

WISKUNDIGE

GELETTERDHEID VOLGENS PISA

HOE STAAT DE VLAG ERBIJ?

1. ANALYSE

Truus Dekker
Kees Lagerwaard
Jan de Lange
Ger Limpens
Monica Wijers

WISKUNDIGE GELETTERDHEID VOLGENS PISA

HOE STAAT DE VLAG ERBIJ?

1. ANALYSE



Freudenthal Instituut
Aldadreef 12
3561 GE Utrecht

Deze publicatie is samengesteld door het PISA-NL-team, gevormd door:
Prof.dr. Jan de Lange, Freudenthal Instituut, projectleider
Drs. Kees Lagerwaard, Citogroep, projectleider

Truus Dekker MA, Freudenthal Instituut
Drs. Ger Limpens, Citogroep
Drs. Monica Wijers, Freudenthal Instituut

Met dank aan Johan Wijnstra voor zijn analyse van de PISA-gegevens

Utrecht/Arnhem, 2006

Layout: Jan de Lange, Kaspar de Lange, Jasper Lisman
Eindredactie: Truus Dekker, Monica Wijers
Redactie: Anneleen Post, Ank van der Heiden
Druk: Wilco, Amersfoort

Inhoudsopgave

0. Voorwoord	5
<i>Jan de Lange, Kees Lagerwaard</i>	
1. Inleiding	7
<i>Kees Lagerwaard, Ger Limpens</i>	
2. Van voor naar achter	21
<i>Jan de Lange</i>	
3. Hoe werden de opgaven in Nederland gemaakt?	47
<i>Truus Dekker, Monica Wijers</i>	
4. PISA-resultaten in het vmbo	57
<i>Monica Wijers, Truus Dekker</i>	
5. Over probleemoplossen en wiskunde	81
<i>Kees Lagerwaard, Ger Limpens</i>	
6. Opmerkingen en Aanbevelingen naar aanleiding van dit Rapport	105
Bijlage	107
Literatuur	109

0. Voorwoord

Jan de Lange, Kees Lagerwaard

In het voorjaar van 2003 vond er een groot internationaal onderzoek plaats naar de staat van onder andere de wiskundige geletterdheid in zo'n veertig landen. Eind 2004 verschenen de eerste internationale en nationale rapporten. Voor Nederland stond de vlag er goed bij: na Finland het tweede Europese land en derde in de OECD/OESO wereld. OESO staat voor Organisatie voor Economische Samenwerking en Ontwikkeling, OECD is het engelse equivalent.

De studie heet PISA: Programme for International Student Assessment. Nederland deed het dus goed in PISA, zoals alle media dan ook trouw meldden. En zo af en toe werden er ook nog details gepubliceerd uit het lijvige PISA-rapport. Maar vaak bleek alleen de plaats en de score het vermelden waard.

Veel mensen vinden dat jammer en enkelen, werkzaam bij het Cito en het Freudenthal Instituut, namen het initiatief om een poging te wagen achter de score te kijken. Immers, alle gegevens van PISA zijn openbaar, en zelfs het leerlingenwerk kan ingekken worden. Het vermoeden was dat een meer inhoudelijke kijk naar de 'resultaten' wellicht meer relevante informatie zou kunnen opleveren. Daarbij werd vooral gedacht aan het 'zorgenkindje' van het Nederlandse wiskundeonderwijs, het vmbo.

Het Ministerie van Onderwijs, Cultuur en Wetenschappen bleek ook geïnteresseerd en zodoende kon het Cito/FI-team aan de slag. Met de voorliggende dubbelpublicatie als resultaat.

Het eerste deel is het resultaat van een analyse van de resultaten in meer inhoudelijke zin.

In overleg met het Ministerie is gekozen voor een zoom-benadering. Allereerst wordt gekeken naar de prestatie van Nederland in relatie met die van een aantal andere West-Europese landen, plus de Verenigde Staten. Daarna zoomen we in op de prestatie van Nederland in de volle breedte om ten slotte naar het vmbo in het bijzonder te kijken.

Aangezien er in 2003 ook een studie was naar het niveau van curriculum overstijgend probleemoplossen, bevat het laatste hoofdstuk een analyse van deze resultaten. Dit omdat de opgaven erg vaak een sterk wiskundige inslag hadden, terwijl de resultaten nogal afweken van die van wiskunde.

Het Cito/FI-team is de uitdaging niet uit de weg gegaan om een aantal concrete aanbevelingen te doen, terwijl er ook in de reguliere tekst al veel, meer op de klaspraktijk afgestemde, aanbevelingen staan.

Het tweede deel levert een schat aan informatie, met name ook voor docenten, gekoppeld aan alle nu vrijgegeven opgaven, uiteraard vergezeld van de relevante gegevens en een toelichting. Gezien het aparte karakter van dit deel is er voor gekozen

dit separaat uit te geven, zodat de totale studie uit twee handzame deeltjes bestaat. De samenstellers realiseren zich dat het geheel waarschijnlijk iets minder spannend overkomt dan bijvoorbeeld de Da Vinci Code, maar voor de in wiskundeonderwijs geïnteresseerde lezer is er veel van waarde te vinden achter de PISA-score, denken wij.

Met deze secundaire analyse hopen we niet alleen een bijdrage te leveren aan een reflectie op en een verbetering van het Nederlandse wiskundeonderwijs. We denken dat deze publicatie een discussie omtrent de gewenste ontwikkelingen binnen het wiskundeonderwijs zal faciliteren, en verwachten dat ook in volgende cycli van PISA een dergelijke analyse gemaakt zal worden. Zo krijgt Nederland meer waar voor zijn PISA-geld.

1. Inleiding

Kees Lagerwaard, Ger Limpens

In 2003 vond de tweede ronde plaats van het driejaarlijkse programma PISA (Programme for International Student Assessment). De OESO besteedt al vanaf het begin veel aandacht aan onderwijs maar had tot voor enkele jaren geen eigen indicatoren om de onderwijsopbrengsten in de diverse lidstaten in kaart te brengen. PISA is gericht op 15- à 16-jarige leerlingen – tegen het einde van de leerplichtige leeftijd dus – en brengt gegevens bijeen over leesvaardigheid, en kennis en vaardigheden in wiskunde en natuurwetenschappen.

In 2000, de eerste ronde, lag de nadruk op leesvaardigheid. In de tweede cyclus lag de nadruk op wiskunde. Behalve leesvaardigheid en natuurwetenschappen was overigens in de 2003-ronde ook het domein *problem solving* opgenomen. Bij dat domein gaat het niet zozeer om een specifiek vak maar meer om het toepassen van algemene probleemoplossende vaardigheden. In 2006, de derde cyclus, zal natuurwetenschappen hoofdonderwerp van PISA zijn.

Aan PISA 2003 namen, behalve de 30 huidige OESO-lidstaten, ook elf niet-lidstaten deel, de zogeheten partnerlanden. Het aantal deelnemende landen aan de derde ronde is opnieuw aanzienlijk toegenomen: in 2006 zullen circa 60 landen meedoen. Voor veel landen is deelname aan PISA een bewijs dat zij de kwaliteit van het onderwijs van groot belang achten. Bovendien krijgen zij een goed beeld van 'waar ze staan' ten opzichte van de rijke, geïndustrialiseerde landen.

De PISA Governing Board, waarin alle deelnemende OESO-landen vertegenwoordigd zijn, bepaalt de grote lijnen van het onderzoek. Op internationaal niveau wordt het project uitgevoerd door een consortium onder leiding van ACER (Australian Council for Educational Research). In dit consortium participeert ook het Cito. Verder is er in alle deelnemende landen een projectorganisatie gevormd die de gegevens van het PISA-onderzoek verzamelt. In Nederland is deze taak door het Ministerie van OCW ondergebracht bij het Cito. Het Cito is dus op twee manieren betrokken bij PISA.

PISA verschilt van andere internationale vergelijkende onderzoeken. Bij PISA wordt niet zozeer gekeken naar de gemeenschappelijke zaken in de verschillende onderwijscurricula, maar ook en vooral naar de mate waarin leerlingen opgedane kennis en vaardigheden kunnen gebruiken in het dagelijks leven. In de verschillende PISA-rapporten die sinds 2000 gepubliceerd zijn wordt dan ook vaak de toevoeging *literacy* gehanteerd bij de diverse domeinen waarover gerapporteerd wordt: *reading literacy*, *mathematical literacy* en *scientific literacy*. In het Nederlandse rapport wordt ten aanzien van wiskunde meestal de term 'wiskundige geletterdheid' gebruikt. Het begrip wiskundige geletterdheid wordt in het Nederlandse rapport als volgt omschreven: 'Wiskundige geletterdheid is de vaardigheid om – met gebruikmaking van wiskundige kennis – vraagstukken in een realistische context te benaderen en op te lossen'. Dit rapport *Praktische kennis en vaardigheden van vijftienjarigen* is in december 2004 verschenen en is via de website van het Cito (<http://www.cito.nl/> of <http://www.pisa.nl/> daarna doorklikken via 'Expertisecentrum') te downloaden. In het PISA-onderzoek is er niet alleen aandacht voor cognitieve vaardigheden maar ook voor andere factoren die van invloed kunnen zijn op toekomstige prestaties. Leerlingen krijgen niet alleen een boekje met opgaven voorgelegd. Er is een uitgebreide leerlingvragenlijst waarin

vragen zijn opgenomen over leerstrategieën en over de attitude van de leerling ten opzichte van het vak wiskunde. Bovenvermeld rapport doet verslag van deze aspecten. De voorliggende publicatie gaat nader in op veel zaken die ook in het rapport *Praktische kennis en vaardigheden van 15-jarigen* opgenomen zijn.

Wiskundige domeinen, competenties en contexten

De vraagstukken die in het PISA-onderzoek aan de 15-jarige leerlingen worden voorgelegd, zijn speciaal voor dit doel ontworpen. Bij de constructie ervan stonden drie aspecten centraal: de wiskundige inhoud, de voor beantwoording noodzakelijke competenties en de context waarbinnen het vraagstuk gesitueerd wordt.

In PISA 2003 zijn vier wiskundige domeinen opgenomen: Vorm en Ruimte (ruimtelijke en geometrische vormen), Veranderingen en Relaties (vergelijkingen, formules, grafieken, tabellen), Onzekerheid (kansrekening en statistiek) en Hoeveelheid (numerieke verschijnselen, kwantitatieve relaties). PISA 2000 beperkte zich tot de eerste twee domeinen, Vorm en Ruimte en Veranderingen en Relaties. Omdat wiskunde hoofdonderwerp was bij het PISA 2003-onderzoek, kon het wiskundige gebied uitgebreid worden.

In het kader van het PISA-onderzoek worden er drie competentieclusters onderscheiden, te weten het Reproductiecluster (routineprocedures, feitenkennis en dergelijke), het Verbindingencluster (het leggen van verbanden tussen diverse representatievormen, interpretatievaardigheden) en het Reflectiecluster (generaliseren, abstraheren, redeneren en bewijzen).

De contexten van de diverse opgaven variëren van contexten in de persoonlijke sfeer via de schoolse en beroepsmatige sfeer tot contexten binnen de publieke sfeer en contexten binnen de wetenschappelijke sfeer.

Voor meer informatie over deze aspecten zie PISA 2003 *Assessment Framework: Mathematics, Reading, Science and Problem Solving Knowledge and Skills* (OECD, 2003, ISBN 9264101721). Deze publicatie is ook te downloaden via <http://www.pisa.oecd.org>.

Verschillende typen PISA-vragen en de beoordeling daarvan

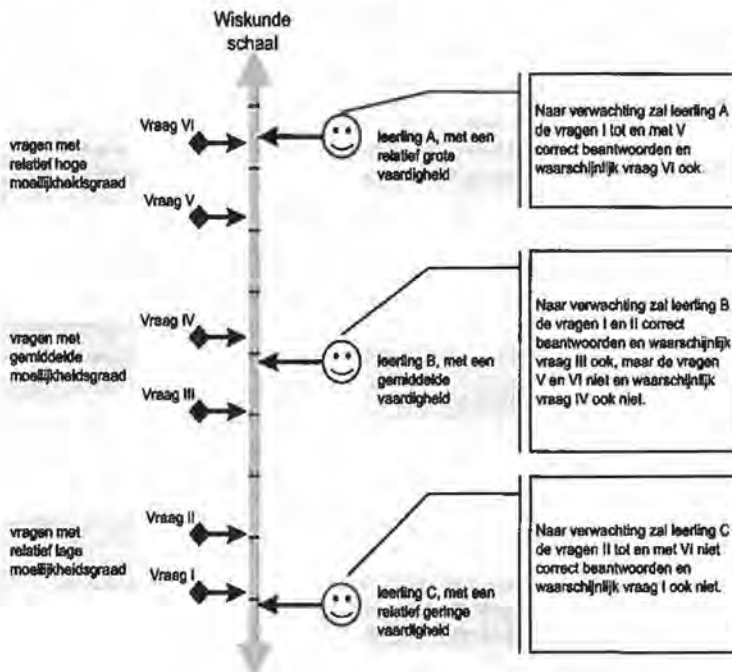
Bij PISA zijn drie vraagtypen te onderscheiden: lang antwoordvragen (antwoord plus toelichting), kort antwoordvragen (antwoord in de vorm van een getal of een andere strikt omschreven vorm) en meerkeuzevragen. Bij de beoordeling van lang antwoordvragen is gebruikgemaakt van speciaal voor dit doel opgeleide correctoren, die aan de hand van een per vraag geformuleerd correctievoorschrift ieder leerlingenantwoord voorzagen van een bijbehorende categoriecode. In het kader van de beoordeling van met name de lang antwoordvragen hebben in 2002, voorafgaand aan het PISA-onderzoek in 2003, proefafnames plaatsgevonden. Op basis van daarbij verkregen leerlingenantwoorden konden de correctievoorschriften nog meer worden verfijnd.

