



Toetsing als kans om te leren

Toetsing kan het leren en de motivatie van leerlingen veel beter dienen dan het nu vaak doet. Alleen maar cijfers geven verarmt de mogelijkheden die toetsing heeft. In bijna elke school beraden mensen zich op het eigen toetsbeleid. Wat helpt bij het nadenken hierover? Dit meinummer over toetsing openen we met een lange bijdrage. Samen met lector Dominique Sluijsmans maken we de balans op van de staat van toetsing, aan de hand van het advies van de Onderwijsraad uit december 2018.

DOOR RENSKÉ VALK EN DOMINIQUE SLUIJSMANS

Als je vader op de kweekschool werkte, als leraar en later als directeur van de PABO, en je moeder de KLOS volgde en in het basisonderwijs werkte, kun je nog van alles worden, later. Maar Dominique Sluijsmans ontwikkelde een onvermijdelijke liefde voor het onderwijs. Na haar studie pedagogiek/onderwijskunde in Nijmegen, en haar promotie op het gebied van peer assessment vaardigheden bij aanstaande leerkrachten, is ze nu lector Professioneel beoordelen aan Zuyd Hogeschool. Daar, zo zegt ze, kan ze voortdurend de verbinding opzoeken tussen onderzoek en onderwijs. Haar specialisme: toetsing, maar dan vooral gericht op leren.

De Onderwijsraad schrijft: 'Toetsen en examens dragen onvoldoende bij aan de kwaliteit van het onderwijs.' En: er wordt te veel gewerkt met summatieve in plaats van formatieve

toetsing, en te vaak met kwantitatieve in plaats van kwalitatieve toetsen. Voor een juist begrip van de gebruikte termen, hoe beschrijf jij: summatieve toetsing, formatieve toetsing, kwantitatieve toetsing en kwalitatieve toetsing?

'Fijn dat je deze vraag stelt, want ik merk in de praktijk dat er snel misvattingen ontstaan over deze begrippen. Als het gaat om formatief en summatief is het van belang te beseffen dat formatief en summatief geen eigenschappen zijn van een toets. Ze verwijzen naar eigenschappen van een beslissing die je neemt op basis van informatie uit een toets. Een toets zie ik dan als een moment waarop je informatie verzamelt over je leerlingen, die je vervolgens als docent gebruikt om die beslissing te kunnen nemen. Om het concreter te maken een voorbeeld. Een docent wil graag weten welke voorkennis leerlingen al hebben over bepaalde lesstof. Om dit te doen stelt de docent de leerlingen

een aantal vragen, die zo zijn gesteld dat de voorkennis goed kan worden bepaald. Op basis van de antwoorden van de leerlingen neemt de docent een beslissing welke lesstof extra aandacht behoeft en welke al bekend is. Dit is een *formatieve* beslissing: het gaat immers om informatie die de docent gebruikt om de instructie en het verdere leerproces aan te passen. De docent had deze beslissing niet kunnen nemen als de vragen ter voorkennisactivatie niet waren gesteld. Het is dus vooral belangrijk om vragen te formuleren die ontlocken wat leerlingen nog *niet* weten of kunnen. Als het gaat om een beslissing met een *summatieve* functie, gaat het over een uitspraak of een leerling daadwerkelijk heeft geleerd. Er is dus informatie nodig of de leerling iets *wel* kan of begrijpt, en dan niet op één specifiek moment, maar ook voor de lange termijn. 'Geleerd' verwijst dus niet naar een momentane prestatie, maar naar een structurele verandering in het langetermijngeheugen. Om een dergelijk verandering te bewerkstelligen zal veel tijd en oefening nodig zijn. Daarom stelt Daisy Christodoulou¹ dat we toetsen met een summatieve functie niet te vaak moeten doen, anders vragen we van leerlingen te snel dat ze een marathon moeten lopen, terwijl ze tijd en oefening nodig hebben deze te kunnen leren lopen. Anders gezegd: het is belangrijk te onderzoeken of we nu in het onderwijs niet te veel momentane prestaties aan het toetsen zijn met een kortetermijneffect, in plaats van te investeren in een visie op leren gericht op een langetermijneffect.

