

In Boulder, Colorado bevindt zich een dependance van het Freudenthal Instituut. In een groot project is gepoogd de kloof tussen goed en minder goed presterende leerlingen te dichten. Een van de manieren die onderzocht is om dat doel te bereiken, is het leren maken van verbeterde toetsen. **David Webb** en **Truus Dekker** schreven hun bevindingen op.

## Beter toetsen, beter leren?

### Internationale samenwerking

Wist u dat er behalve het Freudenthal Instituut (FI) in Utrecht ook een Freudenthal Institute USA (FI-US) bestaat? FI-US is ontstaan uit de jarenlange samenwerking tussen universiteiten in de Verenigde Staten en Nederland, eerst via de University of Wisconsin in Madison en sinds 2005 via de University of Colorado, te Boulder. Beide auteurs zijn bij het Freudenthal Institute USA betrokken. FI en FI-US werkten in de loop van de jaren samen aan diverse onderzoeksprojecten, het ontwikkelen van een wiskundemethode voor de Middle School (leerlingen in de leeftijd van 10–14 jaar), nascholing van docenten en conferenties. De University of Colorado staat bekend om haar expertise op het gebied van toetsing en het is daarom niet zo gek dat veel samenwerkingsprojecten juist over dat onderwerp gaan. We hebben het dan vooral over toetsen die in de klas gebruikt worden en niet over examens of andere externe toetsen. In dit artikel staat het BPEME-project centraal, Boulder Partnership for Excellence in Mathematics Education, gefinancierd door het schooldistrict Boulder. De uitwisseling van ideeën tussen de onderzoekers uit Nederland en de VS bleek, hoewel om verschillende redenen, van groot belang voor beide landen.

Onderwijsonderzoekers hebben al vaak betoogd dat docenten die het leren van de leerlingen willen verbeteren ook aandacht moeten schenken aan de inhoud van de door hen gebruikte formatieve toetsen. Toetsing, in welke vorm dan ook, dient verschillende doelen. Toetsen die vooral gebruikt worden om het hele proces van leren en onderwijzen te monitoren en bij te sturen noemen we ‘formatief’. De docent bekijkt de antwoorden op de toetsvragen en past eventueel zijn onderwijs aan. Er is misschien meer oefening nodig, of de gegeven uitleg bleek nog niet voldoende. De toetsing is een onderdeel van het leerproces en geen onderbreking daarvan. Feedback, zowel voor de docent als voor de leerling, is belangrijk. Toetsen die vooral gebruikt worden om een (afsluitend) oordeel te vellen over de mate waarin de leerling de stof beheerst heten ‘summatief’. Voorbeelden daarvan zijn de centrale examens.

In eerdere projecten zagen we al hoe belangrijk het is dat docenten hun expertise op het gebied van het ontwerpen van toetsen vergroten, zodat ze meer zelfvertrouwen krijgen bij het nemen van onderwijsbeslissingen. De resultaten van hun leerlingen verbeterden daardoor (Webb et al., 2005). Amerikaanse docenten in Middle Schools hebben soms een beperkte wiskundekennis en ook dat heeft een grote invloed op hun mogelijkheden om goede toetsvragen te zoeken, te ontwerpen en een evenwichtige toets samen te stellen (Webb, 2004).