Ontwerp, analyse en schaling van de toetsen

Het PISA 2003-onderzoek omvatte 84 wiskundevragen. Niet iedere leerling kreeg daarbij iedere vraag voorgelegd: er waren verschillende clusters gemaakt die elk een half uur aan leerlingentwerk bevatten en iedere leerling ontving een viertal verschil-

lende clusters. De samenstelling van een set van vier clusters bestemd voor een leerling wisselde zodanig dat iedere vraag even vaak voorkwam in de testboekjes. Daarbij is ervoor gezorgd dat ieder cluster vragen even vaak vóór, achter of midden in een testboekje verscheen.

De relatieve vaardigheid van een deelnemende leerling kan geschat worden door de proportie van vragen te beschouwen die juist beantwoord wordt. De relatieve moeilijkheid van een vraag kan geschat worden door de proportie van testdeelnemers te beschouwen die de betreffende vraag juist beantwoordt. Op basis van deze schattingen kan een continue schaal gedefinieerd worden die de te onderzoeken wiskunde representeert. Op die schaal kunnen zowel individuele leerlingen als individuele vragen worden geplaatst (zie fig. 1). Bij iedere leerling en bij iedere vraag is dan de mate van wiskundige geletterdheid af te lezen. Alle schalen bij PISA hebben dezelfde kenmerken: ze zijn gestandaardiseerd op een gemiddelde van 500 voor alle OESO-landen bij elkaar met een standaardafwijking van 100. De schalen zijn gebaseerd op 'item response theorie' (IRT). Deze theorie maakt het mogelijk om 'leerlingen en opgaven op één schaal te plaatsen'. Dit betekent zoveel als: bij een vaardigheidsniveau op de schaal zijn opgaven te vinden die illustreren wat kinderen op dat niveau goed aankunnen, wat in principe makkelijke opgaven voor hen zijn en welke waarschijnlijk nog te moeilijk zijn. Door leerlingresultaten bij elkaar te nemen is het gemiddelde per land te berekenen en ook op de schaal aan te geven.



figuur 1 Wiskundeschaal

De moeilijkheid van een vraag wordt uitgedrukt in een getal dat we de drempelwaarde van die vraag noemen. De vraag met precies de gemiddelde moeilijkheid heeft dan drempelwaarde 500. Dat is immers ook de gemiddelde score van de deelnemende leerlingen uit de OESO-landen. Hoe moeilijker een vraag, hoe hoger de drempelwaarde. De moeilijkste van de 84 wiskundevragen had een drempelwaarde van 801. De eenvoudigste daarentegen scoorde als drempelwaarde 262.

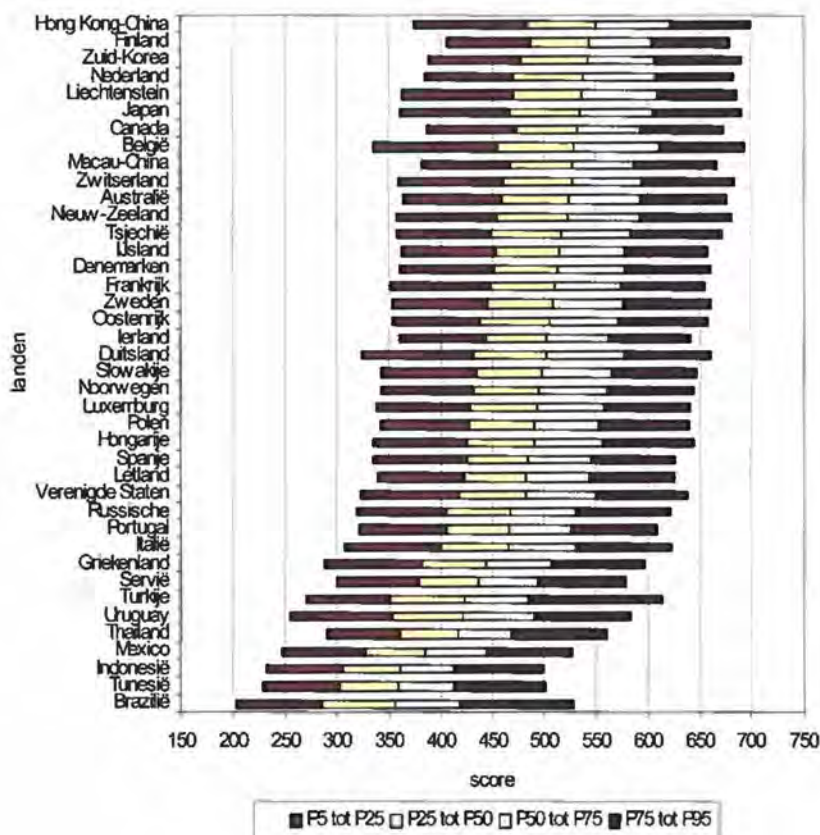
In figuur 1 is te zien dat leerling C net onder de drempelwaarde van vraag 1 zit. Dit betekent niet dat vraag 1 voor leerling C echt veel te moeilijk is, maar meer dan de helft van de leerlingen van het niveau van leerling C zal deze vraag niet goed beantwoorden. Leerling C is een zwakke leerling: de overige vijf vragen zijn voor deze leerling veel te moeilijk.

Nederlandse resultaten voor wiskunde internationaal vergeleken

In figuur 2 zien we voor ieder land dat participeerde (OESO-landen zowel als partnerlanden) de scores op de wiskundeschaal van de middelste 90% van de leerlingen weergegeven. De landen zijn gerangschikt volgens de score op de middelste 50% van de leerlingen, de P50 (de mediaan dus). Iedere balk is verdeeld in vier vakken waarvan het meest linkse vak de afstand aangeeft tussen percentiel 5 en 25 (P5 tot P25), vervolgens een vak met de afstand tussen percentiel 25 en 50 (P25 tot P50), dan een vak met de afstand tussen percentiel 50 en 75 (P50 tot P75) en ten slotte een vak met de afstand tussen percentiel 75 en 95 (P75 tot P95). Zowel de laagst scorende 5% van de leerlingen als de hoogst scorende 5% van leerlingen zijn hier dus niet weergegeven. Dit is gedaan om te voorkomen dat uitschieters teveel nadruk leggen op het geheel aan scores.

Uit figuur 2 is voor Nederland af te lezen dat 90% van de geteste leerlingenpopulatie een score had tussen 385 en 683 en een gemiddelde van 538. Nederland scoort daarmee hoog. Slechts drie landen hebben een hogere gemiddelde score voor wiskunde: Hong Kong, Finland en Zuid-Korea.

De ons omringende landen scoren lager. Duitsland, Frankrijk en Denemarken staan op respectievelijk plaats 20, 16 en 15 en blijven met hun score aanzienlijk achter bij Nederland. België staat op plaats 8 en heeft met 529 een lagere gemiddelde score dan Nederland. Daarbij moet worden aangetekend dat de verschillen tussen de Nederlandstalige en Franstalige leerlingen in België aanzienlijk zijn. Verderop in deze publicatie zullen we zien dat de Vlaamse leerlingen het zelfs beter doen dan alle andere leerlingen, Vlaanderen zou bovenaan de lijst staan.



figuur 2 Scoreverdeling op de wiskundeschaal in de OECD- en partnerlanden

Verder is het interessant op te merken dat in PISA 2000 de Oost-Aziatische landen ook al hoog scoorden. Net als drie jaar eerder wordt Nederland in PISA 2003 op de ranglijst omringd door landen als Hong Kong, Zuid-Korea en Japan.

Vaardigheidsniveaus

Om meer zicht te krijgen op de bij PISA 2003 ontwikkelde schaal is een onderverdeling van deze schaal in verschillende vaardigheidsniveaus ontworpen. Bij PISA 2000 waren de leerlingenscores op leesvaardigheid ondergebracht in vijf vaardigheidsniveaus. Op dezelfde wijze zijn in PISA 2003 voor wiskunde zes verschillende niveaus ontworpen, die een serie vaardigheden bevatten met een stijgende moeilijkheidsgraad. Hierbij is niveau 1 het eenvoudigste en niveau 6 het moeilijkste niveau. Leerlingen met een vaardigheid op de PISA-schaal lager dan 358 zijn geclassificeerd als 'onder niveau 1'. Een uitgebreide beschrijving van de vaardigheidsniveaus is te vinden in de volgende tabel.

Tabel 1

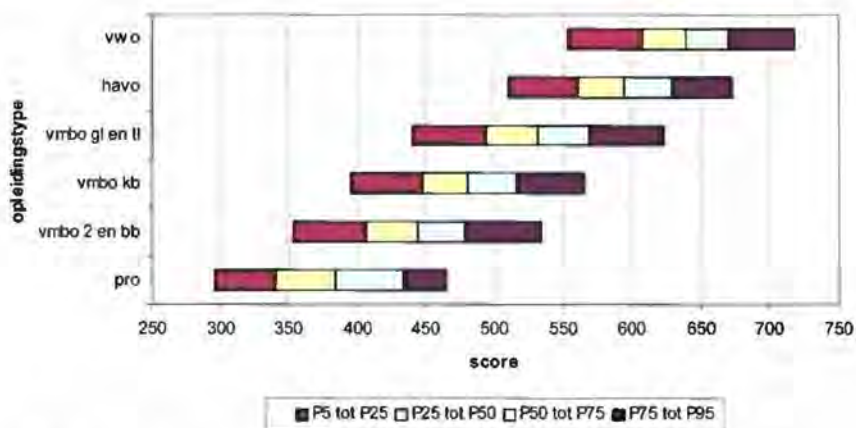
niveau	Vaardigheden	schaal
1	Op niveau 1 kan een leerling: • vragen beantwoorden die betrekking hebben op bekende contexten indien alle relevante informatie gegeven is en de vraagstelling helder is omschreven; • informatie identificeren en routineprocedures uitvoeren die betrekking hebben op directe aanwijzingen in expliciete situaties; • activiteiten uitvoeren die voor de hand liggend zijn en onmiddellijk uit de gegeven stimuli volgen	van 358 tot 420
2	Op niveau 2 kan een leerling: • situaties in contexten interpreteren en herkennen op basis van directe gevolgtrekkingen; • relevante informatie onttrekken aan een enkele bron; • gebruikmaken van een enkele representatievorm; • gebruikmaken van elementaire algoritmes, formules, procedures of afspraken; • gebruikmaken van eenvoudig redeneren; • letterlijke interpretaties maken van resultaten	van 420 tot 482
3	Op niveau 3 kan een leerling: • helder omschreven procedures uitvoeren waaronder procedures op basis van gefaseerde besluitvorming; • selecteren en eenvoudige vraagstukoplossende strategieën toepassen; • interpreteren en gebruik maken van representatievormen gebaseerd op verschillende informatiebronnen; • korte mededelingen doen waarin verslag gedaan wordt van gevonden interpretaties, resultaten en redeneringen	van 482 tot 544
4	Op niveau 4 kan een leerling: • gericht werken met expliciete modellen van ingewikkelde situaties waarbij beperkingen aan de orde kunnen zijn of zelf veronderstellingen gemaakt dienen te worden; • kiezen uit dan wel integreren van verschillende representatievormen, waaronder symbolische vormen, waarbij deze op een directe manier in verband gebracht kunnen worden met realistische situaties; • uitleg en argumenten construeren en communiceren, gebaseerd op eigen interpretatie en redeneringen	van 544 tot 606

Tabel 1

5	Op niveau 5 kan een leerling: • modellen voor ingewikkelde situaties ontwikkelen en daarmee werken waarbij randvoorwaarden geïdentificeerd worden en zelf veronderstellingen gespecificeerd worden; • geschikte probleemoplossende strategieën selecteren, vergelijken en evalueren om complexe vraagstukken die bij vermelde modellen horen op te lossen; • strategisch werken, daarbij gebruikmakend van brede, goed ontwikkelde redeneervaardigheden, geschikte representatievormen, symbolische en formele karakteristieken en inzicht relevant voor de vermelde ingewikkelde situaties; • reflecteren op zijn eigen handelen; • zijn interpretaties en redeneringen formuleren en communiceren	van 606 tot 668
6	Op niveau 6 kan een leerling: • conceptualiseren, generaliseren en informatie benutten, gebaseerd op het onderzoek en het modelleren van een complexe probleemstelling; • diverse informatiebronnen en representatievormen met elkaar verbinden en flexibel overstappen van de een op de ander; • op hoog wiskundig niveau denken en redeneren; • dit inzicht en begrip samen met symbolische en formele wiskundige operaties en verbanden inzetten om nieuwe aanpakken of strategieën te ontwikkelen om ongebruikelijke situaties aan te pakken; • zijn bevindingen, interpretaties en argumenten rond zijn handelingen en overdenkingen en tevens de geschiktheid hiervan met betrekking tot de oorspronkelijke situatie formuleren en helder communiceren	668 en hoger

In figuur 3 worden de diverse landen weergegeven op basis van de percentages leerlingen binnen elk van deze zes vaardigheidsniveaus. Voor de volledigheid is ook bij ieder land weergegeven hoe groot het percentage leerlingen is dat onder niveau 1 blijft.

