

In het vorige nummer van de *Nieuwe Wiskrant* werd gerapporteerd over TIMSS, de vergelijkende internationale landenstudie over het onderwijs in de exacte vakken. Prettig voor de politiek was dat Nederland hier aardig goed uitkwam. Maar **Jan de Lange** kraakt enkele kritische noten over wat TIMSS nu eigenlijk méét.

De betekenis van TIMSS voor het Nederlandse wiskundeonderwijs

TIMSS: voer voor de politiek

TIMSS is vooral van belang voor politici, laat dat vooraf gezegd zijn. En die hebben zich ook danig geroerd. Vooral, uiteraard, in de Verenigde Staten, de grootfinancier van het onderzoek naar de prestaties op het gebied van wiskunde- en science-onderwijs. Daar was het oordeel over de resultaten voor de Verenigde Staten uitermate negatief, zowel in de pers, politiek als wetenschap. Tja, Amerika kwam er ook wel erg mager vanaf, als middenmoter op zijn best. En dat terwijl president Bush nog wel had aangekondigd dat de Verenigde Staten in 2000 nummer 1 zouden zijn in wiskunde en science. Pittige rapporten en commentaren fileerden genadeloos het Amerikaanse onderwijs, vaak de spijker op de kop slaand. Maar hoe het verder moet, is een vraag die nauwelijks beantwoord wordt.

Dat geldt voor veel andere landen. Duitsland dat er marginaal beter uitkomt dan de Verenigde Staten (door bijvoorbeeld een iets diepere behandeling van concepten en iets minder oppervlakkige boeken, alhoewel de testscore dat niet weerspiegelt) verkeert ook in rep en roer. Maar ook Duitsland weet niet goed hoe en wat te veranderen. Dat heeft mede te maken met het feit dat Duitsland niet erg 'veranderingsgezind' is ingesteld – in tegenstelling tot bijvoorbeeld Nederland.

Het aardige, en ook nogal verontrustende, is het feit dat vrijwel niemand verder kijkt dan de testcores, of nog 'enger', de volgorde van de landen. Terwijl je, als je echt wat wilt leren van TIMSS, je toch eerst eens moet kijken naar wat TIMSS nu eigenlijk voorstelt.

In de vorige *Nieuwe Wiskrant* kon u lezen hoe de Nederlandse onderzoekers de testcores beoordeelden en welke gevolgtrekkingen ze trokken. Alhoewel auteur dezes een zekere verantwoordelijkheid draagt voor TIMSS als lid van de Nederlandse begeleidingscommissie, zou ik toch graag enige kritische noten willen kraken. Om daarna te kijken of er iets te zeggen valt dat relevant is voor het veranderingsproces waarin het Nederlandse wiskundeonderwijs verkeert.

Wat meet TIMSS?

Niet de kwaliteit van het wiskundeonderwijs. Alhoewel er veel lovende woorden waren voor de statistisch-methodologische opzet van het onderzoek, zijn er voldoende kanttekeningen te plaatsen die aantonen dat de betekenis van TIMSS zeer beperkt is. Als inleiding op het eerste punt kan opgemerkt worden dat er in de literatuur nogal eens wordt gekeken naar verschillende soorten curriculum. Want wat curriculumontwerpers voor ogen hebben, komt natuurlijk nooit zo in de klas. Eerst gaan auteurs er hun interpretatie aan geven, vervolgens gaat de leraar eigen keuzes niet uit de weg en ten slotte hebben toetsontwerpers ook weer hun eigen gekleurde bril op. Welnu, in het rapport wordt geen enkel verband gelegd tussen het 'beoogde' curriculum (door curriculumontwikkelaars), het 'ervaren' curriculum (bijvoorbeeld door schoolboek-schrijvers), het 'geïmplementeerde' curriculum (door docenten), het 'getoetste' curriculum en de TIMSS-toets.