### Het BPEME-project – ‘closing the gap’

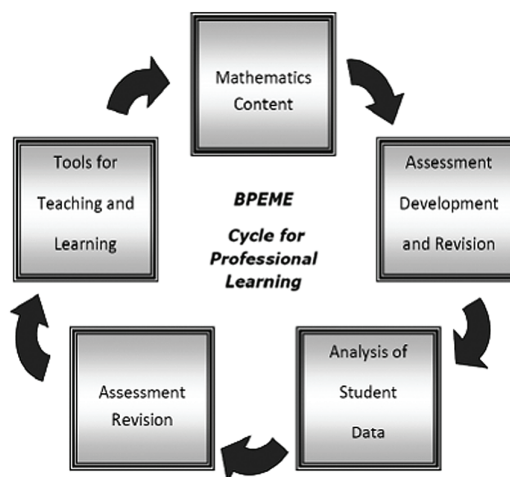


fig. 1 schema nascholing

Het BPEME-project duurde van mei 2006 tot juni 2008. Er waren 32 Middle School-docenten bij betrokken van zes verschillende scholen in wat in de VS een middelgroot schooldistrict heet (totaal ongeveer 28.500 leerlingen). Medewerkers van de universiteit van Colorado (zowel van de toegepaste wiskunde als de onderwijsfaculteit) en van het Freudenthal Instituut waren betrokken bij de planning en uitvoering. Vanuit het schooldistrict deden ‘district instructional leaders’ mee. Dat zijn docenten die lesgeven, maar ook andere taken hebben zoals het begeleiden van leerkrachten in de verschillende scholen van het schooldistrict. De keuze van de scholen was niet wil-

lekeurig, het waren over het algemeen zwak presterende scholen met een groot percentage leerlingen uit groepen met lagere inkomens, en met een groot aantal leerlingen voor wie Engels niet hun eerste taal is. Een van de doelen van dit project was het ‘dichten van de kloof’ tussen deze zwakke leerlingen en de beter presterende leerlingen van de andere Middle Schools in het district. Deze kloof is hardnekkig en bestaat al vele jaren.

De nascholing van de docenten uit het BPME-project bestond uit een cyclus met vijf verschillende fases (zie figuur 1). Het idee om de ontwikkeling, analyse en revisie van formatieve toetsen te gebruiken als onderdeel van nascholing gericht op het verbeteren van de inhoudelijke wiskundekennis van docenten is nieuw.

De cyclus werd verschillende malen doorlopen. De wiskundige inhoud werd beperkt tot de onderwerpen ‘redeneren met verhoudingen’ (ook breuken) en ‘algebraïsche verbanden’. Die beperking had te maken met de beschikbare tijd, maar bovendien zijn dit de gebieden waar de kloof tussen zwakke en goede leerlingen het grootst is.

Het zal u wellicht verbazen dat Amerikaanse docenten nascholing volgen tijdens hun zomervakantie. Dat is echter niet ongebruikelijk in de VS en de docenten worden voor hun aanwezigheid betaald. In 2006 was er een vierdaagse zomercursus, daarna 40 uur nascholing op school gedurende het schooljaar 2006/07, weer een zomercursus in de schoolvakantie in 2007 en 40 uur nascholing tijdens het schooljaar 2007/08. De laatste zomercursus in 2008 duurde een week.

Het onderzoeksdeel van dit project bestond uit het analyseren van alle toetsen die in de loop van het schooljaar door de deelnemende docenten aan hun klassen werden gegeven. De docenten verzamelden al hun toetsen voor dat doel. De toetsen uit het jaar voor het project startte, werden gebruikt als vergelijking (jaar 0). We wilden weten in hoeverre de inhoud van de nascholing van invloed was op de vragen die docenten kiezen, ontwerpen en gebruiken in hun toetsen, en of een verandering in de toetspraktijk van invloed zou zijn op de resultaten van de leerlingen.

## Van informeel naar formeel

### Het ijsbergmodel

Dit model (Webb, Boswinkel & Dekker, 2008) werd gebruikt tijdens de zomercursus van 2007 om het belang te laten zien van toenemende formalisering. Als leerlingen een nieuw wiskundig onderwerp leren, bouwen ze voort op eerder opgedane informele en preformele kennis over het onderwerp. Hieronder staan twee voorbeelden, het eerste werd gebruikt tijdens de nascholing en het tweede werd gemaakt tijdens de ‘ijsbergopdracht’. Groepjes docenten maakten hun eigen ijsberg, gebaseerd op een ana-

lyse van het op hun school gebruikte lesboek. Sommigen van hen merkten op dat hun lesboek voornamelijk leerstof bevatte uit de ‘top van de ijsberg’, dus met een nadruk op formele notaties en procedures. Het ‘drijfvermogen’ van de ijsberg, bestaande uit informele en preformele representaties en wiskundige redeneringen die nodig zijn om de formele wiskunde ook echt te begrijpen, ontbrak daar vrijwel geheel.

