwijs en de eindscreening in het zesde leerjaar nog een aantal tussentijdse metingen inlassen (bijvoorbeeld aan het einde van elk leerjaar) om de groei van de leerlingengroepen te blijven volgen. Zoals uit het eerste artikel in de vorige *VONK* bleek, laten genormeerde leerlingvolgsystemen systematisch toe om de scores van de klasgroepen te vergelijken met de prestaties van grotere groepen leerlingen (meestal genormeerd op landelijk niveau). Met andere woorden, heel wat toetsen die ingezet kunnen worden voor de screening op leerlingniveau, kunnen ook dienst doen voor de screening op schoolniveau. Eerder dan te focussen op individuele leerlingen zal het schoolteam op dit niveau haar aandacht richten op de verdeling en bewegingen van de leerlingen over de normcategorieën en de proportie van de leerlingen die de gewenste normcategorieën (of het gewenste niveau) haalt.

Op tussentijdse screening zoals die door leerlingvolgsystemen worden gehanteerd, komt vaak de kritiek dat ze te zeer op het leerstof-jaarklassensysteem is gebaseerd. Ze wekt met andere woorden de schijn dat alle leerlingen zich op hetzelfde tempo en volgens dezelfde route zouden moeten ontwikkelen. Ze houdt te weinig rekening met het feit dat de natuurlijke ontwikkeling van vele leerlingen geen uniform lineair proces is, maar vaak met bokkensprongen verloopt. Dat kan dan als gevolg hebben dat op sommige tussentijden de school zich over sommige leerlingen (of groepen van leerlingen) te veel of te weinig zorgen maakt. Deze kritiek wijst er opnieuw op dat de school bij de interpretatie van screeningsgegevens ook op schoolniveau heel wat andere gegevens in rekening zal moeten brengen en zich dus niet alleen op toetsen of momentopnames mag beroepen.

## HET BEPALEN VAN DE TOEGEVOEGDE WAARDE

Screening van de prestaties van haar leerlingen kan de school een beeld opleveren van haar effectiviteit. Niet voor alle scholen werken landelijk genormeerde screeningsinstrumenten echter even goed. Scholen met hoge percentages anderstalige leerlingen bijvoorbeeld geven vaak aan dat de steekproef waarop de normering van de leerlingvolgsystemen was gebaseerd, niet representatief is voor de leerlingenpopulatie op de school. Zij hebben dan het gevoel dat de normen (dus: de maatstaf) te hoog ligt. Het is in dit opzicht dan ook veel correcter om bij het in kaart brengen van leerlingresultaten de input- en contextfactoren van de school mee in rekening te brengen. De toelichting bij de CITO-eindtoets basisonderwijs zet dit sterk in de verf:

*"Het kan informatief zijn om het eigen schoolgemiddelde met dat van andere scholen te vergelijken. Een school weet dan of en hoeveel het schoolgemiddelde lager of hoger is dan het landelijk gemiddelde. Van die positie kan echter niet zonder meer de kwaliteit van het onderwijs worden afgeleid! Uiteraard niet. 'Boven het gemiddelde' is niet automatisch 'goed', zoals '(ver) onder het gemiddelde' niet vanzelfsprekend 'slecht' is. Bij de beoordeling van de kwaliteit van het onderwijs komt meer kijken dan alleen het schoolgemiddelde. Er zijn namelijk tal van factoren waar scholen geen invloed op hebben, maar die wel mede de resultaten van hun onderwijs bepalen. Bijvoorbeeld instroomkenmerken zoals de sociaal culturele achtergrond van leerlingen en het gemak waarmee leerlingen leren. De Inspectie van het Onderwijs weet dat als geen ander en gaat bepaald niet alleen op de gemiddelde score af. En dit geldt niet alleen voor de gemiddelde score op de Cito-toets, maar ook voor de gemiddelde score op allerlei andere toetsen zoals de toetsen van leerlingvolgsystemen."*

Betere aanduidingen van de schooleffectiviteit kan men bijgevolg bekomen door de resultaten van de leerlingen van een school te vergelijken met de resultaten van scholen met een gelijkaardige instroom van leerlingen die binnen een gelijkaardige context opereren. De bestaande leerlingvolgsystemen en toetsen in Vlaanderen en Nederland bieden op dit vlak echter weinig steun. Enkel het Kleuterobservatie-instrument Taalvaardigheid (KOBI-TV) bevat naast 'Vlaamse' normen ook 'OVb-normen': deze werden berekend op basis van de allochtone leerlingen in de steekproef die aan de doelgroepindicatoren van het toenmalige OVb (Onderwijsvoorrangsbeleid) voldeden. Aan OVb-scholen wordt in de handleiding van KOBI-TV de raad gegeven om de

scores van hun leerlingen in de eerste plaats met die normen te vergelijken. Dit komt maar gedeeltelijk tegemoet aan het bovenvermelde probleem: de diversiteit binnen de toenmalige OVb-scholen in termen van instroom was immers nog bijzonder groot. Scholen met een specifiek type instroom dreigen dan ook op eigen houtje op zoek te moeten naar scholen met dezelfde input- en contextgegevens om op basis van de outputgegevens van hun leerlingen een beeld te krijgen van hun effectiviteit.

Bovendien krijgt een school pas echt een zicht op haar effectiviteit op het moment dat zij niet alleen context- en inputgegevens verrekent, maar ook een vergelijking maakt tussen de vaardigheid van de leerlingen bij de instap en de vaardigheid van de leerlingen bij het einde van een bepaald traject. Met andere woorden, de school krijgt een zicht op haar *toegevoegde waarde* door een beginscreening te vergelijken met een eindscreening en daarbij context- en inputfactoren in rekening te brengen (bijvoorbeeld door haar verschil tussen begin- en eindscreening te vergelijken met die van gelijkaardige scholen). In onderzoek naar schoolloopbanen en schooleffectiviteit worden op deze manier analyses van de toegevoegde waarde van scholen berekend en wordt systematisch vastgesteld dat er tussen scholen zeer grote verschillen optreden met betrekking tot hun toegevoegde waarde:

***Sommige scholen zijn veel succesvoller in het bevorderen van taalleren dan andere scholen.***

*"De resultaten geven aan dat schooleffecten echt bestaan en dat de verschillen tussen scholen groot zijn. De variantie*