Wat men op zijn minst zou kunnen verwachten, is dat de TIMSS-toets naadloos aansluit bij bijvoorbeeld het 'getoetste' curriculum. Wat men daarbij dan moet hopen, is dat het 'getoetste' curriculum dicht bij het 'beoogde' curriculum ligt. Het zal de lezer duidelijk zijn dat er per land waarschijnlijk een heel boek te schrijven zou zijn over de verschillen tussen al die curricula. Maar evident is dat de verschillen tussen de TIMSS-toets en de curricula per land zeer zullen verschillen (en bijvoorbeeld in de Verenigde Staten sluit de TIMSS-toets bijna naadloos aan bij andere toetsen). Voor Nederland kan zonder aarzeling gesteld worden dat de verschillen zeer groot zijn. Dat was ook de reden waarom besloten is voor ons land een nationale optie te starten, waarover in de vorige *Nieuwe Wiskrant* uitvoerig is bericht. Kortom, een 'normering' op dit aspect zou op zijn plaats zijn, of een grote terughoudendheid in volgorde lijsten.

Een tweede niet onbetekenend aspect is het volkomen ontbreken van de cultuur waarin het onderwijs zich afspeelt. In de onderzoeksliteratuur daarentegen wordt juist steeds meer de rol van de mate van inbedding van het on-

Table M.2.1. Testing Minutes Allocated by Content and Performance Category for Smallest Content Components (Population 2 Mathematics)

	Performance Reporting Categories						
	Knowing	Routine Procedures	Complex Procedures	Solving Problems	Justifying and Proving	Communicating	
Performance Category	MC 35 SA 2	MC 27 SA 6	MC 25 SA 6	MC 42 SA 10 ER 20	ER 15	ER 10	MC 129 SA 24 ER 45
Totals	Tot. 37	Tot. 33	Tot. 31	Tot. 72	Tot. 15	Tot. 10	Tot. 198

Indeling van de toetsitems met tijdsaanduiding naar de verschillende inhoudsgroepen.

MC = multiple choice (1 min per item), SA = short answer (2 min per item), ER = extended response (5 min per item)

derwijs in de cultuur van een land of beschaving benadrukt. De cultuur van de landen in het verre oosten (Japan, Hongkong, Singapore, Korea) is een geheel andere dan die van westelijke landen.

Het is dan ook nauwelijks verbazingwekkend, alhoewel het nergens wordt opgemerkt (dat is wel verbazend!), dat de ranglijsten duidelijk cultuurbepaald zijn. Zo bestaat de kopgroep exclusief uit de Aziatische tijgers, gevolgd door kleine West- en Oost-Europese landen (in die volgorde) met daarachter de grote Oost-Europese landen, vervolgens de Commonwealth landen, met daar weer achter de Scandinavische landen en als bijna hekkensluiter de Zuid-Europese landen. De plaats van Nederland wijkt ook geenszins af van dit patroon, niks bijzonders dus. Dus?

Een derde punt van kritiek betreft de onevenwichtigheid in de toetsopgaven. De tabel laat zien hoe de items zijn ingedeeld naar 'inhoudsgroepen' (zoals iedere toetsenbakker die hanteert in een of andere vorm) die bij TIMSS Performance Reporting Categories heten. Uit de bovenstaande tabel blijkt dat de 'lagere', veelal routinematige categorieën, zwaar oververtegenwoordigd zijn.

Hierbij moet nog bedacht worden dat deze tabel de tijd weergeeft per categorie en de 'hogere' items een langere duur hebben. Dus het aantal items dat wat minder platte doelen toetst, is nog veel minder dan de tabel suggereert. (In de twee categorieën Justifying and Proving en Communicating samen zitten bijvoorbeeld slechts vijf items.) Ook is er geen evenwicht in de soorten 'formats' dat gehanteerd wordt: in de Nederlandse situatie was er sprake van 429 multiple choice items, 43 korte antwoordvragen en 29 open vragen.

Ten slotte kunnen uit inhoudelijk oogpunt vragen gesteld worden bij de 'performance categories', zoals de volgende twee voorbeelden laten zien.