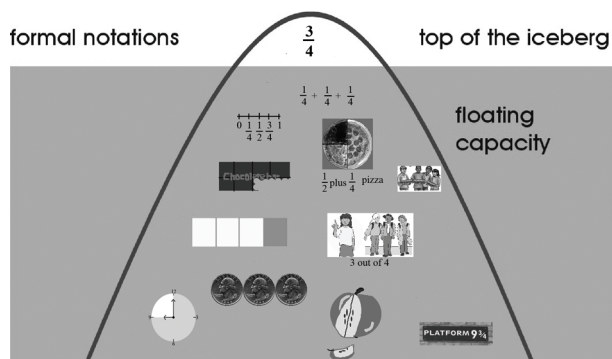


fig. 2 het ijsbergmodel

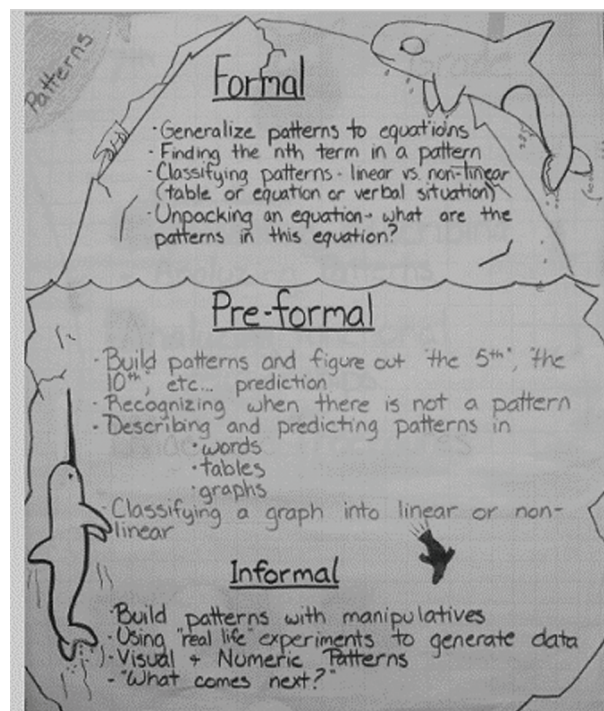


fig. 3 een ijsberg ontworpen door docenten

Omdat onderwijzen en toetsen wederzijds invloed op elkaar uitoefenen, volgde hierna een discussie over de noodzaak van het stellen van vragen die verdergaan dan formele kennis uit het ‘topje van de ijsberg’. Het is immers belangrijk dat leerlingen niet alleen herhalen wat ze precies zo geleerd hebben, maar ook dat ze laten zien dat ze inzicht hebben in de stof en dat ze het geleerde kunnen gebruiken in nieuwe situaties. Docenten kozen geschikte opgaven voor formatieve toetsen uit bestaand materiaal, pasten ze aan of ontwierpen ze zelf. Er werd vooral aandacht besteed aan de veronderstelde voorkennis van hun

leerlingen over het betreffende onderwerp, aan te verwachten informele strategieën van de leerlingen en het mogelijke gebruik van bekende preformele modellen. Die opgaven werden zo samengesteld dat leerlingen door het beantwoorden ervan begrip en inzicht in de stof zouden kunnen laten zien. De docenten bekeken ook aan welke voorwaarden de vragen moesten voldoen opdat alle leerlingen ermee uit de voeten zouden kunnen en welk extra materiaal er bij de voorbereiding nodig zou zijn voor zwakke leerlingen.

Bekijk het door docenten gemaakte voorbeeld in figuur 4 maar eens. Hoeveelheden die beschreven zijn als  $\frac{2}{3}$  deel van iets, kunnen hetzelfde zijn maar ook verschillend, afhankelijk van de bedragen voor zakgeld en het bedrag dat met oppassen verdiend werd. Het is een lastig probleem voor leerlingen en de opgave geeft hen de mogelijkheid om te laten zien dat ze een wiskundige redenering kunnen gebruiken om hun standpunt te verdedigen.

Craig donated  $\frac{2}{3}$  of his allowance to the local pet rescue.  
Paige donated  $\frac{2}{3}$  of her babysitting money to the same pet rescue. Did they both donate the same amount? Explain.



fig. 4 did Craig and Paine donate the same amount?