*tussen scholen voor wat betreft taaltoetsen bedraagt 29%. Ook als verschillen in kenmerken van studenten in acht worden genomen, blijven er substantiële verschillen tussen scholen bestaan. Sommige scholen zijn veel succesvoller in het bevorderen van taalleren dan andere scholen.” (De Fraine e.a., 2003, p. 854; vertaald door KVdB)*

## NAAR EEN VOLWAARDIGE DIAGNOSE OP SCHOOLNIVEAU

Zeker voor scholen waarvan de toegevoegde waarde of de scores van de leerlingen bij de screening (systematisch) te laag uitvallen, kan screening op schoolniveau een eerste stap betekenen naar een fundamentele reflectie over haar school-effectiviteit en een zoektocht naar manieren om deze op te krikken. Scholen worden dan als het ware een spiegel voorgehouden

met betrekking tot de mate waarin ze erin slagen om de leerlingen de gewenste of haalbare groei in taalvaardigheid te doen doormaken. Op het moment dat de school naast de context- en input- ook haar procesvariabelen mee in deze reflectie betreft, zet ze

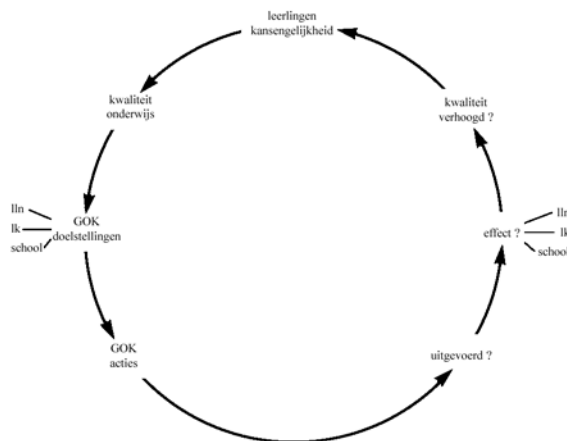
**Taaltoetsen  
vertellen minstens  
even veel over het  
onderwijs zelf als  
over de leerlingen  
die het volgen.**

een belangrijke stap op weg naar een volwaardige diagnose op schoolniveau. De school maakt dan een analyse van de inspanningen die ze getroost om de leerlingen tot leren te brengen met het oog op het verhogen van de kwaliteit van het door haar aangeboden onderwijs. Taaltoetsen

vertellen minstens even veel over het onderwijs zelf als over de leerlingen die het volgen. Net zoals in het geval van leerling-screening zou een gedegen diagnose van een dusdanige kwaliteit moeten zijn dat ze aanwijzingen geeft voor het ondernemen van doelgerichte acties ter verbetering van de kwaliteit van het onderwijs.

In essentie stoten we hier op de basisfilosofie die het ondersteuningsluik van het huidige Gelijke-Onderwijskansenbeleid (GOK) schraagt. De GOK-cyclus met haar ingrediënten van beginsituatieanalyse, zelfevaluatie en inspectiecontrole drukt sterk op het uitvoeren van een grondige diagnose van het onderwijskansenbeleid dat binnen de school wordt gevoerd in functie van het prioriteren van doelstellingen en acties die erop gericht zijn de kwaliteit en effectiviteit van het onderwijs voor alle leerlingen te verhogen. Daarbij wordt scholen sterk aangeraden om de resultaten op leerlingniveau te linken met de onderwijsinspanningen die op leerkrachtniveau en schoolniveau worden gedaan. Zo wordt in de visietekst rond de zelfevaluatie van GOK-scholen (Steunpunt GOK & Vlor-werkgroep Zelfevaluatie, 2005) aan scholen aangeraden om de link te leggen tussen de resultaten van de leerlingen en de kwaliteit van het door haar aangeboden onderwijs, om daaruit lessen te trekken voor eventuele acties ter optimalisering. Op haar beurt moet de proces- en productmatige evaluatie van die acties terug gerelateerd worden aan de verhoging van de algehele onderwijskwaliteit en de output bij de leerlingen. Dit vindt haar visuele neerslag in de zogenaamde ‘grote cirkel’ van een GOK-cyclus (figuur 1).

FIGUUR 1: De kwaliteitscirkel van een GOK-cyclus



Wat taalvaardigheid betreft, houdt diagnose in dat de school de vaststelling (screening) van leerlingresultaten op het vlak van taalvaardigheid integreert in een analyse van haar taalbeleid, dat op zijn beurt een integraal deel uitmaakt van haar breed zorgen gelijke-onderwijskansenbeleid. In binnen- en buitenlandse taalbeleidsinstrumenten, inclusief het beginsituatieanalyse-instrument

voor het thema 'taalvaardigheid' ontwikkeld door het Steunpunt GOK, wordt aan scholen gevraagd de kenmerken van haar leerlingenpopulatie (inclusief prestaties op taaltoetsen) in kaart te brengen en deze te linken aan aspecten van het onderwijsproces, zoals de

invulling van het taalvaardigheidsonderwijs, de rol van taal in de niet-taalkvakken, de samenwerking binnen het schoolteam rond taalbeleid, het nascholingsbeleid rond taal, en de communicatie met ouders hierrond.

Hierbij wordt sterk benadrukt dat het in kaart brengen van het taalbeleid en de reflectie daarover een teamaangelegenheid is. Ten eerste kan het beeld van de school op het vlak van taalondersteuning maar echt scherp worden gesteld als alle leerkrachten inzage bieden in hun percepties en klaspraktijk in verband met taal. Ten tweede zal de kracht van het gevoerde taalbeleid toenemen indien meerdere (ideaaliter: alle) leden van het schoolteam gezamenlijk de doelen bepalen waaraan gewerkt zal worden, gezamenlijk betrokken worden bij de uitvoering en opvolging van het actieplan, en gezamenlijk verantwoordelijkheid dragen voor het evalueren en eventueel bijsturen van het plan aan gewijzigde omstandigheden. Diagnose en acties ter verbetering van de onderwijskwaliteit mogen dus niet afgeschoven worden op een GOK-leerkracht of een gemotiveerde witte raaf, maar moeten samen worden opgenomen.