Hier ligt dan ook een mooie brug naar de begrippen kwantitatieve en kwalitatieve toetsen. Ook hier geldt weer: het zijn geen eigenschappen van de toets. Kwantitatief en kwalitatief verwijst naar de wijze waarop uitdrukking of waarde wordt gegeven aan de beslissing die wordt genomen over het leren van leerlingen. Uit onderzoek weten we dat kwantitatieve cijfers en scores weinig effect hebben op leren en leerlingen ook geen informatie geven hoe verder te gaan. Cijfers kunnen relevant zijn wanneer het gaat om het nemen van een summatieve beslissing, maar dan vooral als een soort van 'benchmark', bijvoorbeeld dat een 8 voor eindexamen Engels even veel waard is op school X als op school Y en Z. Als ik een reeks cijfers zie in Magister, mis ik het verhaal achter die cijfers én het verhaal dat betekenis geeft aan de samenhang tussen die cijfers. Wat betekent bijvoorbeeld een 4 op moment 1 voor wiskunde en een 8 op moment 2? En wat als het omgekeerde het geval is? Werden dezelfde doelen en inhoudsdoelen getoetst? Mag er worden gesproken van vooruitgang dan wel achteruitgang? Zijn het vergelijkbare momenten van informatieverzameling geweest? Wat de moeilijkheid gelijk? Je begrijpt dat alleen zo'n beeld van een reeks cijfers veel vragen oproept.

Om voor een leerling een volgende stap in leren te kunnen vormgeven is er dus een verhaal met rijke informatie en feedback nodig die de leerling in beweging zet. Ervaring en onderzoek leert dat de leerlingen zich nu vooral richten op het cijfer, ook al wordt er rijke feedback gegeven. Door het geven van betekenis aan de waarde van momentane prestaties (die dus zeer kunnen fluctueren), is het mogelijk 'lijm' aan te brengen tussen de vele momenten waarop je als docent tijdens je onderwijs informatie

verzamelt over je leerlingen. Dit biedt vervolgens kansen om de informatie uit deze vele momenten te benutten voor het komen tot een finale, summatieve beslissing of een leerling over mag, en uiteindelijk een diploma van waarde behaalt. We zien nu echter dat we die beslissingen nemen op basis van het optellen, wegen en middelen van een aantal cijfers/scores uit momentane prestaties. De vraag is of dat de beste basis is voor het nemen van zorgvuldige beslissingen over onze leerlingen.'

Hoe kijk jij aan tegen de door de raad betoogde disbalans in toetsmethodieken?

'Met betrekking tot het herstel van de balans tussen de zoëven besproken kwantitatieve en kwalitatieve functie van toetsing, haal ik veel inspiratie uit het werk van correspondent Ontcijferen, Sanne Blauw, die stelt: *'Ik wil cijfers weer op hun plek zetten. Niet op een voetstuk. Niet bij het vuilnis. Maar waar ze horen: naast de woorden.'*² Voor het voortgezet onderwijs zou dat kunnen betekenen dat we cijfers wellicht niet meteen in de ban doen, maar ze wel meer betekenis geven, zowel op het niveau van een individuele toets als op het niveau van een toetsprogramma.³

Verder heeft het zoeken naar de juiste balans altijd te maken met de fundamentele vraag: waartoe dient jouw onderwijs? Je kunt alleen spreken van disbalans in samenhang met visie op leren, de aanwezige deskundigheid van het team, de gedeelde kernwaarden over onderwijs en het docentschap, de aanwezige middelen in termen van tijd en geld en op basis waarvan wordt bepaald of de visie en doelen van de school worden gerealiseerd. Van disbalans is pas sprake als de consistentie tussen de visie en de wijze van toetsen ontbreekt. Ter illustratie: wanneer een school als visie heeft vooral kortetermijnleren te stimuleren, dan is een toetsaanpak met vele toetsen met een summatieve, momentane functie legitiem en is er sprake van een balans tussen visie op leren en de wijze van toetsing. Er is pas sprake van een disbalans als we deze wijze van toetsing handhaven bij een visie op langetermijnleren of een visie waarin we toetsing vooral inzetten als een kans om te leren. Of er al dan niet sprake is van een disbalans, is iets dat ik graag in de context van een specifieke school en de condities aldaar onderzoek.'