De vraag die in figuur 5 wordt gesteld is toegankelijker. Leerlingen kunnen hun eigen strategie gebruiken om een probleem op te lossen dat gaat over ‘gemengde’ getallen zoals  $2\frac{1}{2}$ . Deze opgave is geschikt voor leerlingen die niet zo goed zijn in rekenen-wiskunde.



Kenji hiked  $2\frac{1}{2}$  miles in 50 minutes. How far did he hike in the first ten minutes? Show your work.

fig. 5 how far did Kenji hike?

Leerlingen gebruikten bij het oplossen van Kenji’s probleem een dubbele getallenlijn, een verhoudingstabel en een breukenstrook om het verband te laten zien tussen tijd en afstand van de wandeling.

Het ijsbergmodel werd in het project niet alleen maar gebruikt voor het zoeken en ontwerpen van geschikte toetsopgaven. De docenten gebruikten het model ook als een soort vergrootglas om hun lesmateriaal nader te bekijken. Daarbij diende het model als uitgangspunt voor het ontwerpen van leerlijnen met daarin cruciale leermomenten (bijvoorbeeld belangrijke wiskundige concepten, strategieën en representaties) die niet zomaar overgeslagen kunnen worden door de zwakke leerlingen. Aan die cruciale leermomenten werden door de docenten soms ook voorbeelden toegevoegd van de manier waarop bepaalde onderdelen kunnen worden aangepast voor leerlingen die het Engels nog onvoldoende beheersen. En tenslotte werd het ijsbergmodel gebruikt om correctievoorschriften voor toetsen te schrijven, antwoorden van leerlingen te interpreteren en aan te geven hoe het onder-

wijs op basis daarvan zou moeten worden aangepast in de verschillende leerjaren.

In *Adding it Up* (Kilpatrick, 2001), stellen de auteurs dat voor *mathematical proficiency* nodig is:

1. Begrip op conceptueel niveau, begrijpen van wiskundige concepten, bewerkingen en verbanden.
2. Vaardigheid in het uitvoeren van procedures. Geschikte procedures kunnen kiezen. Die procedures flexibel, nauwkeurig, efficiënt uitvoeren.
3. Goede oplosstrategieën kunnen kiezen. Bekwaamheid in het formuleren, beschrijven en oplossen van wiskundige problemen.
4. Redeneren, logisch denken, reflecteren, uitleggen en verantwoorden.
5. Productief zijn, wiskunde kunnen zien als belangrijk, nuttig en de moeite waard; ijverig en doelmatig werken.

Als docenten kennis en vaardigheden van hun leerlingen willen verbeteren, kunnen de toetsen die ze ontwerpen niet uitsluitend bestaan uit opgaven die gericht zijn op procedurele vaardigheden (niveau I, ‘reproductie’, in het piramidemodel verderop in dit artikel). Toch bestonden de toetsen die we verzamelden en analyseerden uit het jaar voor het project begon voor hoogstens 12% uit opgaven die meer dan reproductie toetsen. (Overigens is er in dit geval weinig verschil tussen Amerikaanse en Nederlandse docenten; ook Nederlandse docenten geven vaak weinig opgaven in hun schriftelijke toetsen die verdergaan dan reproductie). Toetsing was daarom een belangrijk thema, zowel tijdens de zomercursussen als tijdens de tweemaandelijksse bijeenkomsten tijdens het schooljaar. De deelnemende docenten bespraken hoe toetsen het leren van hun leerlingen beïnvloeden en ook hoe die toetsen zouden kunnen verbeteren.

## Het piramidemodel

Voor de nascholing over toetsen gebruikten we de toetspiramide (De Lange & Boertien, 1994), die overigens ook in aangepaste vorm in het ‘Assessment Framework’ van het internationale PISA-onderzoek van de OECD voorkomt. De belangrijkste reden voor het gebruik van dit model was dat het kon helpen om een evenwichtige verdeling van verschillende typen toetsopgaven over het gehele jaar te bereiken. En verder helpt het model om docenten te laten zien hoe ze, als ze leren-met-begrip echt belangrijk vinden, ook kunnen toetsen-met-inzicht.