***Het in kaart brengen van het taalbeleid en de reflectie daarover is een team-aangelegenheid.***

Onderzoek naar de implementatie van de eerste GOK-cyclus wijst echter uit dat het leggen van de link tussen resultaten op leerlingniveau en inspanningen op leerkracht- en schoolniveau, niet systematisch door de huidige GOK-scholen wordt gedaan. Scholen vragen zich met andere woorden te weinig systematisch af of de inspanningen die ze zich getroosten om goed onderwijs te bieden, wel degelijk voor alle leerlingen het verhoopte effect hebben. Dat zou theoretisch te maken kunnen hebben te maken met het feit dat ze het effect van hun inspanningen soms te weinig systematisch in kaart brengen. Uit steekproeven zoals die van Verhelst (2005) in Vlaams-Brabant valt voor taal echter af te leiden dat nagenoeg alle scholen beschikken over en gebruik maken van toetsen om taal te screenen, al moet hierbij opgemerkt worden dat toetsen voor spelling en technisch lezen nog steeds

sterker in gebruik zijn dan toetsen die functionele taalvaardigheid direct in kaart brengen.

**Een hoog beleidsvoerend vermogen verhoogt de kans dat het afnemen van toetsen leidt tot een gedegen diagnose en weloordachte, planmatige actie.**

De bovenstaande onderzoeken naar de implementatie van de eerste GOK-cyclus suggereren dat het probleem in de eerste plaats te maken heeft met het feit dat veel scholen er vooralsnog moeilijk in slagen om de interpretatie van leerling-

resultaten te verbinden aan een planmatige reflectie en actie met betrekking tot het eigen beleidsvoerend vermogen. Dat heeft op zijn beurt met verschillende factoren te maken: schoolteams hebben het moeilijk om SMART<sup>3</sup> doelstellingen te formuleren, creëren te weinig netwerking en draagvlak binnen de school om tot een gezamenlijk

gedragen project te komen, linken doelstellingen voor het lerarenkorps te weinig aan doelstellingen op leerling- en schoolniveau, en zijn niet altijd in staat om een op maat gesneden nascholings- of begeleidingsaanbod te kiezen.

*"Uit de kwantitatieve en kwalitatieve analyses blijkt dat scholen die werken volgens de kwaliteitscirkel en die rekening houden met de drie niveaus school, leraren en leerlingen, met het creëren van een voldoende groot draagvlak, met netwerking en professionele aanpak van vraaggestuurde nascholing en professionalisering juist de voorbeelden van goede praktijk zijn."* (Ministerie van de Vlaamse Gemeenschap, 2005, p. 25)

Met andere woorden, er kan betwijfeld worden of het ter beschikking stellen van (meer) taaltoetsen het beleidsvoerend vermogen van scholen automatisch zal verhogen. De relatie ligt omgekeerd: uit de analyses van de eerste GOK-cyclus kan geconcludeerd worden dat een hoog beleidsvoerend vermogen de kans verhoogt dat het afnemen van toetsen leidt tot een gedegen diagnose en weloordachte, planmatige actie.

De kwaliteit van het beleidsvoerend vermogen van de scholen zal dan ook een sterke invloed hebben op de aard van de conclusies die de school trekt uit de resultaten van taaltoetsen. Onderzoek in binnen- en buitenland toont aan dat scholen bij tegenvallende resultaten niet altijd doelstellingen en acties bepalen die rechtstreeks met de verhoging van de onderwijskwaliteit of met de professionalisering van het leerkrachtenteam te maken hebben. Hieronder schetsen we een aantal valkuilen waarin scholen vaak trappen bij het boeken van tegenvallende leerlingresultaten.

## DE NEIGING TOT HOMOGENISEREN

Een eerste vaak vastgestelde reactie op de screening van leerlingresultaten is een (her)homogenisering van de leerlinggroepen. Dit kan zowel op schoolniveau (bijvoorbeeld in onderwijsstromen of studierichtingen), klasniveau (sterkere versus zwakkere klassen) als op groepsniveau (niveau groepen) gebeuren. Daarbij worden leerlingen van een bepaald niveau samengezet en krijgen ze aangepast onderwijs. Internationaal onderzoek wijst uit dat verregaande vormen van homogenisering negatieve effecten op gelijke onderwijskansen kunnen hebben, vooral indien deze herhomogenisering gepaard gaat met een verlaging van verwachtingen vanwege de leerkrachten ten aanzien van de leerlingen in de zogenaamde 'zwakkere' groepen. Lage verwachtingen kunnen leiden tot een verschralling van het onderwijsaanbod. Dat in Vlaanderen een substantieel gedeelte van de leerlingen uit het leerplichtonderwijs, en met name uit de bso-richtingen, uitstroomt met een onvoldoende hoog niveau van functionele geletterdheid, zou gedeeltelijk vanuit dit mechanisme verklaard kunnen worden, waarbij – in functie van het welbevinden van de leerling of als uitloper van bepaalde verwachtingen van het leerkrachtenkorps ten opzichte van de leerlingen – centraal wordt ingezet op beroeps- of praktijkgerichte competenties en minder op algemeen-maatschappelijke competenties zoals geletterdheid.

Tegelijk kunnen in verregaande vormen van homogenisering de potentiële effecten van heterogeniteit niet meer uitgebuit worden. Uit heel wat onderzoek blijkt dat zogenaamde zwakkere leerlingen sterker gedijen in groepen met een gemiddeld hoog cognitief gehalte of (op taalvlak) met hoger taalvaardige en moedertaalsprekers van de

onderwijstaal. Daar waar het onderwijsstelsel of de keuze van schoolteams tot verregaande gettovorming van kansarme leerlingen leidt, zijn die laatste hiervan dan ook de dupe. (Sterkere presteerders blijken minder last van herhomogenisering te hebben.) Als gevolg dreigt de ongelijke kansenkloof te verbreden in plaats van te versmallen. Het is in dit verband dan ook kenmerkend dat heel wat handleidingen van centrale toetsen die ingaan op welk soort acties geschikt zijn om de taalontwikkeling van laagscorende kinderen te verhogen, pleiten voor het werken in heterogene groepen. Onder andere TAL, KOBI-TV, TAS en TIST, maar ook bijvoorbeeld de handleiding van de Nederlandse Toets Tweetaligheid van het Cito pleiten hier expliciet voor:

***Daar waar het  
onderwijssysteem  
tot verregaande  
gettovorming van  
kansarme  
leerlingen leidt,  
zijn die laatste  
hiervan de dupe.***

*"Groepjes met kinderen van verschillend niveau kunnen positief werken. Kleuters die al wat verder in hun taalontwikkeling zijn gevorderd, kunnen als model fungeren voor de andere kinderen."* (Verhoeven e.a., 1995, p. 52-53)