'Toetsen en examens worden vaak niet door leraren en docenten zelf gemaakt, maar extern vormgegeven. Hierdoor voelen leraren en docenten zich minder eigenaar van hun onderwijspraktijk.' Hoe belangrijk is het dat leraren hun eigen toetsen maken? Waarom?

'Als je toetsing - in lijn met Dylan Wiliam⁴ - definieert als een procedure die je nodig hebt om zorgvuldige conclusies te trekken over het leren en de voortgang van leerlingen, is het heel belangrijk dat een docent hier verantwoordelijkheid in draagt en aanspreekbaar op is. Toetsing gericht op het nemen van formatieve beslissingen is in die zin meer een proces van onderwijzen (*responsive teaching*), waarbij de docent een aantal activiteiten inzet om leerlingen te activeren en informatie verzamelt over waar zij staan, en op basis daarvan bepaalt wat nog nodig is om

¹ https://youtu.be/8kGRS6BX_Kw

² <https://decorrespondent.nl/3323/dit-is-wat-ik-wil-onderzoeken-als-jullie-nieuwe-correspondent-ontcijferen/1460324480087-3dfca647>

³ <http://dnhavo.blogspot.com/2016/09/toetsing-en-cijfers-vanaf-2016.html?m=1>

⁴ <http://toetsrevolutie.nl/?p=475>

de beoogde leerdoelen te behalen. De wijze van toetsing is dan zo verankerd in het dagdagelijkse onderwijs, dat het niet anders kan dan een docent dat doet en niemand anders.

Ik maak daar echter graag twee opmerkingen bij. De eerste is dat het belangrijk is dat docenten als team in de vorm van een *lerend gezelschap* van elkaar leren en ook weten hoe zij tot zorgvuldige beslissingen – zowel formatief als summatief – komen. Al in 1966 betoogde Adriaan de Groot in zijn boek *Vijven en Zessen* dat een leerling niet de dupe mag worden van de subjectiviteit van een individuele docent, en dat beslissingen navolgbaar en onderbouwd moeten zijn. Zeker bij het verminderen van summatieve beslissingen is het belangrijk naar de samenhang tussen een aantal toetsmomenten te kijken, bijvoorbeeld binnen een bepaalde vaksectie met gedeelde leerdoelen over een langere periode. Een tweede opmerking is dat een school ook werkt aan externe validering van toetsen met een summatieve functie. Dit is iets wat we in het hbo al veel doen. Dit kan bijvoorbeeld door als docenten van meerdere scholen samen schoolexamens te ontwikkelen en te valideren. Deze manier van werken laat kwaliteitsbewustzijn zien. Ook geeft het uiting aan het feit dat toetsing niet alleen wordt bepaald door een specifieke school of docent, maar ook een locatieoverstijgende waarde heeft.

De Raad pleit ervoor dat de overheid heldere kaders schept en daar ook een beetje rolvast in is. Dus niet steeds nieuw beleid uitrolt of staand beleid herroept. Breng toetsing beter in lijn met onderwijsdoelen, onderwijsinhoud en onderwijsmiddelen om versmalling van het onderwijscurriculum tegen te gaan. Maak meer gebruik van formatieve toetsing en van kwalitatieve toetsmethoden. Zorg voor leraren met meer toetsdeskundigheid. Hoe lees jij die adviezen?