De vaardigheden die nodig zijn voor het vak wiskunde zijn natuurlijk afhankelijk van de eisen die voor het betreffende leerjaar worden gesteld. In de piramide (figuur 6) zijn drie vaardigheidsniveaus te zien. In tegenstelling tot wat de piramidevorm suggereert is er geen hiërarchie tussen de verschillende niveaus, de indeling geeft een ruwe verdeling weer van de hoeveelheid vragen of beter nog, van de hoeveelheid tijd die aan de opgaven op elk van de niveaus wordt besteed. De toetspiramide dient twee doelen. Hij helpt docenten om opgaven op een be-

paald niveau te classificeren en geeft een indicatie van het aantal vragen op elk niveau dat in een evenwichtige (schriftelijke) toets zou moeten voorkomen. Een vuistregel die in de praktijk veel wordt gebruikt, is Niveau I : Niveau II : Niveau III = 3 : 2 : 1. Hieronder volgt een korte beschrijving van de verschillende niveaus in de piramide.

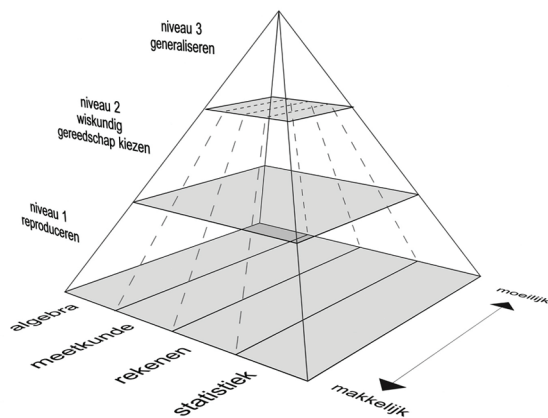


fig. 6 de toetspiramide

**Niveau I: reproductie, gebruik van standaardprocedures en algoritmes, kennis van feiten en definities**

Vragen op niveau I toetsen basiskennis en -procedures die tijdens de lessen veel nadruk kregen en veel zijn geoefend. Zoals het model laat zien, is een vraag op niveau I niet noodzakelijk een gemakkelijke vraag; moeilijke vragen komen op alle niveaus voor. Voor het beantwoorden van een moeilijke niveau I-vraag zijn vaak meerdere stappen nodig. En omdat leerlingen in dat geval meer fouten kunnen maken, is de vraag lastiger.

**Niveau II: verbindingen leggen, wiskundig gereedschap kiezen**

Op dit niveau wordt van leerlingen verwacht dat ze informatie uit verschillende bronnen en uit verschillende wiskundige domeinen kunnen combineren. Ze kiezen zelf het te gebruiken wiskundige gereedschap of de oplossingsstrategie en geven uitleg bij het antwoord. Vragen op dit niveau worden vaak binnen een context gesteld en leerlingen moeten vaak eerst nog een wiskundig model maken van de situatie voor ze het probleem kunnen oplossen. Bedenk overigens dat de context ook uit de wiskunde zelf afkomstig kan zijn.

**Niveau III: mathematiseren, (wiskundig) redeneren, generaliseren en inzicht tonen**

Leerlingen moeten bijvoorbeeld inzien hoe op het oog heel verschillende contexten wiskundig gezien hetzelfde zijn. Voor sommige leerlingen is deze generalisatie heel lastig, ze blijven elk nieuw (context)probleem als helemaal nieuw zien, terwijl ze zouden moeten denken: ‘O ja, dit is net zoiets als de taxi-opgave die we hebben gehad!’ Op dit niveau wordt van de leerlingen verwacht dat ze een wiskundig model van een situatie kunnen maken, bekritisieren en eventueel aanpassen. Ze schrijven wiskun-

dige redeneringen op die hun ideeën ondersteunen en bewijzen bijvoorbeeld dat een bepaalde bewering altijd juist is of nooit juist.

**Enkele resultaten**

Tijdens de nascholing werden de deelnemende docenten betrokken bij een manier van toetsplanning en het ontwerpen van authentieke toetsen die dichtbij hun dagelijkse praktijk staat. Op die manier wordt de toetsing meer iets van henzelf, dat ze direct kunnen toepassen. De docenten gaven in hun evaluatie aan dat ze nu over een groter repertoire aan toetsstrategieën beschikken; wiskunde is toegankelijker geworden voor hun leerlingen en de toetsen ondersteunen het leren van wiskunde.