Het gevaar van een sterke herhomogenisering van de leerlinggroepen doet zich sterk gevoelen bij de overgang van leerlingen van de ene klas naar de andere, of van het ene onderwijsniveau naar het andere. Met andere woorden, screening die zich afspeelt aan het einde van het ene onderwijsniveau en een impact kan hebben op de samenstelling van de groepen van het daaropvolgende niveau, kan sterk bijdragen tot herhomogenisering op dat volgende niveau. In het Vlaamse onderwijsstelsel is de overgang van het basisonderwijs naar het secundair onderwijs in

dit verband een cruciaal moment. Gezien het grote aanbod van secundaire scholen, het bestaan van de 1B-variant, en een *de facto* praktijk waarbij de invulling van de keuze-uren in veel scholen oriënterend/voorbereidend is op de aso/tso/bsc-structuur vanaf de tweede graad, dreigt bij het ingaan van het secundair onderwijs een

sterke segmentering c.q. herhomogenisering van de leerlingpopulatie tot stand te komen.

**Taaltoetsen hebben zelfs het potentieel om de neiging tot homogeniseren te versterken.**

Nicaise & De Smedt (2005) tonen aan – onder andere op basis van LOSO- en PISA-onderzoek – dat het Vlaamse ‘scholenpark’ van het secundair onderwijs

sterk gesegregeerd is, en dat België één van de PISA-landen is waar secundaire scholen het sterkst verschillen qua socio-economische en culturele verschillen van de leerlinginstroom. Allerhande mechanismen, waaronder de wijze waarop scholen, leerlingen en hun ouders reageren op en zich bewegen binnen de onderwijsmarkt blijken hiervoor verantwoordelijk te zijn.

*“Naast scholen, selecteren ook bepaalde groepen ouders (met name de middenklasse). Zij streven naar plaatsing op goed presterende scholen (...); de overige ouders stromen in in de resterende scholen met lagere prestatiecijfers. Deze reacties van ouders en scholen versterken de ongelijkheid, zowel tussen scholen als tussen ouders. Bepaalde groepen ouders krijgen wat ze willen, namelijk hun voorkeurschool (...), anderen niet. Bepaalde scholen werken met leerlingen met een goed perspectief op hoge onderwijsresultaten, andere scholen werken met leerlingen waarvan van meet af aan vaststaat dat hun onderwijscarrière minder ver en minder*

*hoog zal eindigen. Aangezien de onderwijsmarkt niet naar een evenwicht tussen vraag en aanbod tendeeft, staat men voor de keus of men van overheidswege de effecten wil corrigeren (de ‘visible hand’ van de overheid in de publieke markt), dan wel de ongelijkheid accepteert.”* (Onderwijsraad, 1999, p. 66)

Met de negatieve effecten van herhomogenisering in het achterhoofd is het dan ook niet verwonderlijk dat op een recente rondetafelconferentie van de Stichting Gerrit Kreveld een aantal academici een sterke oproep doet voor een echte comprehensieve eerste graad in Vlaanderen. Met dat pleidooi sluiten deze academici aan bij vaststellingen die onder andere door OECD werden gedaan op basis van een analyse van de resultaten van PISA 2000: de gemiddelde leesscores in het PISA-onderzoek liggen hoger in landen die pas vanaf de leeftijd van vijftien aparte onderwijsstromen voorzien.

In dit verband hebben taaltoetsen zelfs het potentieel om de neiging tot homogeniseren te versterken. En dan gaat het vooral over toetsen die een aura van objectiviteit uitstralen vanwege hun uniform, gestandaardiseerd karakter, hun waterdichte scoringswijzers, en hun vastliggende cesuren of normen. Dergelijke taaltoetsen hebben het potentieel om beslissingen aangaande doorverwijzing of oriëntering van leerlingen een objectief en onaanvechtbaar aura te geven, en andere factoren die van belang zijn bij een oriëntering – bredere kijk op de competenties van de leerling, interesse en toekomstperspectieven van de leerling zelf, studiehouding, enzovoort – te overtroeven. Sommige toetsontwikkelaars zijn zich daar terdege van bewust. Zo waarschuwen de auteurs van de TAS-toets, een gestandaardiseerde, genormeerde toets voor het begin

van het eerste jaar secundair onderwijs, dat de toets niet is bedoeld is om (homogene) klassen of leerlingengroepen samen te stellen. Het is de bedoeling dat de toets pas wordt afgenomen in de tweede helft van september, op het moment dat de (heterogene) klassen zijn samengesteld, en de toets kan fungeren als een beginscreening die de leerkracht de kans geeft een beeld te krijgen van de taalvaardigheid van de groep (en met name ook de samenstelling van de groep in termen van hoger- en lagertaalvaardigen).

## NADRUK OP REMEDIËRING

Naast herhomogenisering is een strakke focus op remediëring van individuele leerlingen een andere typische reactie op tegenvallende toetsresultaten. Screening van taalvaardigheid, zelfs indien die het doel heeft een schoolbeeld op te leveren, levert immers informatie op over de prestaties van individuele leerlingen. Het is dan ook verleidelijk om op basis van taaltoetsen een zeer gerichte, individuele remediëring van de probleemleerlingen (door gespecialiseerde leerkrachten) op te zetten en daarvan alle heil te verwachten. Deze praktijk was bijvoorbeeld voor de opstart van het Vlaamse Onderwijsvoorrrangsbeleid in 1990 zeer sterk verbreid (getuige bijvoorbeeld de speciale NT2-remediëringsklasjes voor allochtone leerlingen). Internationaal onderzoek en vooral ook de analyse van de wijze waarop momenteel op internationaal vlak met de bevordering van gelijke onderwijskansen wordt omgegaan, toont aan dat structureel werken aan de basisaanpak (met name het structureel verhogen van de onderwijskwaliteit zodat doorheen het onderwijssysteem krachtige leeromgevingen voor leerlingen ontstaan en veel problemen

kunnen worden voorkomen) een basisvoorwaarde is voor een duurzaam onderwijskansenbeleid.