‘Dit advies verwijst naar een pleidooi voor meer doelgericht in plaats van toetsgericht denken en werken. Alhoewel voor leerlingen de toets vaak leidend is en sturend voor hun leergedrag, zijn het de leerdoelen en de helderheid over succescriteria die leidend zijn voor het ontwerp van het curriculum en het bepalen welke informatie nodig is om te komen tot zorgvuldige beslissingen over (langetermijn)leren van leerlingen. Het garanderen van een goede aansluiting tussen doelen, leeractiviteiten en toetsing (*constructive alignment*) is echter makkelijker gezegd dan gedaan.

Ik zie twee belangrijke uitdagingen voor scholen. De eerste uitdaging is de durf datgene wat je belangrijk vindt wél te onderwijzen maar niet meer ‘hard’ te toetsen. Toetsgericht denken en werken kan leiden tot ‘teaching to the test’, dat leidt tot een versmalling van het curriculum tot alleen datgene wat er in de toets wordt gevraagd. Gevolg is gebrek aan validiteit omdat niet alle te beoordelen aspecten in de toets aan bod komen of dat er zaken worden getoetst die niet er toe zouden mogen doen, bijvoorbeeld begrip lezen in een rekentoets. Toetsing (bijvoorbeeld het eindexamen) gaat het curriculum sturen en niet *dienen*. Uitgaande van de visie dat het onderwijs er is om je te laten (leren) leren en niet om je te leren hoe je toetsen kunt maken, betekent het idee loslaten dat alles meetbaar moet zijn. Maar zolang de toetsing en daarmee samenhangende indicatoren - excuses voor het woord - zoals bijvoorbeeld het aantal zittenblijvers, opstroom- en afstroomcijfers, slagingspercentages en rendementen, een aanzienlijke rol blijven spelen in het vaststellen van onderwijskwaliteit (of datgene is

waar scholen op worden ‘afgerekend’), zullen deze indicatoren ook voor scholen belangrijk blijven. Het gevaar is dat *Campbell's Law* zijn intrede doet: toetsresultaten worden het doel van onderwijzen, en omdat deze resultaten hun waarde verliezen zullen zij ongewenste effecten hebben op het leerproces. Toetsresultaten worden dan uiteindelijk betekenisloos.

Ik denk wel dat we ervoor moeten waken dat we niet teveel focussen op de vorm.

De tweede uitdaging is niet meer normgerichte data over leerlingen leidend laten zijn voor het nemen van beslissingen over leerlingen, maar vooral de focus te leggen op de vraag welke informatie nodig is om een beslissing te kunnen nemen of een leerling een leerdoel heeft behaald of wat er nodig is het leerdoel te behalen. Dit betekent dat toetsen dus niet achteraf worden geconstrueerd, maar dat docenten vanuit de doelen bepalen welke informatie zij tijdens het onderwijs nodig hebben van hun leerlingen om beslissingen over leren zo zorgvuldig mogelijk te nemen. Toetsing is de start, niet de afsluiting. En zolang er geleerd wordt, staat er voor de leerling nog niks op het spel. Zoals ik al eerder betoogde, is het een interessant gedachtenexperiment wat het zou betekenen als het eindexamen alleen nog maar een ceremonieel karakter heeft en niet meer leidt tot een verrassing op de dag van het telefoontje of je je diploma hebt gehaald.

Hoe kunnen scholen aan de slag gaan met meer formatieve toetsing en kwalitatieve toetsing?