Het eerste voorbeeld, een nogal triviaal item (voor I.3/14

jarigen!) wordt gecategoriseerd als 'Problem Solving' terwijl het tweede voorbeeld (een wel zeer eenvoudig schaalprobleempje) wordt ingedeeld bij 'Complex Procedures'.

Measuring the width of a room

Four children measured the width of a room by counting how many paces it took them to cross it. The chart shows their measurements.

Who had the longest pace?

- A. Stephen
- B. Erlane
- C. Ana
- D. Carlos

Name	Number of Paces
Stephen	10
Erlane	8
Ana	9
Carlos	7

Performance Category: Solving Problems

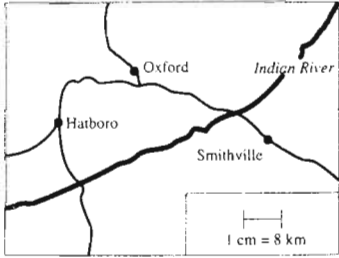
Voorbeeld 1: Problem Solving

Op deze manier wordt de schijn gewekt dat er sprake is van echt 'probleem oplossen' en het toepassen van 'complex procedures'. Zaken die in veel landen in het huidige onderwijs als erg belangrijk worden gevonden, terwijl er in werkelijkheid nauwelijks sprake van is.

Samengevat kan gesteld worden dat TIMSS niet meer meet dan vooral 'lagere' doelen, met voornamelijk multiple choice vragen, dat TIMSS maar matig aansluit bij het curriculum, een discutabele doelenmatrix hanteert en geen rekening houdt met de lokale cultuur.

Distance on a map

One centimeter on the map represents 8 kilometers on the land.



About how far apart are Oxford and Smithville on the land?

A. 4 km
 B. 16 km
 C. 35 km
 D. 50 km

Performance Category: Using Complex procedures

Voorbeeld 2: Complex Procedures

Hoe representatief is TIMSS?

Na het voorgaande zal duidelijk zijn dat er niet al te veel waarde aan de uitkomsten van TIMSS moet worden gehecht. TIMSS meet echt niet de kwaliteit van het Nederlandse wiskundeonderwijs. Dat blijkt ook uit een heel interessant deelaspect dat specifiek voor Nederland geldt en waarover de Twentse onderzoekers in de vorige *Nieuwe Wiskrant* berichtten: de nationale optie.

Nog even kort de essentie van de nationale optie samengevat: omdat de internationale toets het Nederlandse wiskundeonderwijs slecht representeert, is er gekozen om een aantal Nederlandse items te ontwikkelen die wel passen bij ons curriculum. Een groot aantal (representatief) Nederlandse leerlingen heeft gewoon de internationale toets gedaan. Daarnaast heeft een kleiner aantal (ook representatief) de 'gemengde' toets gedaan: een aantal Nederlandse items gekoppeld (ter ankerings) aan een aantal internationale items. De verwachting – vanuit methodologisch oogpunt – was dat de ankeritems (de doorsnede uit beide versies) even goed gemaakt moesten worden in beide versies. Beide zijn immers representatief. Wie schetst de verbazing toen niets minder waar bleek! De ankeritems in de Nederlandse toets worden veel beter gemaakt dan diezelfde items als onderdeel van de internationale toets.

Uit het artikel van de Twentse onderzoekers blijkt tussen de regels door ook wel de lichte paniek die deze resultaten teweeg brachten. De conclusie volgens de onderzoekers: 'de meest waarschijnlijke oorzaak voor de gevonden verschillen is de positie van de verschillende opgaven in beide toetsen'. Ik kan daarin een eind meegaan bij gebrek aan alternatieven: het feit dat de nationale toets als geheel veel 'leuker en uitdagender' was en ook herken-

baarder voor de leerlingen heeft ongetwijfeld een rol gespeeld.

Waar ik de onderzoekers in het geheel niet kan volgen, is hun conclusie dat we de resultaten op de ankeritems van de Nationale Optie maar niet verder in de beschouwing moeten meenemen. Terwijl die nu juist de Nederlandse situatie enigszins juist weergeven.