Nederland doet het, bij een weergave zoals in figuur 3, goed. Van alle Europese landen staat Nederland weer op de tweede plaats, voorafgegaan door Finland. Onze buurlanden Duitsland en België zijn een stuk lager in de internationale rangorde te vinden.



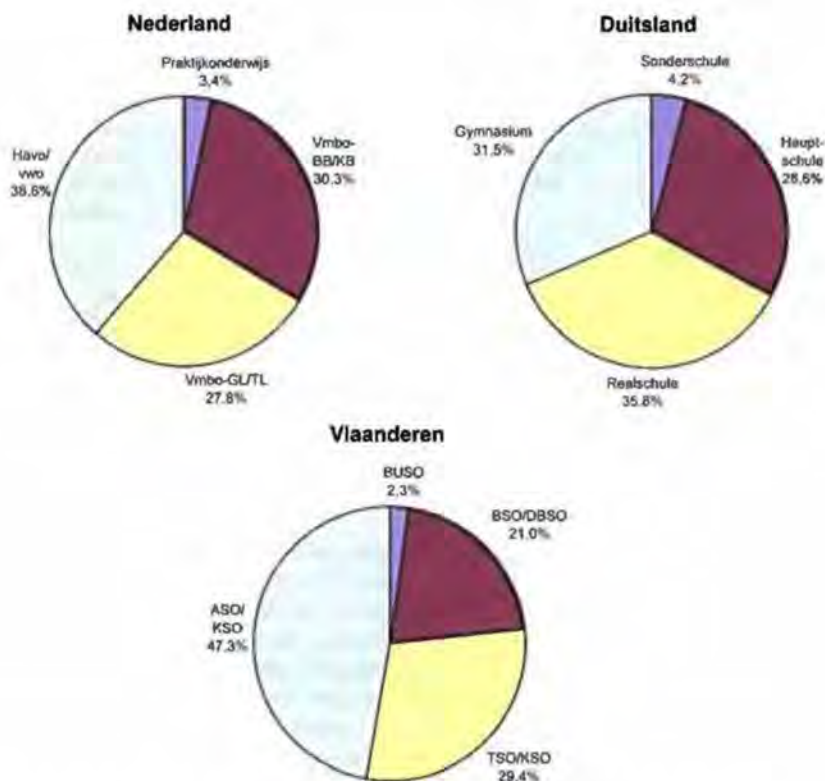
figuur 4 Scoreverdeling op de wiskundeschaal in de Nederlandse opleidingstypen

Nederland vergeleken met Vlaanderen en Duitsland

Een vergelijking met Vlaanderen en Duitsland ligt voor de hand. Verderop in deze publicatie zullen de prestaties van de Nederlandse leerlingen worden vergeleken met hun leeftijdsgenoten net over de grens. Door de verschillen in schoolsystemen in Vlaanderen, Duitsland en Nederland is een vergelijking per schooltype onmogelijk. Maar toch kunnen we de leerlingen in de drie landen zodanig in groepen verdelen dat enige vergelijking wel mogelijk is.

Als we in Nederland drie vrij grote groepen onderscheiden, havo/vwo, vmbo GL/TL en vmbo KB/BB, blijft er nog de kleine groep van het praktijkonderwijs over.

In Duitsland kunnen we dan een soortgelijke indeling maken met de schooltypen Gymnasium, Realschule, Hauptschule en Sonderschule. Voor Vlaanderen lukt een vergelijkbare indeling alleen als we het kleine Kunstsecundair onderwijs (KSO) verdelen over het Algemeen Secundair Onderwijs (ASO) en het Technisch Secundair Onderwijs (TSO). Vergelijkbaar met ons KB/BB hebben we dan in Vlaanderen Beroeps Secundair Onderwijs (BSO) en Deeltijd Beroeps Secundair Onderwijs (DBSO). Het kleine BUSO (Buitengewoon Secundair Onderwijs) completeert het geheel. Zie de cirkeldiagrammen in figuur 5.



figuur 5 Verdeling van de PISA-steekproeven 2003 in Nederland, Duitsland en Vlaanderen over programmatypen

In de volgende hoofdstukken van deze publicatie zullen we diverse keren gebruikmaken van de Nederlandse indeling in drie typen onderwijs (meestal laten we het Praktijkonderwijs buiten beschouwing) om prestaties van groepen Nederlandse leerlingen te vergelijken met hun Vlaamse en Duitse leeftijdgenoten.

Als voorproefje kijken we even naar de groepen bestpresterende leerlingen die algemeen voortgezet onderwijs volgen. De gemiddelde score voor wiskunde is voor havo-leerlingen 594 en voor vwo-leerlingen 639. Het gemiddelde van deze twee groepen samen is 617. Het betreft hier bijna 39% van alle 15-jarigen. In Vlaanderen volgt zo'n 46% van de leerlingen het Algemeen Secundair Onderwijs. Hun gemiddelde score is 624. Het betreft hier dus een groter deel van de leerlingenpopulatie, terwijl de gemiddelde score van deze groep duidelijk hoger is. Dat geeft de hoge positie van Nederland op de internationale ranglijst toch wat minder glans.

De scores van jongens en meisjes

In PISA 2000 bleken de Nederlandse meisjes gemiddeld 11 punten lager te scoren dan de jongens. Ook in de ons omringende landen bleven de scores van de meisjes iets achter bij die van de jongens.

In PISA 2003 is het verschil tussen Nederlandse jongens en meisjes slechts 5 punten

op de PISA 2003-schaal. Jongens scoren gemiddeld 540 en meisjes 535. Het OESO-gemiddelde is voor de jongens 506 en voor meisjes 494 (zie ook tabel 3).

Tabel 3

	Gemiddelde	Gemiddelde jongens	Gemiddelde meisjes	Vershil jongens - meisjes
België	529	533	525	8
Duitsland	503	508	499	9
Nederland	538	540	535	5
Hong Kong	550	552	548	4
Korea	542	552	528	24
OESO	500	506	494	12

We kunnen constateren dat het verschil tussen Nederlandse jongens en meisjes ten opzichte van de meerderheid van de andere OESO-landen klein is. Ook gemeten ten opzichte van de ons omringende landen komen de niveaus van jongens en meisjes in Nederland redelijk dicht bij elkaar. Overigens is het wel intrigerend om hier te vermelden dat het enige OESO-land waar het verschil tussen jongens en meisjes in het voordeel van de meisjes uitvalt IJsland blijkt te zijn, waar jongens gemiddeld 508 voor wiskunde scoren en meisjes maar liefst 523, hetgeen een verschil van 15 punten oplevert. In Finland, de OESO-koploper, scoren de jongens 548 en de meisjes 541. Ook hier is het verschil tussen de sexen dus vrij marginaal.

Bij het hoogst scorende Hong Kong is het verschil slechts 4 punten. In Korea, dat ook een hoge gemiddelde score heeft, is het scoreverschil tussen jongens en meisjes veel groter.

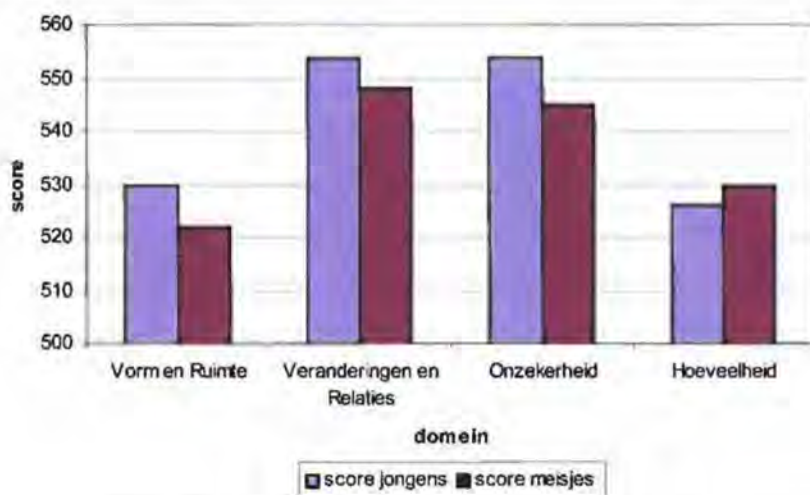
Vershil in scores tussen jongens en meisjes bij de wiskundedomeinen in Nederland

In het domein Vorm en Ruimte scoren Nederlandse jongens gemiddeld 530 en meisjes 522; in Veranderingen en Relaties scoren jongens gemiddeld 554 en meisjes 548; in Onzekerheid scoren jongens gemiddeld 554 en meisjes 545; in Hoeveelheid scoren jongens gemiddeld 526 en meisjes 530. Dus in drie domeinen gebeurt wat verwacht kan worden op basis van de totale wiskundescores: jongens scoren gemiddeld iets beter dan meisjes. In het domein Hoeveelheid is de verhouding andersom (zie fig. 6).

Dat is overigens precies ook het domein waar de verhouding tussen jongens en meisjes in de zes vaardigheidsniveaus wat afwijkend is. Bij de domeinen Vorm en Ruimte, Veranderingen en Relaties en Onzekerheid zien we dat in de lagere niveaus iets meer meisjes zijn dan jongens.

Bij de hoogste niveaus is dat omgekeerd. De verschillen zijn klein, maar consistent. Bij Hoeveelheid bevinden zich op het laagste niveau ook wat meer meisjes dan jongens, maar in de niveaus 1, 2 en 3 meer jongens dan meisjes.

Meisjes zijn in de niveaus 4 en 5 net in de meerderheid, terwijl het op niveau 6 vrijwel 50/50 is.



figuur 6 Gemiddelde score van jongens en meisjes per domein

Een vergelijking met de uitkomsten van PISA 2000

Nederland scoorde in 2000 met wiskunde heel hoog, in feite het hoogst van alle deelnemende landen. Maar omdat het aantal deelnemende scholen en leerlingen in 2000 niet volledig aan de strenge OESO-eisen voldeed, werd deze score niet in alle officiële ranglijsten opgenomen.

Een van de doelen van het PISA-onderzoek is om ontwikkelingen in de tijd te kunnen volgen. Daarom worden in elk volgend onderzoek vragen uit eerdere PISA-onderzoeken meegenomen, waardoor de vaardigheid van leerlingen op de verschillende tijdstippen kan worden vergeleken.

Het feit dat Nederland in 2000 voor wiskunde de hoogste gemiddelde score had en in 2003 met een vierde positie genoeg moet nemen, kan als een aanwijzing worden gezien dat de vaardigheid in wiskunde in Nederland achteruit is gegaan. De mogelijkheid dat andere landen beter zijn geworden, moet op voorhand natuurlijk niet worden uitgesloten. In hoofdstuk 2 van deze publicatie wordt echter betoogd dat deze mogelijkheid niet voor de hand ligt. Een ander aspect dat een rol kan spelen is het feit dat er in 2000 voor wiskunde slechts 1 uur toetstijd beschikbaar was tegen $3\frac{1}{2}$ uur in 2003. In 2000 zijn slechts twee domeinen aan bod gekomen, Vorm en Ruimte en Veranderingen en Relaties. Misschien liggen de twee domeinen die in 2003 zijn toegevoegd de Nederlandse leerling wel wat minder, zo zou men kunnen denken. Maar ook dat wordt min of meer gelogenstraft door de gegevens in tabel 2.

Er zijn analyses uitgevoerd om de prestaties in 2000 en in 2003 met elkaar te kunnen vergelijken. Daartoe zijn de scores van 2000 omgerekend naar de schaal van 2003, dus die met een gemiddelde voor wiskunde totaal van 500 met een standaardafwijking van 100.

Dan blijkt dat Nederland op het domein Veranderingen en Relaties zowel in 2000 als

in 2003 de hoogste gemiddelde score heeft van alle 28 deelnemende OESO-landen. Dat klinkt goed. Maar in 2000 was die gemiddelde score 568 en in 2003 is dat gezakt naar 551. Volgens de analyses is dit een significante afname. Van de overige 27 landen was er in 16 landen geen significant verschil en scoorden 11 landen in 2003 significant hoger dan in 2000.

Op het domein Vorm en Ruimte was de gemiddelde score in 2000 in Nederland 537. Nederland nam daarmee in 2000 een vierde plaats in onder de 28 OESO-landen. In 2003 blijkt die gemiddelde score te zijn teruggelopen tot 526. Daarmee zakten we naar een zevende positie. Het verschil is niet significant. Maar er zijn wel meer landen waarvan de score omhoog is gegaan dan landen waar de score daalde.

Samengevat kan worden gezegd dat de vaardigheid in wiskunde van 15-jarigen in Nederland tot de hoogste in de wereld behoort, maar dat er aanwijzingen zijn dat die vaardigheid in 2003 minder is dan in 2000. Een meer uitgewerkt standpunt kan teruggevonden worden in hoofdstuk 2 van deze publicatie.