‘Een goed begin is te kijken wat er al allemaal gebeurt in de dagelijkse onderwijspraktijk. Mijn ervaring is dat docenten al veel activiteiten organiseren waar we nu het label ‘formatief’ aan hangen. Eerst kijken waar je staat is belangrijk. Ten tweede vind ik het altijd interessant te kijken wat er kan worden ‘opgeruimd’ in het curriculum⁵. Wat doen we misschien te veel, te vaak of te intensief? Waar kan het een onsje minder? Welke vergaderingen kunnen we schrappen? Eerst lucht en ruimte creëren is belangrijk voordat je aan iets nieuws begint of met nieuwe dingen ervaring gaat opdoen. Daarnaast ben ik erg voorstander van een type aansturing waarin zowel onderwijskundig leiderschap (de inhoud, waar gaan we naar toe en waarom?) als de managementkant (het organiseren van de processen om visie en doelen te realiseren) is geborgd. Tot slot is het zinvol te investeren in een *formatieve denk- en werkcultuur*⁶, waarbij je als team op dezelfde wijze kunt werken zoals je dat met je leerlingen doet: waar willen we als school naar toe? Waar staan we nu? Wat hebben we nodig om daar te komen? Het kan heel inspirerend werken om met elkaar na te denken over haalbare doelen op de korte, middellange en lange termijn, de succescriteria (waaraan zien we dat we onze doelen hebben behaald?) en over activiteiten en rollen die nodig zijn de doelen te behalen en in welk tijdsplan. Ik vind het werk van Stichting LeerKRACHT een mooi voorbeeld hoe je formatief werken aan je schooldoelen met een vast ritme kunt organiseren.

Wat betreft onderzoek zie ik nog vele kansen hoe we als scholen, onderzoekers/lectoraten van hogescholen en andere relevante

partners kunnen samenwerken. Het biedt mogelijkheden om complementaire perspectieven met elkaar in contact te brengen en elkaar te versterken. Een onderzoeker of onderwijskundige kan niks betekenen zonder een rijke context van een school. Het Platform Leren van Toetsen (lerenvantoetsen.nl) en de projecten binnen het LOF (LerarenOntwikkelFonds) zijn mooie voorbeelden van hoe er al verbanden worden gelegd.

Tim is biologieleerleraar en wil zijn eigen toetspraktijk zo inrichten dat deze meer bijdraagt aan de kwaliteit van leren. Wat zijn belangrijke vragen die hij zichzelf (en anderen) kan stellen?

- Wat zijn mijn drijfveren geweest om docent te worden?
- Wie zijn mijn leerlingen?
- Wat zijn de doelen die ik mijn leerlingen wil laten bereiken?
- Wat zijn de doelen die ik de les van morgen heb, over zes weken en over een jaar?
- Wat betekent dat voor de inhouden die aan de orde moeten komen?
- Welke activiteiten moet ik organiseren om te bepalen wat mijn leerlingen al van deze inhouden weten en begrijpen?
- Welke informatie heb ik nodig om mijn leerling actief te houden?
- Welke vragen moet ik stellen om te bepalen of de leerlingen de lesstof hebben begrepen?
- Waaraan kan ik zien dat mijn feedback mijn leerlingen verder helpt/aan het denken zet?
- Hoeveel en welke informatie heb ik nodig om vast te stellen dat leerlingen hebben geleerd en op welke wijze communiceer ik dat naar hen?

Waar zal hij wellicht tegenaan lopen en aan welke oplossingsrichtingen moeten we dan denken?

'Leerlingen in het voortgezet onderwijs zullen verwachtingen hebben over de wijze van toetsing, en onzekerheid ervaren als bewijs van leren niet meteen wordt uitgedrukt in een cijfer. Het is belangrijk dat Tim goed blijft werken aan een doelgerichte benadering en daarbij zorgt dat leerlingen snel succes ervaren. Dit kan doordat zij positieve feedback krijgen op het behalen van leerdoelen, of omdat zij al tijdens de les veel meer worden 'gezien' en geactiveerd.

Een ander probleem kan zijn dat ouders vragen om cijfers, omdat ook hun dat houvast geeft in hoe hun kinderen het doen op school. Ook hier kan Tim werken aan een goed evenwicht tussen het cijfer en het verhaal erachter. Het inzichtelijk maken van groei is hierin belangrijk.