Dit voorval maakt duidelijk dat TIMSS geen goed instrument is om de kwaliteit van het Nederlandse wiskundeonderwijs te bepalen. Tevens vormt het een indicatie van de – grote – afstand tussen het door TIMSS gemeten 'getoetste' curriculum en het Nederlandse 'getoetste' curriculum, waarbij ook de 'cultuur' klaarblijkelijk een grote rol speelt. Een lekenberekening leert dat, als we de nationale optie als maat nemen voor het Nederlandse onderwijs, er een score van rond de 585 uit zou rollen in plaats van de huidige 541. Dat is nogal een verschil, zeker als vanuit statistisch standpunt bekeken beide resultaten representatief zijn.

Valt er toch wat te leren? Misschien wel, als we wat verder kijken dan de rangordellijsten. Ik ga op twee aspecten nader in.

Wat om te beginnen erg opviel aan de ranglijst van landen, is de lage plaats van alle Scandinavische landen. Wat die landen gemeenschappelijk hebben, is een soort middenschoolgedachte en de discussie is weer actueel in hoeverre deze filosofie een verklaring biedt voor de wel zeer magere resultaten. Nu is de middenschool wel aan Nederland voorbijgegaan, maar dankzij de inspanningen van Wallage en Netelenbos hebben we toch nog een wat verwaterde variant binnengehaald die leeft onder de naam 'basisvorming'. De basisvorming is door de basis (scholen, docenten) zeker niet innig omarmd. Recente rapporten laten zich dan ook kritisch uit over de basisvorming, maar de Staatssecretaris staat pal, zoals zo vaak. De vraag is of er een link te leggen valt tussen de middenschool in Scandinavië en de zeer slechte resultaten op TIMSS aan de ene kant en de invoering van de basisvorming en de goede resultaten van Nederland. Waarschijnlijk of vermoedelijk wel.

Wat is het geval? Nederland scoort zeer goed in de brugklas (en in het primair onderwijs), maar al een stuk minder in de tweede klas. Sterker nog: Nederland is een staartland voor wat betreft de vorderingen van leerlingen van klas 1 (brugklas) naar klas 2. Dat betekent een geringe leerwinst. En dat was precies wat door velen voorspeld was bij de start van de basisvorming: een (hemels)breed programma zonder enige diepte. Keihard kan die conclusie niet gemaakt worden. Maar de 'circumstantial evidence' zou de verantwoordelijken toch op zijn minst zenuwachtig moeten maken. Over een jaar of drie zal hier meer over bekend worden, want in 1999 doet Nederland mee aan een 'repeat' van TIMSS.

Het lijkt echter verstandig niet zolang te wachten met het opstarten van de discussie. Duidelijk is in ieder geval dat we met een probleem zitten.

Het tweede probleem is nog meer van inhoudelijke aard: de algebra. Nederland scoort nogal laag bij de algebra (middenmoter). Daar valt in ieder geval iets zinnigs over te zeggen.

Begin jaren negentig was de politiek ervan overtuigd dat wiskunde belangrijk, nee, essentieel is voor iedere leerling – op z'n minst tot hij of zij zestien is. De Commissie die toen de opdracht kreeg nieuwe leerplannen te ontwikkelen voor het gebied 12-16, kreeg dan ook de opdracht een programma 'voor allen' te ontwikkelen. Dat was geen eenvoudige klus – met name de algebra zorgde voor veel hoofdbrekens. Zoveel hoofdbrekens zelfs, dat in het eindrapport voor kenners duidelijk te zien is dat dit gedeelte van het curriculum nog niet 'af' was.

De Commissie beval dan ook aan om hierop verder te studeren. Maar de algebra was inmiddels wel ontdaan van veel van zijn angels (zo zien leerlingen het tenminste vaak): weinig formalismen, weinig formulegemanipuleer. Daartegenover staat een veelbelovend begin van een meer intuïtieve en informele benadering die het redeneren centraal stelt. Helaas is de uitlijning van het algebra-deel curriculum nooit helemaal uit de verf gekomen – een zorgpunt dat onmiddellijke actie lijkt te eisen.