In dit verband schoof CREDE (*Centre for Research on Education, Diversity and Excellence*), het federaal gesubsidieerd onderzoekscentrum in de Verenigde Staten dat meta-analyses uitvoerde naar het onderwijssucces van nationale minderheden) de volgende kwaliteitsindicatoren van een kansenvorderende basisaanpak naar voren (vertaling door KVdB):

***Structureel werken aan de basisaanpak is een basisvoorwaarde voor een duurzaam onderwijskansenbeleid.***

- (1) *Leerkrachten en leerlingen werken samen.* Hanteer werkvormen waarbij leerkrachten en leerlingen samen toewerken naar een gezamenlijk project, product of idee;
- (2) *Ontwikkel geletterdheid doorheen het hele curriculum.* Bevorder taalvaardigheid en geletterdheid in alle vakken, en werk in alle vakken met schriftelijke teksten;
- (3) *Verbind de lessen met het leven van de leerlingen.* Contextualiseer het onderwijs en verbind het curriculum en het onderwijs met de ervaringen van de leerlingen (thuis, op school, in de gemeenschap);
- (4) *Creëer uitdagende lessen voor de leerlingen.* Leg de lat hoog voor alle leerlingen; ontwerp activiteiten die alle leerlingen tot complex denken aanzetten;
- (5) *Hanteer werkvormen die interactie lossen.* Maak meer gebruik van dialoog tussen leerlingen en leerkrachten, en van werk in kleine groepjes dan van lange leerkrachtmonologen.

Binnen zulk een structurele basisaanpak kan een doordachte, systematische inpassing van individuele remediëring uiteraard een functionele plaats innemen, maar de aanpak van gelijke onderwijskansen mag niet tot remediëring beperkt worden. Van essentieel belang voor schooleffectiviteit blijkt, volgens de analyse van de internationale PISA-resultaten, de mate waarin schoolteams professioneel en vanuit hoge verwachtingen met de heterogene samenstelling van hun klassen weten om te gaan en de resultaten van toetsen aangrijpen om

fundamenteel te leren nadenken over hoe het onderwijs (nog) beter kan. Met andere woorden, schoolteams hebben nood aan beleidsruimte en stimulansen c.q. ondersteuning om hun professionele deskundigheid te verhogen.

***Vertrekkend van duidelijke doelstellingen kunnen toetsen een informatierijk element vormen in een brede evaluatiepraktijk die ook minder meetbare aspecten van leerlingontwikkeling in kaart brengt.***

Vertrekkend van duidelijke doelstellingen kunnen toetsen (waaronder een beginnende eindscreening waarbij scholen hun output kunnen vergelijken met die van vergelijkbare scholen) een informatierijk element vormen in een brede evaluatiepraktijk die ook minder meetbare aspecten van leerlingontwikkeling in kaart brengt en een brede reflectie op het onderwijsgebeuren centraal stelt. Geheel conform de wijze waarop in binnen- en buitenland tijdens de afgelopen jaren vorm is gegeven aan een gelijked onderwijskansenbeleid dient die reflectie zich centraal op een structurele verhoging van de kwaliteit van de basisaanpak (en dus de preventie van leerproblemen) te concentreren en de remediëring van individuele leerproblemen daarin een gepaste plaats te

geven. Voor het secundair onderwijs is het gevaar op perverse neveneffecten van screening op het vlak van herhomogenisering, reductie van het curriculum, daling van verwachtingen en (verdere) streaming van leerlingen veel groter dan in het basisonderwijs. Hier zullen flankerende beleidsmaatregelen die een stevige dam opwerpen tegen deze neveneffecten van essentieel belang zijn.

## NAAR PUBLIEKE TOETSEN?

In een aantal landen gaat men heel ver met het gebruik van centrale toetsen, zelfs zover dat resultaten van centrale taaltoetsen publiek worden gemaakt en scholen in een publieke ranglijst terechtkomen. Dit geeft aan dat de term 'centraal' veel betekenissen kan hebben. Van Dale wijst in zijn definitie van het woord niet alleen op het middelpunt zelf, maar ook op 'werkingen die vanuit een middelpunt uitgaan'. Daarmee is nog niet precies bepaald hoe de relatie tussen dat middelpunt en het veld er precies uitziet. De volgende varianten komen in de huidige internationale onderwijspraktijk het meest voor:

- (a) Een centrale overheid ontwikkelt taaltoetsen en neemt de toetsen in steekproeven van scholen af om op systeemniveau informatie te krijgen over de kwaliteit van het onderwijs of de haalbaarheid van bepaalde standaarden (doelstellingen);
- (b) Een centrale overheid ontwikkelt taaltoetsen die aan het veld ter beschikking worden gesteld. Een variant hierop is dat de centrale overheid een gespecialiseerde instelling subsidieert om taaltoetsen te ontwikkelen en die vervolgens

aan het veld ter beschikking te stellen. In deze variant is het veld vrij om van de toetsen al dan niet gebruik te maken en bepaalt het veld ook wat ze precies met de resultaten doet en naar wie ze die communiceert;

- (c) De centrale overheid stelt taaltoetsen niet alleen ter beschikking, maar verplicht of stimuleert het veld tevens om deze toetsen op een bepaald moment af te nemen. Het staat de scholen wel vrij om zelf te beslissen wat ze precies met de resultaten doet en met wie ze die communiceert;
- (d) De centrale overheid stelt taaltoetsen ter beschikking, verplicht of stimuleert het veld om de toetsen af te nemen, en vraagt de resultaten die werden geboekt, op;
- (e) De centrale overheid stelt taaltoetsen ter beschikking, verplicht of stimuleert het veld om de toetsen af te nemen en verplicht de scholen ook om de resultaten niet alleen bekend te maken aan de overheid, maar ook publiek te maken. Een variant op deze situatie is dat de overheid zelf (of een instelling op een tussenniveau) de resultaten bekendmaakt. Deze laatste variant leidt in de praktijk vaak tot publieke 'ranking' van scholen.

Op deze vijf standaardsituaties zijn talloze varianten mogelijk, maar de cruciale dimensies waarmee rekening moet worden gehouden, zijn: (1) wie ontwikkelt de toetsen en wie stelt ze ter beschikking?, (2) hoe verplicht/vrijblijvend is het afnemen? en (3) aan wie worden de resultaten bekendgemaakt? We stellen vast dat in tal van Europese en niet-Europese landen er momenteel een tendens valt waar nemen

van (a)/(b) naar (d)/(e). Onder andere in het Verenigd Koninkrijk, Zweden, Noorwegen en de Verenigde Staten is een systeem van publieke 'accountability' ingevoerd; in Hongarije wordt het binnenkort ingevoerd.