Probleemoplossen en wiskunde

In het PISA-onderzoek van 2003 waren ook vragen in het domein Probleemoplossen opgenomen. PISA definieert dit domein als 'de vaardigheid om cognitieve processen te gebruiken bij het aanpakken en oplossen van levensechte, curriculumoverschrijdende problemen waarbij de oplossing niet voor de hand ligt en waarbij de te gebruiken vaardigheden niet behoren tot één enkel domein van wiskunde, natuurwetenschappen of lezen'.

Ook bij wiskunde spreken we vaak over probleemoplossen. In zijn meest uitgebreide vorm gaat het dan over het vertalen van een probleem in een wiskundig model, binnen dat model een oplossing vinden en die oplossing weer terugvertalen naar het oorspronkelijke probleem. In de PISA-opgaven voor wiskunde treffen we vaak probleemoplosactiviteiten aan. Omdat het onderwerp probleemoplossen een eenmalig domein was bij PISA 2003, zijn hiervan alle opgaven vrijgegeven. Aan de hand van deze opgaven kunnen we laten zien dat er zeker een verwantschap is met opgaven uit de wiskundedomeinen. Maar uit de analyses blijkt dat de Nederlandse leerlingen het bij probleemoplossen aanzienlijk minder goed doen dan bij wiskunde. In hoofdstuk 5 wordt daar uitgebreid op ingegaan.

2. Van voor naar achter

Jan de Lange

Van hoog tot laag

De horse race

Hoe je ook over PISA denkt, één ding is overduidelijk: in de ogen van de pers is en blijft het een horse-race. Het PISA-rapport van de OESO zelf is geen uitzondering: al in de eerste alinea's is te lezen dat Finland de winnaar is. En dat was zeker verrassend, omdat het inmiddels een wijdverbreid gegeven was dat in internationaal vergelijkend onderzoek altijd de *East Asian Tigers* (Hong Kong, Singapore, Zuid-Korea, Taiwan) de dienst uitmaken. Sterker nog, het feit dat Nederland in het verleden bij de TIMSS studie direct na de vier of vijf *Tigers* kwam, gaf aan dat Nederland eigenlijk wel heel goed was: *the best of the rest*. Ook internationaal werd dit vaak zo gezien, want die *Tigers*, dat was immers een geheel andere cultuur.

Maar PISA 2003 zorgde voor grote onrust onder de tijgers. Niet alleen stond Finland eerste, maar zeker zo erg: Japan was vierde, na Nederland! ('Gastlanden' niet meegerekend, dus uitsluitend de OESO-leden). Nederland vóór Japan, dat was nog nooit vertoond. De tevredenheid straalde eraf bij de gezagshouders in het kikkerlandje. En de achterstand was ook nog eens heel beperkt (zie tabel 1 van alle OESO-deelnemers).

Finland 544 punten, Korea 542 en Nederland 538. En Japan volgde ons op 4 punten. De vraag ligt voor de hand wat dit nu eigenlijk betekent. Zijn we erg goed? Goed? Redelijk?

Veel antwoorden zijn mogelijk, maar vooraf dient opgemerkt te worden dat alle scores relatief zijn. De gemiddelde totaalscore van alle OESO landen (OECD average) is per definitie vastgesteld op 500 punten en alle landen worden even zwaar gewogen. Naar de klassenpraktijk vertaald: als in een klas alle leerlingen een onvoldoende zouden scoren, zou de OESO het gemiddelde cijfer van die klas (bijvoorbeeld 4,3) aangeven als 500. Daarbij stelt de OESO de standaarddeviatie gelijk aan 100.

Tabel 1

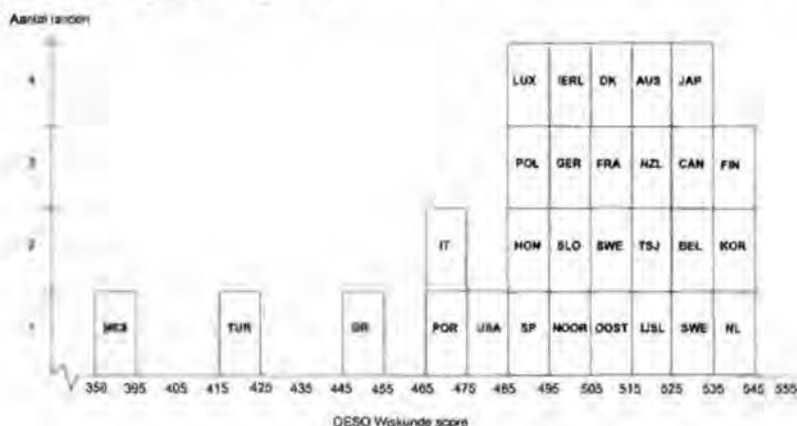
Wiskunde algemeen			
Land	Mean	S.E.	Diff.
Finland	544	(1,9)	O
Korea	542	(3,2)	O
Nederland	538	(3,1)	
Japan	534	(4,0)	O
Canada	532	(1,8)	O
België	529	(2,3)	O
Zwitserland	527	(3,4)	O
Australië	524	(2,1)	<
Nieuw-Zeeland	523	(2,3)	<
Tsjechië	516	(3,5)	<
IJsland	515	(1,4)	<
Denemarken	514	(2,7)	<
Frankrijk	511	(2,5)	<
Zweden	509	(2,6)	<
Oostenrijk	506	(3,3)	<
Duitsland	503	(3,3)	<
Ierland	503	(2,4)	<
Slowakije	498	(3,3)	<
Noorwegen	495	(2,4)	<
Luxemburg	493	(1,0)	<
Polen	490	(2,5)	<
Hongarije	490	(2,8)	<
Spanje	485	(2,4)	<
Verenigde Staten	483	(2,9)	<
Portugal	466	(3,4)	<
Italië	466	(3,1)	<
Griekenland	445	(3,9)	<
Turkije	423	(6,7)	<
Mexico	385	(3,6)	<
OESO gemiddelde	500	(0,6)	<

> = Significant hoger dan Nederland

O = Niet significant verschillend

< = Significant lager dan Nederland

Dat betekent dat de landen aan de bovenkant van de curve (dus scores hoger dan 500) allemaal nogal dicht bij elkaar zitten: minder dan een halve standaarddeviatie van het gemiddelde. Dat de verdeling niet echt 'normaal' is (geheel volgens verwachting), blijkt uit de grafiek in figuur 1.



figuur 1 Verdeling landen over scores

Dus, we moeten ons bij het kijken naar de resultaten van de horse-race vooral realiseren dat het resultaat relatief is. Men kan zeker verdedigen dat Nederland weer 'hoog' scoort, zoals de media ons regelmatig willen doen geloven, maar dat impliceert dus niet dat Nederland ook erg 'goed' scoort. Op dit punt zullen we in deze publicatie regelmatig terugkomen. Als onderdeel van een poging meer inzicht te krijgen in het 'absolute' niveau, zal onder andere worden uitgegaan van de 'echte' scores op de diverse items, de zogenaamde *p*-scores.

Wat vooral opvalt, en telkens weer in dergelijke studies, is dat er een 'verband' lijkt te zijn tussen de scores voor wiskundige geletterdheid en de sociaal/culturele achtergrond van een land. Voor het eerst werd bij PISA 2003 de *East Asian Tiger* groep enigszins uit elkaar gespeeld, maar Korea en Japan zijn 2 en 4. Finland, Nederland, België en Zwitserland zijn nummer 1, 3, 6 en 7. Duitsland en Oostenrijk zijn 15 en 16, en Spanje, Portugal, Italië, Griekenland en Turkije liggen op positie 23, 25, 26, 27 en 28. Dit is wellicht voor de hand liggend, maar zelden als zodanig herkend en erkend in de talrijke discussies.

De subschaal horse race

Zoals bekend, zijn er vier subschalen binnen de wiskundige geletterdheid: die voor Vorm en Ruimte, voor Veranderingen en Relaties, voor Hoeveelheid en die voor Onzekerheid. De scores binnen die schalen leveren zeker stof voor discussie. Zo hadden veel deskundigen een mooie score voor Nederland bij het onderdeel Vorm en Ruimte (vergelijkbaar met meetkunde) verwacht. Mooi niet: Nederland blijft bij dit onderdeel juist opmerkelijk achter. Japan (553) en Korea (548) scoren hoog, op tamelijk ruime afstand van Zwitserland (540) en Finland (539). Daar steekt Nederland met 526 punten wat schril bij af. Hier ligt zeker een punt dat nader onderzoek rechtvaardigt. Ook bij Hoeveelheid (zeg maar rekenvaardigheid en dergelijke) gaat het niet ge-

weldig met Nederland. Maar dat punt is voor veel deskundigen minder verrassend: het rekenonderwijs is niet meer wat het geweest is. Dat de wereld niet zo simpel in elkaar steekt weet iedereen ook wel – en dat zullen we zo ook zien – maar toch, dat rekenen ...

Finland is hier een klasse apart en heeft haar toppositie vooral aan dit onderdeel te danken: met 549 punten ligt het 12 punten voor op Korea en 16 op Zwitserland. En Nederland ligt weer 5 punten achter Zwitserland, op een gedeelde 5e plaats (528 punten).

Maar nu komt het verrassende positieve nieuws. Zowel bij Veranderingen en Relaties als bij Onzekerheid staat Nederland bovenaan! 551 voor algebra en 549 voor Onzekerheid zijn (relatief) hoge scores, en de score voor algebra zal heel wat wenkbrauwen doen fronsen. Maar zoals later zal blijken, dient men alvorens deze wenkbrauwen te fronsen eerst kennis te nemen van wat Veranderingen en Relaties precies inhoudt binnen PISA. Dat is zeker niet hetzelfde als wat veel mensen buiten PISA gemakshalve 'algebra' noemen. Veranderingen en Relaties gaat vooral over verbanden tussen grootheden, en dan vaak nog tamelijk kwalitatief en weinig op formeel niveau. Als voorbeeld kan het item omtrent de hartslag dienen (zie fig. 2): een allerweggen als 'mooi' item bestempeld vraagstuk, dat niet alleen authentiek is maar ook mooie relevante vragen bevat en heel dicht in de buurt komt van formelere algebra.

HARTSLAG

Vanwege gezondheidsredenen moeten mensen bij inspanning, bijvoorbeeld tijdens het sporten, zorgen dat hun hartslag niet boven een bepaald maximum aantal slagen per minuut uitkomt.

Jarenlang werd het verband tussen de aanbevolen maximale hartslag en de leeftijd van een persoon beschreven met de volgende formule:

$$\text{Aanbevolen maximale hartslag} = 220 - \text{leeftijd}.$$

Recent onderzoek laat zien dat deze formule bijgesteld moet worden. De nieuwe formule is de volgende:

$$\text{Aanbevolen maximale hartslag} = 208 - (0,7 \times \text{leeftijd})$$

Vraag 1: HARTSLAG

Een krantenartikel meldde: 'Een gevolg van het gebruik van de nieuwe in plaats van de oude formule voor de aanbevolen maximale hartslag, is dat het aantal hartslagen per minuut voor jonge mensen licht afneemt en voor ouderen licht toeneemt.'

Vanaf welke leeftijd neemt de aanbevolen maximale hartslag toe als gevolg van het gebruik van de nieuwe formule?

Vraag 2: HARTSLAG

De formule $\text{Aanbevolen maximale hartslag} = 208 - (0,7 \times \text{leeftijd})$ wordt ook gebruikt om te bepalen welke fysieke training het meest effectief is. Onderzoek laat zien dat fysieke training het meest effectief is als de hartslag 80% van de aanbevolen maximale hartslag is.

Geef een formule om de meest effectieve hartslag te bepalen tijdens een training, uitgedrukt in leeftijd.

figuur 2 Opgave: Hartslag

Met dit fraaie PISA-probleem werd, zoals bij alle items, een field test uitgevoerd. Daarbij bleek dat maar 10% van de leerlingen dit item met succes kon afronden. Dit feit was voor het PISA-consortium reden het item niet mee te laten doen in de uiteindelijke itemcollectie.

Alleen al dit item roept om een nadere studie: hoe komt het dat een dergelijk item door 15-jarigen in alle landen van de OESO (zeg maar de rijkste landen) niet gemaakt kan worden? Is Nederland dus 'goed' of zijn we alleen maar beter dan de rest?

De toppositie bij Onzekerheid zal minder mensen verrassen: de voorsprong op andere landen is klein (Nederland 549, Finland 545, Canada 542) en de inhoud wordt wel eens gekarakteriseerd als vooral *common sense* redeneren. En dat zouden Nederlandse leerlingen redelijk kunnen, hoewel de resultaten bij de studie Probleemoplossen dit standpunt niet ondersteunen. Kortom, zo transparant als de resultaten lijken – wanneer gepresenteerd als een *horse race* – liggen de zaken zeker niet. En ook nadere analyses, zoals getracht als onderdeel van deze studie, roepen naast aanwijzingen voor verklaringen evenzoveel nieuwe vragen op.