Dezelfde Commissie kreeg parallel als haar taak om eindtermen basisvorming te ontwikkelen – een taak waarbij zij gerechtvaardigde vraagtekens zette, omdat óf de algebra óf de basisvorming geen recht gedaan zou kunnen worden. Uiteindelijk worden de leerlingen altijd de dupe van dit soort politiek-inhoudelijk discussies.

Hoe dan ook: de algebra in de basisvorming is een nogal verwaterd geheel en nogal informeel. Echt een onderdeel dat een grondige heroverweging verdient, net als de basisvorming. Maar een verklaring voor een aspect van de TIMSS-resultaten ligt hierin wel.

Helaas is de situatie inmiddels alweer verder aan het verslechteren. De plannen voor het VBO/MAVO maken de positie van het vak wiskunde er niet beter op. Niet alleen maakt Nederland zich nogal belachelijk, doordat wiskunde straks niet meer voor allen verplicht zal zijn in een tijd waarin iedereen de plaats van wiskunde steeds meer benadrukt, maar ook zijn de voorgestelde curricula van adembenemend slechte kwaliteit, zeker als men in aanmerking neemt dat wiskunde dus niet meer voor ALLEN is.

De verantwoordelijkheid voor deze slechte ontwikkelingen ligt in de eerste plaats bij de politiek – maar dat neemt niet weg dat allen die in Nederland medeverantwoordelijk zijn voor goed wiskundeonderwijs, de krachten moeten bundelen – vooral die mensen, waarbij inhoud voorop staat en die zich niet bij iedere nieuwe wind vanuit Zoetermeer neerleggen. Als dat een resultaat is waaraan TIMSS heeft bijgedragen, is de Nederlandse deelname alweer gelegitimeerd.

Jan de Lange
Freudenthal Instituut

Literatuur

- Kuiper, W., K. Bos en Tj. Plomp (1997). 'De TIMSS nationale optietoets wiskunde'. *Nieuwe Wiskrant* 17(1), pp. 39-46.
- Beaton, Albert E. e.a. (1996). *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study*. Boston College: Chestnut Hill.
- Lange, J. de (1997). 'Looking through the TIMSS mirror from a teaching angle'. In: *Selected Papers from TIMSS-Symposium*, NRC, Washington.
- homepage TIMSS: <http://wwwwcsteep.bc.edu/timss>

Sporen conferentie – Rekenen en Wiskunde in het ROC

Op donderdag 23 en vrijdag 24 april 1998 wordt in de Leeuwenhorst te Noordwijkerhout de conferentie *Sporen 1998* gehouden. Deze conferentie is bedoeld voor docenten rekenen en/of wiskunde in ROC's. Het vak rekenen & wiskunde kent binnen het ROC een uiterst breed en gevarieerd spectrum. Tijdens deze conferentie willen we ingaan op de nieuwste ontwikkelingen. Ook kunt u zelf aan de slag.

Thema's

In werkgroepen, presentaties en lezingen komen de volgende thema's aan de orde. Rekenen&Wiskunde en: implementatie van de KSE, de functie voor het dagelijks leven en/of beroep, didactische vernieuwingen, zelfstandig leren, toetsing, ICT en NT2.

Deze thema's worden uitgewerkt naar verschillende 'werksoorten' (basiseducatie, VAVO, beroepsonderwijs) en niveau's binnen het ROC. Bij elk thema wordt veel aandacht geschonken aan de praktische bruikbaarheid op de werkvloer.

Organisatie

Sporen 1998 wordt georganiseerd door het Cinop en het Freudenthal Instituut.

Inlichtingen en aanmeldingsformulier:
Sporen 1998, p/a Freudenthal Instituut
t.a.v. Betty Heijman, Tiberdreef 4, 3561 GG Utrecht
tel: 030-261 16 11, fax: 030-266 04 30
email: sporen@fi.ruu.nl