De Vlaamse overheid heeft zich tot hiertoe op domeinen (a) en (b) bewogen. Via periodieke peilingen heeft de overheid systeem-informatie verzameld over de mate waarin eindtermen met betrekking tot taal, wiskunde en wereldoriëntatie op het einde van het basisonderwijs worden gehaald. Ook heeft de Vlaamse overheid bepaalde instellingen gesubsidieerd die een aantal taaltoetsen hebben ontwikkeld en genormeerd, die dan vervolgens vrijblijvend aan de scholen werden aangeboden. Recentelijk heeft de Vlaamse overheid te kennen gegeven om middels haar programma voor Onderwijskundig Beleids- en Praktijkgericht Wetenschappelijk Onderzoek (OBPWO) een beweging in de richting van situatie (c) te maken: binnen dat programma wil de overheid tijdens de periode 2006-2011 een programma uitvoeren voor periodieke peiling van leerlingprestaties in het leerplicht-onderwijs.

Tot hiertoe dienden periodieke peilingen een functie op systeemniveau, namelijk het leveren van objectieve, betrouwbare en landelijke informatie over de mate waarin de getoetste eindtermen of ontwikkelingsdoelen bereikt zijn door de betrokken leerlingen. Dat laat de overheid toe een grondige analyse te maken van *"de sterke en zwakke punten in het onderwijscurriculum en daaruit de nodige lessen te trekken. De peilingresultaten leveren informatie op die mogelijk leidt tot bijstellingen van de eindtermen en/of van de leer- en onderwijsmethodes."* In het nieuwe OBPWO-programma voegt de Vlaamse overheid daar een functie aan toe die rechtstreeks te

maken heeft met de 'interne kwaliteitszorg van de scholen'. De overheid geeft aan dat ze parallelversies van de genormeerde, aan de eindtermen/ontwikkelingsdoelen gelinkte peilingsinstrumenten vrijblijvend aan de scholen zal aanbieden en via een centraal schoolfeedbacksysteem aan de scholen informatie zal bezorgen omtrent 'gecorrigeerde prestaties' die rekening houden met achtergrondkenmerken van leerlingen en eventueel schoolkenmerken. De potentiële winst van dit voornemen bevindt zich dus op verschillende niveaus:

- (1) De scholen krijgen valide en betrouwbare instrumenten om de eindtermen/ontwikkelingsdoelen te toetsen. Indien

***De introductie van directe toetsen voor functionele taaldoelstellingen kan ook een positief backwash-effect hebben op de didactiek van het taalonderwijs.***

voor sommige domeinen of eindtermen zulke instrumenten niet bestaan, kunnen de scholen via dit overheidsinitiatief voorbeelden krijgen van hoe zulke toetsen eruit zien, en tot een kwaliteitsvollere evaluatie van de betrokken doelstellingen overgaan. Dat kan mogelijk ook een positief *backwash*-effect hebben op de eigen (tussentijdse) toetsen die de school hanteert. Dit potentieel wordt met name groot op scholen die zich voor eindtermen taal beroepen op louter indirecte toetsen of toetsen die enkel deelvaardigheden in kaart brengen.

- (2) De introductie van directe toetsen voor functionele taaldoelstellingen kan ook een positief backwash-effect hebben op de didactiek van het taalonderwijs. Volgens Weir (1990) is *"a communicative approach to language teaching more likely to be adopted when the test*

*at the end of a course is itself communicative. A test can be a very powerful instrument for effecting change in the language curriculum..."* (p. 27)

- (3) De school ontwikkelt een 'correcter zicht' op haar interne kwaliteitszorg en een meer informatierijke omgeving door haar uitkomsten te vergelijken met de resultaten van andere scholen die onder vergelijkbare omstandigheden aan dezelfde doelstellingen werken.

Op basis van de boven vermelde OBPWO-oproep kan afgeleid worden dat de Vlaamse overheid vooralsnog geen expliciete plannen heeft om op te schuiven richting (d) en (e) waarbij afname van centrale toetsen verplicht wordt en resultaten publiek worden gemaakt.

## KRITIEK OP PUBLIEKE SYSTEMEN

Tegen het systeem van verplichte toetsing en publieke resultaten wordt in tal van onderzoekspublicaties bijzonder veel kritiek geuit. De meest voorkomende punten van kritiek worden hieronder opgesomd, waarbij kan worden opgemerkt dat een aantal punten ook van toepassing is op situatie (c):

- (1) De koppeling van 'high-stakes' centrale toetsing aan de evaluatie van individuele leervorderingen dwingt leraren om hun onderwijs af te stemmen op de toets en tot specifieke training over te gaan op het afleggen van de toets. Dat kan negatieve gevolgen hebben aangezien in centrale toetsing, vanuit het sterk uniform en gestandaardiseerd karakter van de toetsen en hun afnamemodaliteiten, vooral gefocust wordt op wat makkelijk meetbaar is.

Dat kan op termijn leiden tot curriculumverenging (doelen die niet gemeten kunnen worden, worden uit het oog verloren of worden als minder belangrijk ervaren) of tot een grotere eenvormigheid in het didactische aanbod van verschillende scholen. In de periode voor de toets dreigen 'niet relevante' vakken verwaarloosd en 'inefficiënte' activiteiten opgeofferd te worden.