Eén van de doelen die de OESO zich heeft gesteld, is een zogenaamde trendanalyse over de tijd. Hoe ontwikkelen de prestaties zich binnen tijdsintervallen van drie jaar? Uiteraard hier ook vanuit het relativiteitsbeginsel, in die zin dat voor *science* en *reading* het uitgangspunt de 500 (gemiddeld) van het jaar 2000 is, het jaar van het eerste PISA-onderzoek. Door bepaalde items hetzelfde te laten over de periode van drie jaar, kon men vaststellen dat het gemiddelde voor *science* in de drie jaar tussen 2000 en 2003 in vergelijkbare landen hetzelfde is gebleven. Doordat er in 2003 meer landen hebben meegedaan, is het gemiddelde in 2003 gedaald naar 496 met een standaarddeviatie van 105 scorepunten.

De trendkwestie

Nederland komt in de genoemde vergelijkende studies niet aan bod omdat we in 2000 werden gediskwalificeerd door de OESO: Nederland voldeed niet aan bepaalde statistische normen. Nederland vond dat het resultaat wél representatief was en in het eigen nationale rapport staat Nederland voor wiskunde dan ook fier aan top met 564 punten. In 2003 is dat 538 punten. Dat lijkt een flink verschil, maar hoe dat geïnterpreteerd moet worden is onduidelijk. Het Nederlandse rapport spreekt van 'waarschijnlijk een geringe achteruitgang'. Maar er zijn ook argumenten te bedenken die spreken voor een significante achteruitgang.

Allereerst moet opgemerkt worden dat voor wiskunde de norm van 500 voor het gemiddelde pas in 2003 kon worden gesteld omdat in de studie van 2000 slechts twee deelgebieden werden getest, namelijk Vorm en Ruimte en Veranderingen en Relaties.

Als we alleen de subschalen van Vorm en Ruimte en Veranderingen en Relaties nemen zien we dat Nederland in 2000 dus 564 scoorde en in 2003 slechts 538 (namelijk 526 voor Vorm en Ruimte en 551 voor Veranderingen en Relaties).

Dit is een geschatte achteruitgang van 26 punten. Tenminste, als het gemiddelde van 2000 en 2003 ongeveer hetzelfde is op deze twee subschalen. Het is de vraag of deze schalen wel goed te vergelijken zijn.

In de eerste plaats lijkt er geen reden te zijn om aan te nemen dat PISA 2000 niet een representatief beeld geeft van wiskunde als geheel, dus als geheel het domein wiskunde representerend. Het gemiddelde van Vorm en Ruimte en Veranderingen en Relaties in 2003 scheelt nauwelijks van het gemiddelde van Hoeveelheid en Onzekerheid. Ten tweede blijkt dat de gemiddelde score voor de beste achttien OESO landen van 2000 naar 2003 vrijwel constant is gebleven: zes landen gingen vooruit, zes bleven gelijk en zes gingen achteruit, met Nederland en Japan als negatieve uitschieters. Dit lijkt een aanwijzing te zijn die aangeeft dat de scores van 2000 en 2003 wel degelijk vergelijkbaar zijn.

Nu is het de vraag of de bewering van het Ministerie van OCW bij de resultaten van 2000 terecht heeft geclaimd dat de score toen representatief was. In reflectie kan gesteld worden dat de OESO gelijk heeft gekregen met haar bewering dat de Nederlandse score niet geheel representatief is geweest en wellicht tot maximaal tien punten naar beneden had moeten worden bijgesteld.

Dan nog is Nederland in de twee gemeten domeinen gedaald: bij Veranderingen en Relaties ging Nederland van 568 naar 551, bij Vorm en Ruimte van 537 naar 526. De eerste achteruitgang is significant, de tweede niet.

Ten slotte verwijzen we naar een andere belangrijke internationale studie: TIMSS (Trends in International Mathematics and Science Study). Deze constateert voor curriculaire wiskunde een niet-significante achteruitgang voor de leeftijdsgroep van 14-jarigen en een significante achteruitgang voor 8-jarigen, samen met Noorwegen (2003).

Met alle reserves kan gesteld worden dat het beeld van een eventuele trendontwikkeling in ieder geval niet positief te noemen is. Op de meest positieve manier geïnterpreteerd is er sprake van een lichte achteruitgang (over de gehele breedte). Waarschijnlijk meer realistisch is de conclusie dat het peil van het wiskundeonderwijs in Nederland, zoals gemeten met PISA danwel TIMSS, dalende is.

Van makkelijk naar moeilijk

De 84 items

Er zijn 84 items gebruikt bij de PISA-studie. Deze zijn allemaal uitgetest en besproken, ze zijn allemaal door nationale teams geaccepteerd, ze voldoen aan eisen van methodologische aard en ze zijn getoetst aan het PISA-framework. De opgaven zijn keurig verdeeld over de vier wiskundige deelgebieden, het gebruik naar context en de verschillende competenties: Reproductie, Verbindingen en Reflectie.

84 items met uiteraard een verschillende moeilijkheidsgraad – dat is een standaardvoorwaarde waaraan niet getornd kan worden. Juist de steeds groter wordende verzameling landen van buiten de OESO verlaagt het gemiddelde niveau (zie de discussie hierover voor *science* in het officiële OESO-rapport) en maakt het noodzakelijk dat er voldoende makkelijke opgaven zijn. Aan de andere kant is het zo dat moeilijkere opgaven meer inzicht kunnen geven over wat nu eigenlijk door leerlingen als moeilijk wordt ervaren. Zeker voor de betere leerling zijn moeilijkere items noodzakelijk om niet met allemaal 'uitstekende' resultaten te komen zitten.

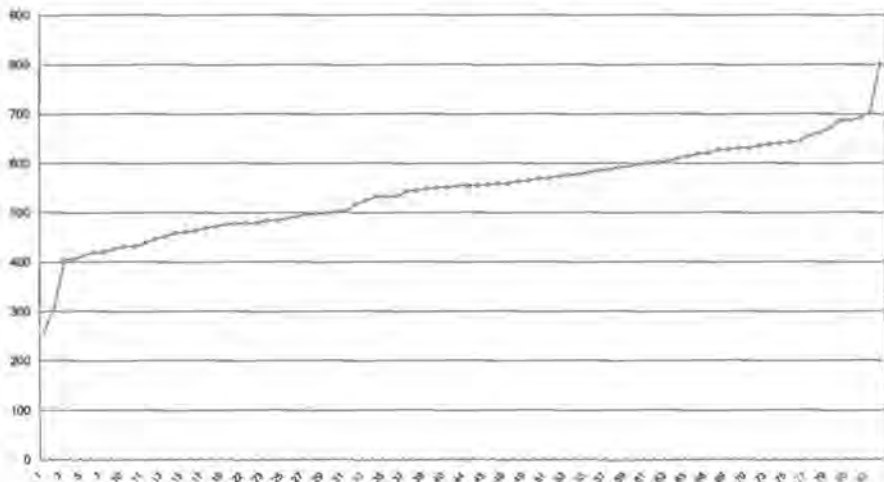
Voordat we dieper ingaan op de Nederlandse 'problemen', kijken we eerst naar de verdeling van de items naar moeilijkheidsgraad en of er iets te zeggen valt, qua inhoud of vorm, over wat een item moeilijk maakt. Dat dit een exercitie wordt op glad ijs is duidelijk, mede omdat de schrijvers van deze studie die items wel kennen maar ze, op enkele uitzonderingen na, niet bekend mogen maken. Toch denken we dat deze analyse de moeite waard is.

Allereerst enkele technische observaties: de waarden die de moeilijkheid van een item bepalen, zijn de zogenaamde drempel- of threshold-waarden. Deze lopen bij PISA 2003 van 262 voor een zeer eenvoudig item tot 801 voor het moeilijkste item. Om inzicht te krijgen in de resultaten in termen van niveaus van geletterdheid is een schaal ontworpen met een onderverdeling in verschillende vaardigheidsniveaus. Daarbij is de locatie van een specifieke leerling op de PISA-schaal aangewezen door specifiek die vraag waarbij de leerling een kans van 62% heeft om die vraag correct te beantwoorden.

Geletterdheidsniveau 1 begint bij 358 en loopt tot 420, niveau 2 van 420 tot 482 en niveau 3 van 482 tot 544. In dit gebied zit dus 500, het gemiddelde van de OESO. Daarna komt niveau 4 van 544 tot 606 en vervolgens niveau 5 van 606 tot 668. Daarna volgt het hoogste geletterdheidsniveau vanaf 668 zonder bovengrens.

Dit lijkt allemaal heel ingewikkeld, en dat is het ook. Niet voor niets publiceert de OESO een dik boek dat de technische achtergronden van deze indeling in detail schetst. Later in dit hoofdstuk zullen we ook naar wat meer laag-bij-de-grondse waarden kijken: de *p*-waarden, of het percentage van goede antwoorden.

De grafiek in figuur 3 geeft een verdeling van de items van makkelijk naar moeilijk: 81 van de 84 items liggen heel regelmatig over de drempelwaarden verspreid, maar de makkelijkste twee en de moeilijkste vallen echt buiten deze lineaire boot:



figuur 3 Verdeling items naar drempelwaarde

Makkelijk

Om een indruk te krijgen van bepaalde aspecten die de moeilijkheid zouden kunnen beïnvloeden, hebben we een analyse gemaakt van de 10 (15) makkelijkste en 10 (15) moeilijkste items.

Van de 10 gemakkelijkste items zijn er vier publiek gemaakt zodat u zelf mee kunt analyseren. Het zijn (van makkelijkst af gerekend) nummer 3: Wisselkoers vraag 1; nummer 6: Jongeren steeds langer vraag 2; nummer 7: Trap en nummer 8: Export vraag 1. In tabelvorm:

Tabel 2

Item rang /drempel	Naam	Domein	Competentie
3 / 406	Wisselkoers	Hoeveelheid	Reproductie
6 / 420	Jongeren steeds langer	Ver. en Relaties	Verbindingen
7 / 421	Trap	Vorm en Ruimte	Reproductie
8 / 427	Export	Onzekerheid	Reproductie

Zo'n klein tabelletje vertelt al heel veel. Wat opvalt (hoe bestaat het!) is dat we alle vier domeinen te pakken hebben en overwegend één competentie: Reproductie. Dit laatste blijkt aardig structureel: van de makkelijkste tien items vallen er acht onder Reproductie, van de makkelijkste vijftien zijn dat er twaalf. Ook verderop in dit hoofdstuk en in het rapport ziet men deze observatie bevestigd: PISA bevat weinig items op het reproductieniveau die moeilijk zijn. Verder valt op dat bij de vijftien makkelijkste items drie van de vier domeinen goed voorkomen – alleen het domein Onzekerheid komt maar éénmaal voor.

Om te kijken welke uiterlijke kenmerken deze makkelijke items hebben, reproduceren we ze in 'kale' vorm, dus alleen de opgave:

WISSELKOERS

Mei-Ling uit Singapore bereidt zich voor op een verblijf van drie maanden in Zuid-Afrika in het kader van een uitwisselingsprogramma voor studenten. Ze moet Singapore dollars (SGD) wisselen in Zuid-Afrikaanse rands (ZAR).

Vraag 1: WISSELKOERS

Mei-Ling ontdekte dat de wisselkoers van de Singapore dollar en de Zuid-Afrikaanse rand de volgende is:

$$1 \text{ SGD} = 4,2 \text{ ZAR}$$

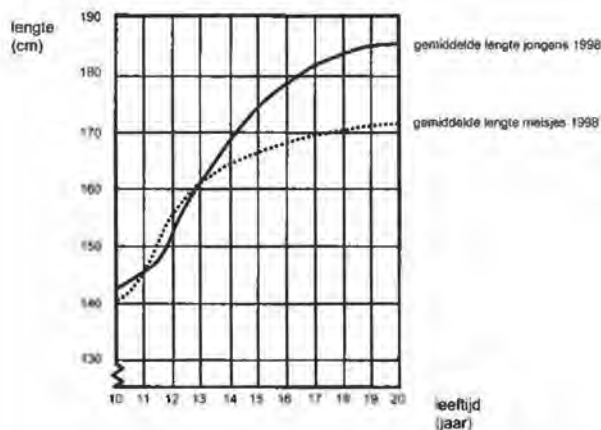
Tegen deze wisselkoers wisselde Mei-Ling 3000 Singapore dollars voor Zuid-Afrikaanse rands.

Hoeveel Zuid-Afrikaanse rands kreeg Mei-Ling?

figuur 4 Opgave: Wisselkoers, vraag 1

JONGEREN STEEDS LANGER

In deze grafiek wordt de gemiddelde lengte van zowel jongens als meisjes in Nederland in 1998 weergegeven.



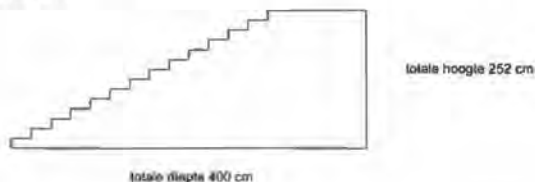
Vraag: JONGEREN STEEDS LANGER

Leg uit hoe je aan de grafiek kunt zien dat het gemiddelde groeitempo van meisjes langzamer wordt na hun 12de jaar.

figuur 5 Opgave: Jongeren steeds langer, vraag 2

TRAP

Hieronder zie je de afbeelding van een trap met 14 treden met een totale hoogte van 252 cm:



Wat is de hoogte van elk van de 14 treden?