De eerste analyse van de door de docenten verzamelde toetsen (jaar 0 tot jaar 1), liet zien dat de docenten een steviger basis ontwikkelden om niet alleen kennis maar ook inzicht te toetsen. Zo gebruikten ze bijvoorbeeld meer contextopgaven, een toename van gemiddeld 25,2% aan het begin van het project tot 34,0% na het eerste jaar. Ook het aantal vragen dat verdergaat dan reproductie nam toe, hoewel wat minder sterk, van 15,1% in het begin tot 18,6% na het eerste jaar. Bovendien zagen we hoe docenten misconcepties en fouten uit eerdere toetsen opnieuw gebruikten in de vorm van een vraag om de fout te verbeteren, de leerling in de rol van de docent dus.

Maar de meest in het oog springende verbetering is misschien dat de kloof tussen de prestaties van BPEME- en niet-BPEME-scholen kleiner werd. Uit de gegevens van de externe CSAP-test (Colorado Student Assessment Program) blijkt dat de verbetering van jaar-tot-jaar in toetsresultaten groter is bij de BPEME-scholen. Dat laatste is ook voor de Nederlandse situatie een belangrijk resultaat. Er zijn immers ook in Nederland docenten die menen dat vooral reproductie belangrijk is en dat inzicht pas daarna komt; het BPEME-project laat met concrete resultaten op een onafhankelijke test zien dat aandacht voor leren en toetsen met inzicht de resultaten van de leerlingen verbetert.

In de eerste paragraaf van dit artikel vertelden we dat het BPEME-project belangrijk was vanwege de uitwisseling van ideeën tussen onderzoekers in Nederland en de VS. Voor het Boulder School district was dit al duidelijk uit het bovenstaande; de nascholing maakte gebruik van in Nederland ontwikkelde en beproefde modellen en de CSAP-test die de staat verplicht stelt op de scholen in het district liet een verkleining van de kloof tussen BPEME- en niet-BPEME-scholen zien. Maar ook Nederland heeft belang bij de samenwerking, de ideeën die ontwikkeld zijn bij het FI kunnen nu met grotere aantallen leerlingen en docenten worden getoetst en omdat Nederland geen verplichte toetsen kent zoals de CSAP-test, kunnen de resultaten nu ook met onafhankelijke data worden toegelicht. Formatieve toetsing is een belangrijk thema voor de constructivistische leertheorieën, zoals bijvoorbeeld ‘realis-

tisch wiskundeonderwijs', en de resultaten van het BPE-ME-project lijken dit belang te onderstrepen.

Tenslotte nog dit. Het uitvoeren van nascholing, gebaseerd op de ontwikkeling, analyse en revisie van formatieve toetsen om zo de inhoudelijke wiskundekennis van docenten te verbeteren en uit te breiden, bleek een goed idee waarmee recht werd gedaan aan de al bestaande expertise van de docenten. Als je weet wat je wilt toetsen en op welke manier je dat wilt doen geeft dat inzicht in wat daarvoor nodig is dus inzicht in wat de leerlingen aan leerstof en lessen nodig hebben om de toetsdoelen te kunnen bereiken.

Truus Dekker was tot haar pensionering in 2008 bestuurslid van FI-US. David Webb is nog steeds directeur van FI-US.

Truus Dekker, Maarssen  
David Webb, FI-US

## Literatuur

Dekker, T. (2004). Toetsen op een hoger niveau? *Nieuwe Wiskrant*, 23(3), 33-36.

Dekker, T. (2004). Een evenwichtige toets maken. *Nieuwe Wiskrant*, 23(4), 28-30.