- (2) Training voor de centrale testen kan niet alleen bij de leerkrachten aangewakkerd worden, maar ook bij ouders. Er kan dan een risico ontstaan dat net kinderen uit kansrijkere gezinnen 'beter' en intensiever worden getraind door hun ouders, waardoor hun kansen om goede testcores te halen, verhogen en de sociale kloof vergroot.
- (3) Didactisch kan het invoeren van centrale, publieke toetssystemen leiden tot een stap terug in de tijd. Samenhangend met de direct meetbare items die in de toets voorkomen, kan de didactiek zich terug sterker richten op diezelfde items die makkelijk onderwijsbaar zijn. Dat staat haaks op de huidige overheersende visies op leren, waarin met name het ontwikkelen van complexe competenties die kunnen aangewend worden in levensechte situaties centraal staat. De Vlaamse respondenten in een Delphi-studie rond dit onderwerp gaven dit als één van hun allerbelangrijkste zorgpunten aan (Van Petegem & Vanhoof, 2006).
- (4) Toetsenmakers krijgen een sterke machtspositie over wat er in het onderwijs gebeurt. Dat staat haaks op het basisprincipe van de pedagogische autonomie.
- (5) Leraren gaan op de duur testcores meer vertrouwen dan hun eigen professionele oordeel van leerlingontwikkeling, dat nochtans gebaseerd is op langdurige en diepgravende observaties. Centrale toetsen dreigen leerkrachten op die manier te 'deprofessionaliseren'.
- (6) Samenhangend met het vorige punt dreigen leraren steeds minder te reflecteren op hun eigen werk en steeds minder van collega's van andere scholen te (willen) leren. In plaats daarvan gaan ze zich afschermen en verdedigen tegen de concurrenten van de andere scholen.
- (7) Verplichte centrale taaltoetsen – met hun nadruk op strak meetbare output, verantwoording en doorgedreven transparantie – verhogen het economisch bedrijfskarakter van een school en voeden het beeld dat het onderwijskorps door de overheid wordt gewantrouwd. Bij sommige leraren heeft dat als gevolg dat ze geen risico's meer durven te nemen en hun praktijk niet meer vernieuwen; bij andere leraren en directeurs is dit een reden om het onderwijs te verlaten.
- (8) De grote druk en mogelijke sancties bij lage scores verhogen de kans op frauduleuze praktijken bij de toetsafnames.
- (9) Leraren ontwikkelen een negatieve houding tegenover laag scorende kinderen, omdat ook zij op die lage scores dreigen te worden afgerekend.

***Centrale toetsen dreigen leerkrachten te 'deprofessionaliseren'.***

Dat kan in sommige scholen leiden tot een verhoging van het 'wegdraineren' van slecht scorende kinderen om de eigen scores te verfraaien.

- (10) Het beleid in de school en het overleg daarover dreigt te worden versmald tot de inzet van de centrale toetsen. Het beleidsvoerend vermogen dreigt zich centraal toe te spitsen op prestaties die rechtstreeks met de toetsen samenhangen, eerder dan op het brede, eigen pedagogisch project.

- (11) De leerlingen zelf dreigen een zeer eenzijdig beeld van ontwikkeling op te bouwen. Je goed ontwikkelen op school dreigt gelijk te komen staan met goed scoren op bepaalde toetsen.

- (12) Leerlingen dreigen geïdentificeerd te worden met hun toetsscores. Leerlingen die aan centrale toetsen worden onderworpen, dreigen een etiket opgeplakt te krijgen of 'vanuit objectieve gronden' met lage verwachtingen vanwege de leerkrachten bejegend te worden. In diezelfde geest dreigt centrale taaltoetsing de homogeniseringsdrang in het onderwijs aan te wakkeren.

***Centrale taaltoetsing dreigt de homogeniseringsdrang in het onderwijs aan te wakkeren.***

- (13) Als centrale toetsen met normcategorieën werken waarbij kinderen met elkaar worden vergeleken, zullen er altijd winnaars en verliezers zijn, namelijk kinderen die relatief beter scoren en kinderen die relatief zwakker scoren. Dat kan in het nadeel werken van scholen die veel kinderen inschrijven die uit socio-economische milieus komen die een grotere kloof

met het schoolleven vertonen, want zij maken veel kans om bij de lagere scoorders terecht te komen.

- (14) Het is voor leraren aantrekkelijker om op scholen te werken met hoge scores dan op scholen met lage scores die onder druk van sancties moeten leven. Laagscorende scholen zouden het daardoor op termijn moeilijker kunnen krijgen om (goede) leerkrachten aan te trekken en te behouden. Dat is paradoxaal, omdat achter het idee van 'accountability' precies het voornemen zit om vooral in die scholen die laag scoren, beter onderwijs te stimuleren.

- (15) Het publiek maken van de resultaten verhoogt het risico dat de pers eenzijdig op (bepaalde) cijfers zal focussen en scholen op eenzijdige wijze zal stigmatiseren. Sommige scholen dreigen op die manier afgestraft te worden ondanks hun harde werk.

Het is opvallend dat de bovenvermelde critici geen tegenstanders van toetsen tout court zijn, maar vooral de uitwassen aan de kaak stellen van een te strak doorgedacht systeem van centrale toetsing. Veel van de boven aangehaalde punten zijn gebaseerd op kleinschalige studies of op theoretische pleidooien. Toch bestaan er ook studies op grotere schaal.

Zo besluiten studies in Texas dat centrale toetsing die gepaard gaat met ranking van scholen, negatieve effecten heeft voor scholen met sociaal-economisch achtergestelde kinderen: zo versmalt het curriculum in deze scholen sterk tot toets-training en 'basic skills'. Daarentegen blijft in scholen met een hogere instroom op het vlak van SES (sociaal-economische status) volop ruimte voor probleemoplossende

activiteiten, verkenning van de wereld en 'higher-level critical thinking'. Hillocks (2002) stelt vast dat vooral in scholen met veel doelgroepeleringen het schrijfonderwijs sterk verengd wordt tot het aanleren van schrijfconventies en vele vormen van creatief denken worden geweerd. Deze scholen worden bovendien sterk onder druk gezet om bepaalde lesmethoden en materialen die rechtstreeks op de toets voorbereiden, in te voeren en met niveau-groepen te gaan werken. In Arizona stellen onderzoekers vast dat in hoge SES-scholen de verwachtingen van leerkrachten stijgen, terwijl die in lage SES-scholen onder de invloed van testcores dalen.

In hun grootschalige studie onderzochten Hanushek & Raymond (2004) of 'school accountability' heeft geleid tot verbeterde leerlingprestaties en gelijkere onderwijskansen. Zij richtten zich op de grootschalige 'No Child Left Behind'-policy in de Verenigde Staten (NCLB). Een uitgangspunt van dit beleid is dat toetsen leiden tot betere resultaten. Staten werden opgeroepen om 'accountability systems' op te richten waarbij alle kinderen tussen de '3th and 8th grade' van de basisschool jaarlijks worden getest op wiskunde, Engels en wetenschappen. Staten maken de resultaten van hun scholen publiekelijk bekend.