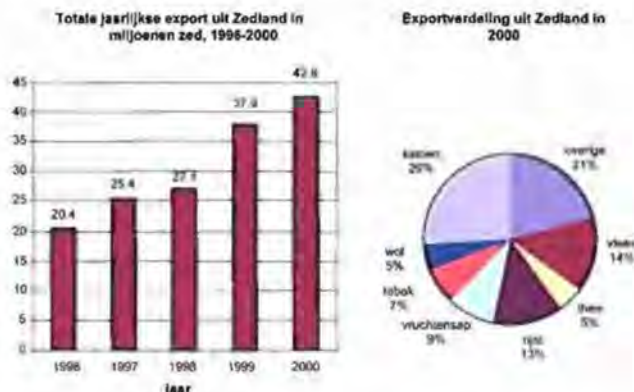
Hoogte: cm.

figuur 6 Opgave: Trap

Een enkelvoudig wisselkoersprobleem, een grafiekinterpretatie, een enkelvoudig deelprobleem, en het aflezen van een grafiek. Je kunt van leerlingen van vijftien jaar eenvoudig niets anders verwachten dan vrijwel 100% scores. En daar komen ze ook dicht bij in de buurt. Nederlandse leerlingen komen tot *p*-waarde scores van 82,4%,

EXPORT

In de onderstaande grafische voorstellingen vind je informatie over de export uit Zedland, een land dat de zed als munteenheid gebruikt.



Vraag 1: EXPORT

Wat was de totale waarde (in miljoenen zed) van de export uit Zedland in 1998?

figuur 7 Opgave: Export, vraag 1

89,2%, 84,4% en 90,7%. Dat is dus het gemiddelde over de hele Nederlandse populatie, van praktijkonderwijs via vmbo tot en met gymnasium. Overigens zijn er soms grote verschillen tussen de resultaten van de verschillende landen voor wat betreft moeilijkheid. Daarover later meer, maar het wisselkoersprobleem leidde in de Verenigde Staten tot een score van slechts 53,6% correct, wellicht doordat veel Amerikaanse jongeren zelden geld hoeven te wisselen. Ten slotte: Nederlandse leerlingen scoren voor de makkelijkste 28 items rond de 80%. Even anders geformuleerd: één derde van de items wordt door de leerlingen rond de 80% gescoord. Dit is aan de ene kant mooi, maar aan de andere kant betekent het dat we hier weinig informatie uit kunnen halen.

Moeilijk

Van de tien moeilijkste items zijn er slechts twee publiek gemaakt, waardoor u zelf heel beperkt mee kunt analyseren. Het zijn (van moeilijkst af gerekend) nummer 5: Timmerman en nummer 8: De beste auto vraag 2.

In tabelvorm:

Tabel 3

Item rang / drempel	Naam	Domein	Competentie
5 / 687	Timmerman	Vorm en Ruimte	Verbindingen
8 / 657	De beste auto	Ver. en Relaties	Reflectie

En als we doorgaan tot de moeilijkste vijftien:

13 / 636	Chatten	Hoeveelheid	Reflectie
----------	---------	-------------	-----------

Ook in dit geval reproduceren we de items in 'kale' vorm, dus alleen de opgave:

TIMMERMAN

Een timmerman heeft 32 meter planken en wil daarmee een rand om een bloemperk maken. Hij overweegt de volgende ontwerpen voor het bloemperk.

A

B

C

D

Omcirkel 'ja' of 'nee' voor elk ontwerp om aan te geven of het bloemperk met 32 meter planken gemaakt kan worden.

Ontwerp bloemperk	Kan met dit ontwerp het bloemperk worden gemaakt met 32 meter planken?
Ontwerp A	ja / nee
Ontwerp B	ja / nee
Ontwerp C	ja / nee
Ontwerp D	ja / nee

figuur 8 Opgave: Timmerman

DE BESTE AUTO

Een autoblad gebruikt een beoordelingssysteem voor nieuwe auto's en kent de prijs van 'Auto van het Jaar' toe aan de auto met de hoogste totaalscore. Er zijn vijf nieuwe auto's getest. In de onderstaande tabel staan de toegekende waarderingscijfers.

Auto	Veiligheid (V)	Brandstof- verbruik (B)	Ontwerp (O)	Interieur (I)
Ca	3	1	2	3
M2	2	2	2	2
Sp	3	1	3	2
N1	1	3	3	3
KK	3	2	3	2

De waarderingscijfers hebben de volgende betekenis:

3 punten = Uitstekend

2 punten = Goed

1 punt = Redelijk

Vraag 2: DE BESTE AUTO

De fabrikant van de auto 'Ca' vond de formule voor de totaalscore oneerlijk.

Geef een formule om de totaalscore zo te berekenen dat auto 'Ca' er als winnaar uitkomt.

Jouw formule moet elk van de vier variabelen bevatten. Geef jouw formule door positieve getallen in te vullen op de vier stippelijntjes in de onderstaande formule.

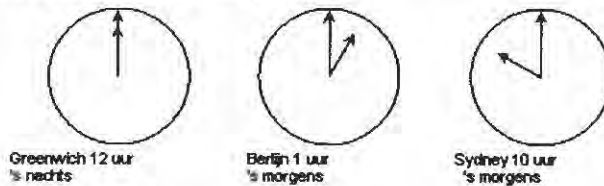
Totaalscore = V + B + O + I

figuur 9 Opgave: De beste auto, vraag 2

CHATTEN VIA INTERNET

Mark (uit Sydney in Australië) en Hans (uit Berlijn in Duitsland) hebben vaak contact met elkaar door te 'chatten' op het internet. Ze moeten zich tegelijkertijd op het internet bevinden om te kunnen chatten.

Om een goed tijdstip te vinden om te chatten, heeft Mark een overzicht van de wereldtijdzones opgezocht en heeft het volgende gevonden:



Vraag 2: CHATTEN VIA INTERNET

Mark en Hans kunnen niet chatten tussen 9 uur 's morgens en half 5 's middags hun lokale tijd, want dan zitten ze op school. Van 11 uur 's avonds tot 7 uur 's morgens hun lokale tijd kunnen ze ook niet chatten, want dan slapen ze.

Wat is een geschikt tijdstip voor Mark en Hans om met elkaar te chatten? Noteer de lokale tijden in de tabel.

stad	tijd
Sydney	
Berlijn	

figuur 10 Opgave: Chatten via internet, vraag 2

Timmerman bestaat in feite uit vier problemen. Dit is een merkwaardig formaat en daarom komen we hier op terug. De beste auto is een nogal constructieve vraag (dat is vaak al wat lastig) en bovendien moet er een formule gevonden worden. Het vinden van een raam in de tijd waarop twee ver van elkaar verwijderde personen kunnen internetten, is tamelijk complex om te lezen en te begrijpen, en vereist wat denkorganisatie van de leerlingen. Bij de laatste twee opgaven is het dan ook niet echt verwonderlijk dat deze als moeilijk werden gekwalificeerd.

Nederlandse leerlingen scoorden als volgt op deze drie items: 24,1%; 32,2% en 36,6%. Niet best dus. Al zijn het maar drie van de vijftien items, het volgende beeld klopt: alle vijftien items behoren of tot het competentie cluster Verbindingen, of tot Reflectie. Geen enkel item valt onder Reproductie, en dit is dus complementair aan het beeld dat we hadden van de tien (vijftien) makkelijkste items. Maar bij een analyse van de moeilijke items vallen nog andere zaken op. In de eerste plaats de zeer onevenwichtige distributie naar domein. Zes van de tien moeilijkste items toetsen competenties behorend bij het domein Vorm en Ruimte, terwijl Hoeveelheid helemaal niet voorkomt. Waarom scoort meetkunde zo hoog in moeilijkheid? Deze vraag zullen we straks nader aan de orde stellen.

Een ander opvallend feit is dat het toetsformaat bij deze opgaven bijna uitsluitend het open vraag-formaat is. Dertien van de vijftien items zijn van deze soort en dan bijna allemaal van het lang antwoord-type (slechts een van kort antwoord-type). De andere twee zijn van het complexe multiple choice-type, en gelukkig kunnen we dat formaat hier bespreken, want Timmerman is van dit, naar ons inzicht weinig waardevolle, formaat. Kijken we nader naar Timmerman, dan zien we dat hier vier vragen gesteld worden. Dat is op zich natuurlijk geen probleem. Het probleem is dat als je één antwoord fout hebt, het geheel als fout wordt gewaardeerd. Met andere woorden: dit item voldoet niet aan de voorwaarde dat je weet wat je meet. En dat is een nogal ernstige consequentie van dit formaat. Het feit dat dit item zo moeilijk is, is gewoon de lage score (zeggen methodologen). Maar cognitief gesproken lijkt het allemaal wel mee te vallen. Het lijkt niet al te gek te veronderstellen dat een grote meerderheid van de leerlingen zeker drie van de vier vragen goed zal beantwoorden.

Maar er is nog een probleem dat vaker geldt voor 'moeilijke' meetkunde problemen, namelijk dat de context er met de haren bij is gesleept. Timmerman bestaat helemaal niet; het probleem is gewoon een kaal wiskundig probleem. En dit kan ook een rol spelen in de relatieve moeilijkheid van het probleem: als het probleem niet als relevant en/of authentiek wordt ervaren, kan de motivatie van de leerling afnemen. De 'meetkunde' problemen uit de vijftien moeilijkste items hebben meer dan eens een gekunstelde context. Het zijn gewoon *inner mathematical* problemen, en wellicht ligt hier een oorzaak van de lage scores. Opvallend is ook dat drie van de vijftien moeilijkste vragen onderdeel zijn van dezelfde taak. Ook deze taak (die geheim is) is weer van meetkundige aard, maar nu eens met een zeer authentieke context uit de wereld van de atletiek. Maar hier zijn de vragen duidelijk van hoog wiskundig niveau en de score ligt min of meer in de lijn der verwachtingen.

Van overeenkomst tot verschil

Deze studie beperkt zich in eerste instantie tot Nederland en de buurgebieden Vlaanderen en Duitsland. Daarbij dient te worden opgemerkt dat Vlaanderen koploper is en zelfs beter scoort dan Finland. Naast de logische keuze voor onze burens binnen deze studie, kan men er niet omheen ook de winnaar van de *horse race*, Finland, in de vergelijking een prominente rol te laten spelen. Op sommige momenten nemen we ook Denemarken, Zweden en Noorwegen mee, zodat we een aaneengesloten regio in Noordwest Europa bestuderen. Ten slotte worden hier en daar de Verenigde Staten meegenomen in de analyse, mede vanwege hun leidende rol in de OESO. We kunnen dus de volgende clustering waarnemen:

E7:

Nederland	Nederland	Nederland	Nederland	Nederland
	Vlaanderen	Vlaanderen	Vlaanderen	Vlaanderen
	Duitsland	Duitsland	Duitsland	Duitsland
		Finland	Finland	Finland
		Noorwegen	Noorwegen	Noorwegen
		Denemarken	Denemarken	Denemarken
		Zweden	Zweden	Zweden
				Verenigde Staten

Voor het eerste deel van de vergelijking van de verschillende prestaties letten we vooral op verschillen tussen landen: zijn er items die zich niet 'normaal' gedragen; zijn er bepaalde structuren te ontdekken die relevant zouden kunnen zijn?

Allereerst lijkt het relevant de scores van deze acht landen in perspectief te zien:

Tabel 4

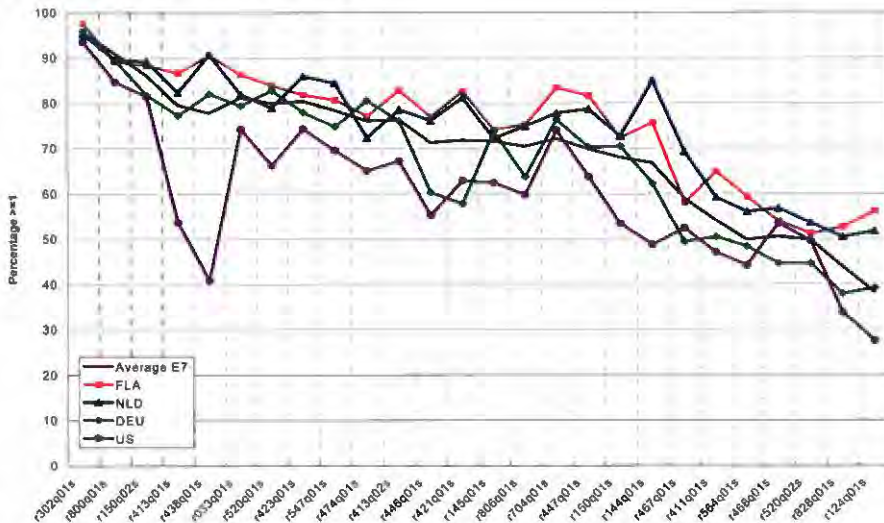
Vlaanderen	553
Finland	544
Nederland	538
Denemarken	514
Zweden	509
Duitsland	503
OESO	500
Noorwegen	495
Verenigde Staten	483

Het blijkt uit tabel 4 dat we een duidelijke top drie hebben die zich kan meten met de besten van de wereld. Sterker nog: Vlaanderen en Finland zijn wereldwijd nummer een en twee. Denemarken, Zweden, Duitsland en Noorwegen zitten in de grote middenmoot, en de Verenigde Staten houden alleen Zuid-Europese landen en Mexico achter zich.