- Dekker, T. (2007). A Model for Constructing Higher Level Classroom Assessments. *Mathematics Teacher*, 101(1), 56-61.
- Dekker, T., & Webb, D.C. (2003). Toetsen op niveau. *TDB, Tijdschrift voor Didactiek der  $\beta$ -wetenschappen*, 20(2), 91-104.
- Lange, J. de, & Boertien, H. (1994). *Model voor wiskundetoets nationale optie TIMSS 1995*. Enschede: Universiteit Twente, OCTO (unpublished)
- Webb, D.C., Boswinkel, N., & Dekker, T. (2008). Beneath the Tip of the Iceberg: Using Representations to Support Student Understanding. *Mathematics Teaching in the Middle School*, 14(2), 110-113.
- Webb, D.C., Romberg, T.A., Burrill, J., & Ford, M.J. (2005). Teacher Collaboration: Focusing on Problems of Practice. In T.A. Romberg & T.P. Carpenter (Eds.), *Understanding Math and Science Matters* (pp. 231-251). Mahwah, NJ: Lawrence Erlbaum Associates.
- Webb, D.C., Romberg, T.A., Dekker, T., Lange, J. de, & Abels, M. (2004). Classroom Assessment as a Basis for Teacher Change. In T. A. Romberg (Ed.), *Standards-Based Mathematics Assessment in Middle School: Rethinking Classroom Practice* (pp. 223-235). New York: Teachers College Press.

## Klassenposter 'Wiskunde in Perspectief'

Als wiskundige kun je aan het werk als leeraar of onderzoeker. Maar daarnaast werken wiskundigen ook bij banken en verzekeringsmaatschappijen, bij ziekenhuizen of bij de overheid. Wiskunde in Perspectief is een project uitgevoerd onder toezicht van de Taakgroep Landelijke Wiskunde PR van het KWG, met als doel de minder bekende beroepen van wiskundigen beter zichtbaar te maken. Dit gebeurt door middel van de website [www.wiskundeinperspectief.nl](http://www.wiskundeinperspectief.nl), gevuld met interviews met wiskundigen, en een begeleidende klassenposter voor in het (wiskunde)lokaal. Deze vindt u bij deze uitgave van de *Nieuwe Wiskrant*.



Op [www.wiskundeinperspectief.nl](http://www.wiskundeinperspectief.nl) vertellen wiskundigen over hun beroep, en hoe ze daar wiskunde bij nodig hebben. Zo licht Nienke Valkhoff haar werk toe met het volgende voorbeeld: 'Stel je voor dat je een postkantoor hebt. Je denkt na over hoeveel klanten er binnenkomen. En als je ze binnen een bepaalde tijd van dienst wilt kunnen zijn,

hoeveel medewerkers je dan nodig hebt. Dat zou je wiskundig kunnen gaan onderzoeken'. Zij werkt als Manager training and education bij Incontrol Enterprise Dynamics, een bedrijf dat simulatiesoftware ontwikkelt.

De geïnterviewden vertellen ook over hun overwegingen bij de studiekeuze. Zo vertelt Kit Roes, werkzaam als Vice President Global Clinical Information bij Schering-Plough (voorheen Organon): 'Wat ook heeft meegespeeld, is dat het heel breed toepasbaar is. Er zitten in allerlei hoeken wiskundigen, of je het nou hebt over het ontrafelen van genetische basis van ziekten, IT-toepassingen (Google!) of het verzekeringswezen, het is echt heel breed'.

Dat het vak breed toepasbaar is, blijkt ook uit de bijzondere stage die Marry Ouwehand deed in Bolivia: 'Er is daar een gigantisch afvalprobleem. Wij hebben ons bezig gehouden met het optimaliseren van de routes van vuilniswagens. (...) We hopen dat ons verslag heeft bijgedragen, zodat de vuilniswagens beter ingepland kunnen worden'. Tegenwoordig werkt Ouwehand als Analyst Quantitative Analysis bij PriceWaterhouseCoopers.

De wiskundigen op deze website laten zien dat hun opleiding lang niet altijd gepaard gaat met het imago van 'stoffige professor'. Wiskundigen blijken overal te zitten, zelfs in de politiek: ook Jan Terlouw, jeugdboekenschrijver en politicus, heeft in een ver verleden wiskunde gestudeerd. Hij vertelt: 'Ik ben in de eerste plaats wetenschapper, pas in de tweede plaats politicus. (...) Ik kijk ook als exacteling. Ik denk eigenlijk dat ik dat nooit echt kan laten'.