Op het moment van het onderzoek hadden 39 staten van de Verenigde Staten een systeem van publieke 'accountability' ingevoerd. De onderzoekers brachten informatie over de door de staten ingevoerde systemen in verband met de scores van de leerlingen van '4th grade' en van '8th grade' op de afgenomen centrale taaltoetsen. Zij vergeleken daarbij de scores van blanke leerlingen met die van Hispanics en zwarte leerlingen. De onderzoekers komen tot de conclusie dat in de Verenigde Staten het

systeem van 'accountability' heeft geleid tot hogere gemiddelde scores op de toetsen. In staten waar het systeem sneller werd ingevoerd, is de winst hoger dan bij de zogenaamde 'late adopters'. Opvallend is dat het volume aan overheidsgeld dat in het onderwijsbeleid wordt gepompt, geen cruciale impact heeft op de effectiviteit van de scholen.

De resultaten zijn echter niet onverdeeld positief. Vrij verontrustend vinden de onderzoekers dat Hispanics en zwarte leerlingen veel minder vooruitgaan dan blanke leerlingen. Vooral zwarte leerlingen blijken de dupe van het systeem te zijn. De onderzoekers stellen zelfs vast dat de inhaalbeweging die Hispanics en zwarte leerlingen aan het maken waren voor het invoeren van het 'accountability' systeem, nu wordt teruggedraaid. De onderzoekers concluderen dat publieke 'accountability' systemen niet de kracht hebben om de ongelijke-onderwijskansenkloof te dichten en soms zelfs het tegenovergestelde effect blijken te hebben. Dat heeft in de onderzochte staten te maken met het feit dat heel wat andere factoren op het onderwijs inwerken (en sommige van de bovenvermelde negatieve neveneffecten opspelen), waarmee binnen een beleid dan ook rekening moet worden gehouden. Zo blijkt uit de verdere analyses van Hanushek & Raymond onder andere dat de sterke concentratie van zwarte leerlingen in bepaalde gettoscholen een zeer negatieve impact heeft op de resultaten.

## CONCLUSIE

De bovenstaande Amerikaanse meta-analyse toont aan dat een beleid dat, in mindere of meerdere mate, nadruk legt op

**Taaltoetsen vormen een element dat gericht is op een stimulering van de groei van scholen tot beleidseffectieve, hoogreflectieve en daadkrachtige teams.**

outputverantwoording door scholen, omringd moet worden met tal van andere overheidsmaatregelen die er sterk op gericht zijn mogelijke uitwassen van het systeem tegen te gaan. Met name blijken maatregelen ter bevordering van het beleidsvoerende vermogen van de scholen, de vrijwaring van heterogene school- en klassamenstellingen, en de professionalisering en verhoging van de doelmatigheidsbeleving van het lerarenkorps volgens de bovenvermelde bronnen van bijzonder groot belang te zijn.

De invoering van centrale taaltoetsen kan met andere woorden enkel bijdragen tot een informatierijke omgeving indien de informatie waarop de school zich baseert om na te denken over leerlinggroei en de onderwijsinspanningen die ze daartegenover zet, niet wordt verarmd tot de resultaten van de taaltoetsen alleen. Opschuiven naar de meest doorgedreven vormen van publieke 'accountability' verhoogt de kans dat dat laatste gebeurt. Taaltoetsen vormen vanuit deze vaststelling dan ook beter niet de spil van een gelijke-onderwijskansenbeleid, maar eerder een element dat gericht is op een stimulering van de groei van scholen tot beleidseffectieve, hoogreflectieve en daadkrachtige teams die op deskundige wijze de kansen van alle leerlingen op volwaardige ontplooiing vanuit hoge verwachtingen gaaf houden.

*Marleen Colpin, Sara Gysen, Robin Heymans,  
Koen Jaspaert, Kris Van den Branden & Machteld Verelst  
Centrum voor Taal en Onderwijs  
Blijde-Inkomststraat 7  
3000 Leuven*

## Noten

- 1 Dit artikel is een samenvatting van hoofdstuk 2 van

*Colpin, M., Gysen, S., Jaspaert, K., Heymans, R., Van den Branden, K. & Verhelst, M. (2006). Studie naar de wenselijkheid en haalbaarheid van de invoering van centrale taaltoetsen in Vlaanderen in functie van gelijke onderwijskansen. Leuven: Centrum voor Taal en Onderwijs.*

*De volledige tekst van dit hoofdstuk (inclusief alle bibliografische referenties) kan gratis gedownload worden van de website van het Centrum voor Taal en Onderwijs: <[www.cteno.be](http://www.cteno.be)>.*

- 2 SMART staat voor Specifiek, Meetbaar, Acceptabel, Realistisch, Tijdsgebonden.

## Bibliografie

*Om de leesbaarheid van dit artikel te verhogen, werd ervoor geopteerd om de vele bibliografische referenties die in de oorspronkelijke tekst van de haalbaarheidsstudie zaten, niet op te nemen. U kunt de integrale tekst met referenties downloaden van <[www.cteno.be](http://www.cteno.be)>. Hieronder verwijzen we enkel naar de werken die in dit artikel geciteerd werden.*

*De geciteerde toetsen in het artikel kunnen opgezocht worden via de bijlages van het rapport die evenzeer van de website van het Centrum voor Taal en Onderwijs downloadbaar zijn. In die bijlages worden alle toetsen op korte, handige fiches beschreven.*

### Geciteerde werken

De Fraine, B., Van Damme, J., Van Landeghem, G., Opdenakker, M.-C. & Onghena, P. (2003). The effect of schools and classes on language achievement. *British Educational Research Journal*, 29, p. 841-859.

Hanushek, E. & Raymond, M. (2004). Does school accountability lead to improved student performance? *NBER Working Papers*, National Bureau for of Economic Research.

Hillocks, G. Jr. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.

Ministerie van de Vlaamse Gemeenschap (2005). *Onderwijsspiegel 2004-2005*. Brussel: Departement Onderwijs.

Onderwijsraad (1999). *Handle with care. Buitenlandse ervaringen met het publiceren van schoolprestatiegegevens*. Den Haag: Onderwijsraad.

Steunpunt GOK & Vlor-werkgroep Zelfevaluatie (2005). *Aanzetten tot zelfevaluatie binnen GOK*. Te downloaden van de website van het Steunpunt GOK: <[www.steunpuntgok.be](http://www.steunpuntgok.be)>

Van Petegem, P. & Vanhoof, J. (2006). *Kwaliteit van scholen (openbaar) vergelijken? Het zwart-wit denken voorbij*. Mechelen: Wolters-Plantyn.

Weir, C. (1990). *Communicative language testing*. London: Prentice Hall.