Een derde probleem kan zijn dat Tim zich alleen voelt in zijn weg naar meer formatief werken. Als veel collega's een andere manier van werken hebben, zullen leerlingen meer weerstand hebben en verward raken door de verschillende wijze van toetsing. Dan is het belangrijk dat Tim tijdig bij zijn teamleider aangeeft wat hij nodig heeft om deze beweging verder door te zetten en te benoemen dat teamleren belangrijk is om tot

echte verandering te komen. Als er binnen de school geen of nauwelijks ruimte is, is het een overweging waard om - mogelijk binnen dezelfde scholenkoepel - een school te zoeken die beter bij de visie van Tim past.'

Resumerend, hoe sta je tegenover het advies van de Onderwijsraad en hoe liggen de kaarten: heeft het onderwijs de macht/kracht om in deze lijn te gaan handelen?

'Waar ik heel blij mee ben, is dat toetsing, en dan met name de formatieve functie, nu een structureel onderwerp van gesprek is binnen alle onderwijssectoren. Inmiddels zijn er tal van initiatieven die uiting geven aan het feit dat scholen toe zijn aan een opfrisbeurt van hun wijze van toetsing. Het advies van de Onderwijsraad vind ik belangrijk omdat het toetsing integraal benadert en beschrijft voor alle onderwijssectoren. Wat ik tevens sterk vind, is dat het advies op hoofdlijnen benoemt dat het steeds gaat om het werken aan balans, omdat disbalans ertoe leidt dat iets of iemand uit evenwicht raakt. Tegelijkertijd blijven er genoeg 'Maar hoe dan?'-vragen liggen. Deze zijn met dit advies niet te beantwoorden, maar ik zie wel dat door de opkomst van verschillende Facebookgroepen, de leernetwerken formatief evalueren, en vele andere vormen van kennisdeling scholen elkaar steeds meer weten te vinden. Ik denk wel dat we ervoor moeten waken dat we niet teveel focussen op de *vorm*, bijvoorbeeld het aantal summatieve toetsen verminderen of op een losse wijze een aantal werkvormen of technieken inzetten. Het is vooral belangrijk eerst aandacht te schenken aan het samen beantwoorden van de vraag welke *functie* toetsing in het voortgezet onderwijs dient op basis van een visie op leren en wat ervoor nodig is om dat doel te realiseren. René Kneyber benoemt dit in zijn blog⁷ *functioneel toetsen*. Als scholen te snel maatregelen nemen in het stoppen met cijfers of het terugdringen van toetsen met een summatieve functie, zonder eerst goed na te denken wat een andere visie op leren en toetsing betekent in de gehele planning en organisatie, de samenhang in het curriculum en de communicatie naar ouders en leerlingen, is de kans op terugval naar het oude groot. Tot slot hoop ik dat we elkaar veel meer gaan vinden in de aansluiting binnen de onderwijsskolom. Centrale vragen zijn dan: Hoe kan de wijze van toetsing tussen primair, voortgezet, middelbaar en hoger (beroeps)onderwijs beter op elkaar aansluiten, zodat er optimale onderwijsloopbanen voor leerlingen worden gecreëerd? En wat vraagt dat aan leiderschap en professionalisering van docenten? Ik hoop dat wij als onderwijslectoren nog lang een bijdrage kunnen en mogen leveren aan het beantwoorden deze vraagstukken.' ●

► **Dominique Sluijsmans is lector Professioneel beoordelen aan Zuyd Hogeschool. Renske Valk is hoofdredacteur van Van Twaalf tot Achttien. Lees het advies van de Onderwijsraad op <https://www.onderwijsraad.nl/publicaties/2018/toets-wijzer/item7707>**



⁵ Zie ook onze column hierover op <https://www.professioneelbegeiden.nl/voorjaarsschoonmaak-in-het-toetshuis>

⁶ <https://www.amazon.co.uk/Leadership-Teacher-Learning-Dylan-Wiliam/dp/1941112269>

⁷ <http://toetsrevolutie.nl/?p=512>