Volgens het Duitse nationale rapport van PISA 2000 vindt men de conceptualisering met de competentieclusters een belangrijke bijdrage. De clusters Reproductie, Verbindingen en Reflectie dragen ertoe bij dat er verschillende *cognitive demands* zijn. We zagen al dat het niet overdreven is vast te stellen dat er veel Reproductie-items zitten bij de makkelijke items, en veel Reflectie-items bij de wat moeilijkere opgaven. De volgende drie grafieken (zie figuren 11, 14, en 17) leveren enkele interessante inzichten op: Het zijn de grafieken met *p*-waarden per cluster, gerangschikt naar de afnemende makkelijkheid over alle E7-landen.

De landen waarvoor deze waarden zijn uitgezet zijn Nederland, Vlaanderen, Duitsland en de VS.

De eerste grafiek (Reproductie)

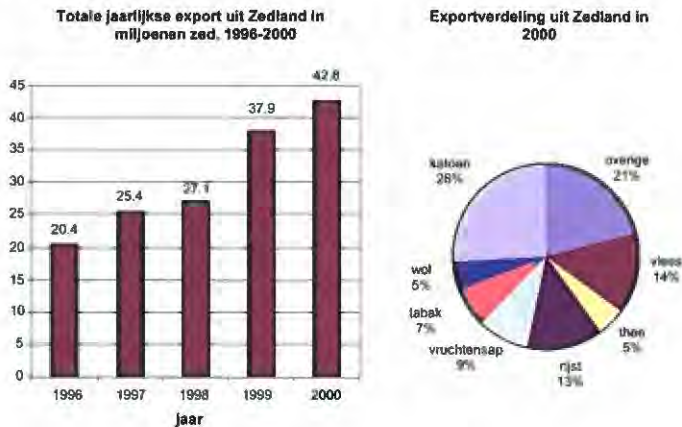


figuur 11 p-waarden Reproductie Vlaanderen-Nederland-Duitsland-VS

Wat valt er zoal te zeggen over deze grafiek? Voor de eerste elf items (de makkelijkste) liggen de scores van NVD (Nederland, Vlaanderen, Duitsland) zeer dicht bij elkaar. Dat is natuurlijk ook wel te verwachten omdat de items zo makkelijk zijn dat er maar een smalle band voor de scores mogelijk is. Maar daarna wordt de grafiek aanzienlijk onrustiger: Duitsland haakt nu duidelijk af van Nederland en Vlaanderen, op een paar items na. En Nederland doet enkele heel goede zaken hier. Maar de grootste afwijkingen komen op rekening van de VS. We hadden eerder al waargenomen dat de Amerikanen erg veel moeite hadden met het geldwissel-probleem (413), waarbij ze maar tot een p -waarde van 53,6 komen. En bij het item 438:1 wordt het nog erger: 40,8 tegenover een gemiddelde van dik boven de 80 voor de E7. Gelukkig is dit item niet geheim:

EXPORT

In de onderstaande grafische voorstellingen vind je informatie over de export uit Zedland, een land dat de zed als munteenheid gebruikt.



Vraag 1: EXPORT

Wat was de totale waarde (in miljoenen zed) van de export uit Zedland in 1998?

figuur 12 Opgave: Export

Het is vooralsnog niet eenvoudig, in tegenstelling tot bij het wisselkoersprobleem, een verklaring te vinden voor het collectief falen van de VS in dit geval. Het wordt nog raadselachtiger als we zien dat Onzekerheid het sterkste Amerikaanse domein blijkt te zijn (score 491). Als het toch moet, zou een verklaring kunnen zijn dat Amerikaanse leerlingen wellicht meer moeite hebben met het feit dat je voor het beantwoorden van de vraag het sectordiagram helemaal niet nodig had. Maar het blijft een verbaazingwekkende score.

Nederland doet het goed: met enige goede wil kan men stellen dat Nederland bij maar liefst dertien van de zesentwintig items koploper of vrijwel medekoploper is. De conclusie lijkt gewettigd dat Nederland het goed doet bij Reproductie. Dit feit denken we later nog te kunnen bevestigen. Als er één item boven uitsteekt voor Nederland, is het item 144:1. Helaas is dit een geheim item, maar er valt wel iets over te zeggen. Het betreft hier een item dat Nederlandse leerlingen erg bekend zal voorkomen, namelijk een opgave omtrent blokjes en een kubus. Het is eerder verrassend dat andere landen hier zo op Nederland achter blijven: Vlaanderen bijna tien punten, Duitsland maar liefst 23 en de VS zelfs 37 punten, terwijl daar dergelijke problemen toch ook frequent in de schoolboeken voorkomen.

Afgezien van het allermakkelijkste item (waar de scores maar 4 punten uit elkaar lopen) is het opvallend dat bij item 520:2 ook de scores van de vier landen nauwelijks verschillen: drie landen binnen vier punten en Duitsland daar weer vier punten onder.

SKATEBOARD

Erik is een fanatiek skateboarder. Hij bezoekt een winkel met de naam SKATERS om wat prijzen te bekijken.

In deze winkel kun je een compleet skateboard kopen. Of je kunt een plank, een set van 4 wieltjes, een set van 2 wielassen en een set losse onderdelen kopen en je eigen skateboard monteren.

De prijzen van de artikelen uit de winkel zijn:

Product	Prijs in zed	
compleet skateboard	82 of 84	
plank	40, 60 of 65	
een set van 4 wieltjes	14 of 36	
een set van 2 wielassen	16	
een set losse onderdelen (lagers, rubberen matjes, bouten en moeren)	10 of 20	

Vraag 2: SKATEBOARD

De winkel verkoopt drie soorten planken, twee soorten sets met wieltjes en twee soorten sets losse onderdelen. Er is slechts één mogelijke keuze voor een set wielassen.

Hoeveel verschillende skateboards kan Erik monteren?

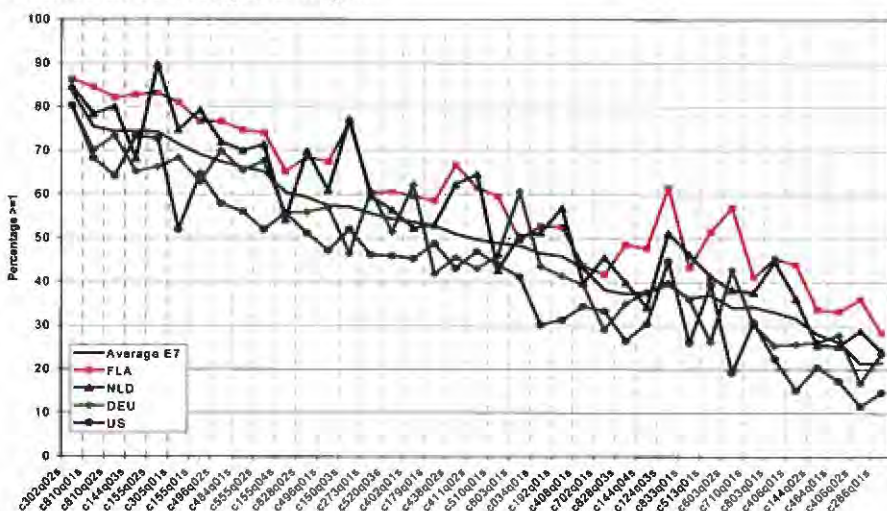
- A 6
- B 8
- C 10
- D 12

figuur 13 Opgave: Skateboard, vraag 2

Dit item (zie fig. 13) heeft klaarblijkelijk nauwelijks te lijden van cultuurverschillen of de verklaring kan erin liggen dat de context hier nauwelijks een rol speelt.

De tweede grafiek (Verbindingen)

In dit cluster bevinden zich heel veel items (dat was ook zo bedoeld volgens het Framework) en de grafiek is verrassend regelmatig. Duidelijk is dat Vlaanderen zich enigszins los begint te weken van Nederland en de rest. Ook daalt de grafiek van Vlaanderen heel regelmatig, totdat het echt moeilijk begint te worden: dan begint Vlaanderen ook hevig te bewegen.



figuur 14 p-waarden Verbindingen Vlaanderen-Nederland-Duitsland-VS

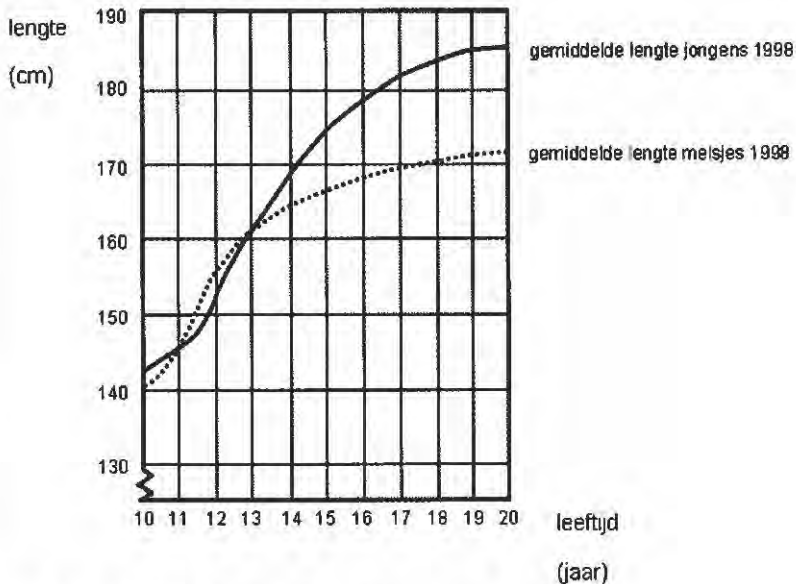
Nederland weigert zich op te stellen als een gemiddeld land: de volgorde van de moeilijkheid van items is voor Nederland een geheel andere dan die van de E7. Er zijn met name uitschieters naar boven: 155:2 levert ons een score van 90 op, een koppositie kortom. Dit item betreft (weer) het correct uitzoeken, lezen en interpreteren van een bijzondere grafiek, en dat kunnen Nederlands leerlingen goed. Het feit dat een dergelijk item werd ondergebracht in het domein Veranderingen en Relaties heeft weer bijgedragen aan de koppositie van Nederland op dit domein.

Een ander toppunt is de prestatie op item 150:3 waar Nederland nipt Vlaanderen voorblijft (77,1 versus 76,4), maar waarbij beide ver voor liggen op de E7 (51,9). Ook hier is de prestatie geen grote verrassing: weer dient een grafiek geïnterpreteerd te worden en deze keer is er echt sprake van een open vraag (zie fig. 15).

Ook dit item valt onder het domein Veranderingen en Relaties, en dat versterkt het gevoel dat we enig grip beginnen te krijgen op het feit dat Nederland 'wereldkampioen' Veranderingen en Relaties is geworden. Dat lijkt mede verband te houden met het hoge gehalte aan grafieken lezen en interpreteren en aan het ontbreken van wat formele algebra, zoals we dat kennen van de schoolcurricula. Het uiterst slecht scoren, ook van Nederland, op het eerder vermelde Hartslag-probleem, versterkt deze gedachte.

JONGEREN STEEDS LANGER

In deze grafiek wordt de gemiddelde lengte van zowel jongens als meisjes in Nederland in 1998 weergegeven.



Vraag 3: JONGEREN STEEDS LANGER

Tijdens welke periode in hun leven zijn, volgens deze grafiek, meisjes gemiddeld langer dan jongens van dezelfde leeftijd?

figuur 15 Opgave: Jongeren steeds langer

Maar Vlaanderen slaat terug bij met name items 124:2 en 603:2, waarbij we ook nog opmerken dat Nederland bij 124:2 nog goed meedoet maar bij 603:2 Duitsland voor zich moet dulden. Dat is nogal verrassend. 124:2 is gelukkig openbaar (zie fig. 16).

LOPEN

De afbeelding toont de voetafdrukken van een wandelende man. De paslengte P is de afstand tussen de achterkanten van de twee opeenvolgende voetafdrukken.

Voor mannen geeft de formule $\frac{n}{P} = 140$ bij benadering het verband weer tussen n en P , waarbij:

n = aantal stappen per minuut, en

P = paslengte in meters.

Vraag 2: LOPEN

Bernard weet dat zijn paslengte 0,80 meter is. De formule past bij het lopen van Bernard.

Bereken Bernards loopsnelheid in meter per minuut en in kilometer per uur. Laat zien hoe je tot je antwoord bent gekomen.

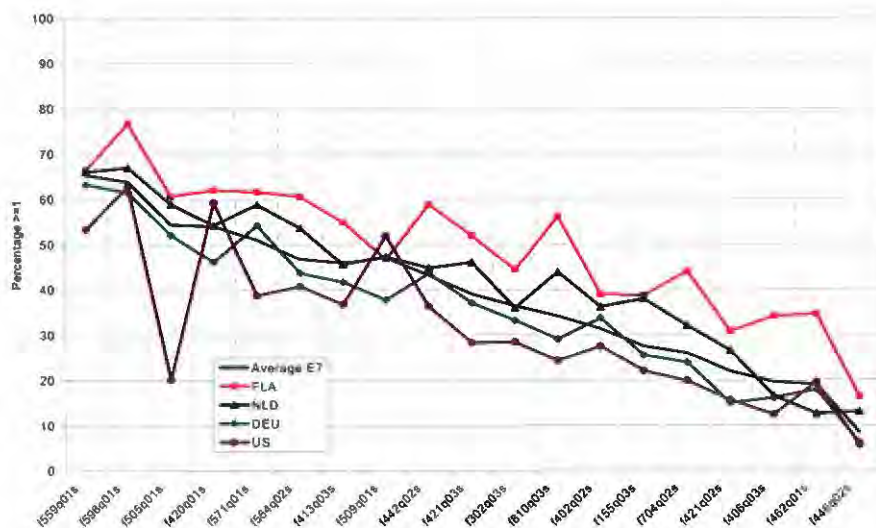
figuur 16 Opgave: Lopen

Het vermoeden wordt bevestigd: als het iets formeler wordt, haken Nederlandse leerlingen af – tot op zekere hoogte natuurlijk. Bovendien moet ook nog nauwkeurig met eenheden gewerkt worden. Al met al een relatief moeilijker opgave met voor de Nederlandse leerlingen een 51,0 score: ruim tien punten minder dan onze zuiderburen, maar altijd nog veel beter dan de oosterburen (39,7). Opmerkelijk genoeg hetzelfde resultaat als kampioen Finland.

Over opgave 603:2 kan opgemerkt worden dat dit een probleem is dat betrekking heeft op codes, en dan met name de rol van het controlegetal. Het is een nogal formeel probleem, met enkele regels die gevolgd moeten worden. De Vlamingen blinken nu uit: ze scoren voor dit lastige item nog altijd 56,9. Dit is echt uitzonderlijk gezien de scores van Finland en Duitsland: respectievelijk 43,2 en 42,6. Nederland heeft een score van 38,0 en de VS zakt helemaal weg met 19,1. Wat je noemt een discriminerend item. Deze scores lijken te bevestigen wat de Belgische Professor Janssens eens opmerkte: 'We hebben goed naar Nederland gekeken en overgenomen wat ons heel goed leek (contexten, toepassingen) maar zorgvuldig bewaard wat wij belangrijk vonden (meer formele aspecten).' 'Maar', voegde hij eraan toe: 'onze leerlingen stoppen nog voor een rood stoplicht en die houding maakt ook veel uit.'

De derde grafiek (Reflectie)

Deze grafiek (zie fig. 17) bevat duidelijk het minste aantal items. Tijdens het ontwerp-proces bleek het moeilijk om korte reflectie-items te ontwerpen, evenals moeilijke reproductie-items die zinvol zijn.



figuur 7 p-waarden Reflectie Vlaanderen-Nederland-Duitsland-VS

Bij de grafiek betreffende het Verbindingencluster viel al op dat Vlaanderen zich los ging maken van het veld, in de eerste plaats van Nederland. Deze trend wordt sterker doorgezet bij het Reflectiecluster: het is in dit competentiegebied dat Vlaanderen zich onderscheidt van de rest. Van de negentien items is er maar één (509:1) waarin Vlaanderen niet het beste is. Extra verrassend is dat de VS daar topscorer is. Dit is een nadere analyse waard: de VS boven Vlaanderen, Nederland en Duitsland. Gelukkig is dit wonderbaarlijke item openbaar: Aardbeving (zie fig. 18).

Dit is toch echt geen complex probleem lijkt het, maar ook elders blijkt het lastig om de moeilijkheid van items uit met name het domein Onzekerheid te voorspellen.

De grafiek heeft een rustig verloop: Vlaanderen ligt overal boven, op één uitzondering na. Nederland ligt tamelijk soeverein tweede, maar vaak dicht bij het E7-gemiddelde, wat geen hoopvol signaal lijkt. Duitsland ligt consistent onder Nederland, op een enkele uitzondering na. Wel is het opvallend dat de VS Nederland drie keer te slim af is. De eerste keer bij Aardbeving, dan bij 420:1 en bij 462:1. Het eerste item is redelijk te maken, gezien de E7-score van 53,3. De VS scoort 59,4, Nederland 54,4. Helaas wordt de moeilijkheid van dit item bepaald door het formaat, namelijk *complex multiple choice*. Het probleem op zich heeft een aardig uitgangspunt: er worden waarden en hun gemiddelden vergeleken, waarna enkele interpretatievragen worden gesteld. Maar omdat we niet kunnen zeggen (met enige mate van zekerheid) wat het *cognitive demand* is, is dit item niet geschikt om zinvolle uitspraken over te doen.

Voor item 462:1 worden verantwoorde uitspraken al helemaal moeilijk gezien de lage scores: het E7-gemiddelde is slechts 19,0. Nederland scoort 12,7 en de VS 19,7. En ja, we zien weer een eerder geuite gedachte bevestigd: Nederland heeft moeite met *inner mathematical* problemen. We treffen hier een standaardprobleem uit de vlakke meetkunde aan, abstract en wat formeel. Vlaanderen doet het hier uitzonderlijk goed in vergelijking met anderen: een score van 34,7 is ongeveer vijftien punten meer dan

AARDBEVING

Er is een documentaire uitgezonden over aardbevingen en hoe vaak die voorkomen. Hierin vond een discussie plaats over de voorspelbaarheid van aardbevingen. Een geoloog beweerde: 'De kans is twee op drie dat er een aardbeving in Zedstad zal plaatsvinden in de komende twintig jaar'.

Welke van de volgende beweringen geeft het beste weer wat de geoloog bedoelt?

- A. Aangezien $\frac{2}{3} \times 20 = 13,3$ zal zich in Zedstad over 13 tot 14 jaar een aardbeving voordoen.
- B. $\frac{2}{3}$ is meer dan $\frac{1}{2}$. Het is dus zeker dat op een bepaald moment in de komende 20 jaar een aardbeving zal plaatsvinden in Zedstad.
- C. Het is waarschijnlijker dat in Zedstad in de komende twintig jaar een aardbeving zal plaatsvinden dan dat er geen zal plaatsvinden.
- D. Je kunt niet voorspellen wat er gaat gebeuren, want niemand weet met zekerheid op welk moment zich een aardbeving zal voordoen.

figuur 18 Opgave: Aardbeving

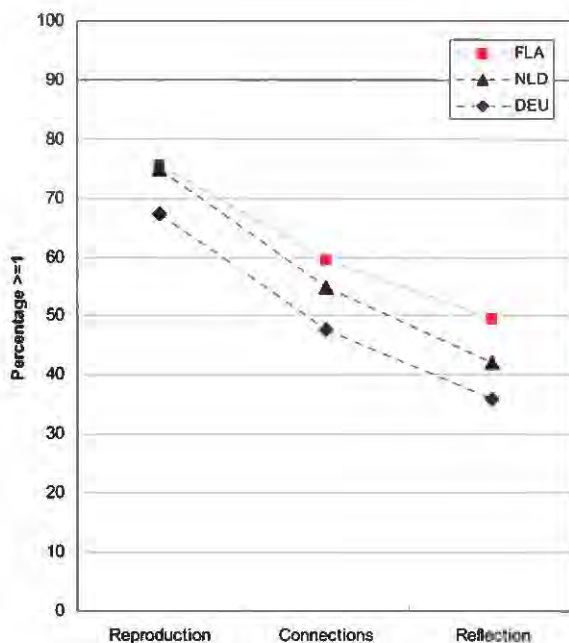
welke concurrent ook. Maar dat hadden we eigenlijk wel verwacht: hoe moeilijker het item, des te meer licht er komt tussen Vlaanderen en de rest. Het lijkt erop dat Vlaanderen niet zozeer excelleert in het Reproductiecluster, maar zich pas echt onderscheidt in het Verbindingencluster en nog pregnanter in het Reflectiecluster. Nederland en Finland moeten het vooral hebben van het Reproductiecluster. Dit is voor Finland wellicht te verwachten, maar voor Nederland zeker niet. Maar het achterblijven van Nederland bij het Reflectiecluster past wel bij de matige score van Nederland bij probleemoplossen.

We zullen eens wat verder kijken hoe de diverse landen zich gedragen over de competentieclusters.

Van reproductie tot reflectie

De analyse van de data tot dusver deed vermoeden dat Vlaanderen zich vooral onderscheidde bij items die behoren tot het competentiecluster Reflectie. Bij het Reproductiecluster waren er geen opmerkelijke verschillen tussen Vlaanderen en Nederland. Bij het Verbindingencluster maakte Vlaanderen zich los van Nederland en bij het Reflectiecluster werd dit effect nog duidelijker merkbaar, met als toegevoegde indicatie het feit dat hoe moeilijker de Reflectie-items werden, hoe groter de afstand Nederland – Vlaanderen. Deze observaties worden bevestigd als we verder speuren, met name bij de resultaten op het gebied van probleemoplossen.

Allereerst kijken we naar de grafiek die de gemiddelde *p*-waarden van de items per competentiecluster van Vlaanderen, Nederland en Duitsland weergeeft (zie fig. 19). Deze grafiek bevestigt duidelijk onze eerdere observaties: bij het Reproductiecluster scoren Vlaanderen en Nederland vrijwel gelijk (76 en 75), bij het Verbindingencluster



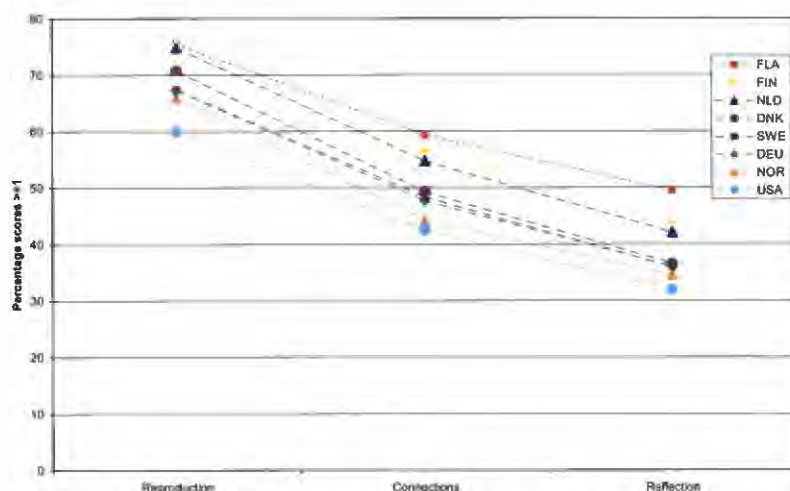
figuur 19 Gemiddelde p-waarden Vlaanderen, Nederland en Duitsland per itemcategorie

ligt Vlaanderen duidelijk voor (ruim 59 punten tegenover bijna 55) en bij het Reflectiecluster is dit opgelopen tot ruim zeven punten: bijna 50 en ruim 42. Dit verschil is vrij groot gezien de laagte van de scores: ruim vijf punten is rond 14% blijkt uit de precieze scores. Dit mag zeker opvallend genoemd worden.

Dat Nederland relatief niet zo goed scoort bij Reflectie blijkt ook uit het feit dat de achterstand van Duitsland op Nederland bij Reproductie het grootst is, iets minder bij Verbindingen en het kleinst bij Reflectie. Nu kan men terecht tegenwerpen dat deze trend logisch is, omdat hoe lager de drie scores worden, hoe dichter ze bij elkaar komen te liggen. Dit is uiteraard het geval, maar hier is het misschien dienstig om de wiskunderesultaten te vergelijken met die van probleemoplossen. Deze éénmalige studie werd in 2003 uitgevoerd en zeker 50% van de items heeft een erg wiskundig karakter. Vandaar dat we hier later in deze publicatie uitgebreider op terugkomen.

Wat opvalt als we naar de Nederlandse burens en Nederland kijken, is dat Vlaanderen in ieder geval weer hoog scoort en Duitsland heel dicht tegen Nederland aan zit. Het bevestigt dus de observatie dat naarmate het wiskundeonderwijs meer mathematisering en probleemoplossing en de daarbij behorende reflectie vereist, Nederland ten opzichte van de burens (en ook verdere burens) in een niet benijdenswaardige positie verkeert.

Wellicht is het aardig op dit moment de grafiek die we zojuist bekeken hebben in figuur 19, uit te breiden tot de E7 en de VS (zie fig. 20).



figuur 20 Gemiddelde p-waarden E7 en de VS

De tabel, behorende bij deze grafiek is tabel 5:

Tabel 5

	VL	FI	NL	DK	SE	DU	NO	VS
Reproductie	76	75	75	71	67	67	66	60
Verbinding	58	56	54	49	48	47	44	42
Reflectie	46	42	41	35	35	35	34	31

	VL	FI	NL	DK	SE	DU	NO	VS
OESO-score	553	544	538	514	509	503	495	483

Een preciezere analyse van de tabel en de grafiek kan ons leren dat de bijdrage van de scores in het Reproductiecluster het grootst is in Nederland, een uitkomst die nauwelijks nog als een verrassing kan komen. Als goede tweede komt Finland uit de bus, ook daar valt de winst te verklaren uit een hoge Reproductiescore. Denemarken blijft in het geheel redelijk op de been door alweer de scores in het Reproductiecluster. Daarna komt Vlaanderen pas.

Bij Zweden, Duitsland en Noorwegen is de bijdrage van het Reproductiecluster aan de totaalscore relatief hetzelfde. De VS doen het niet zo goed in dit cluster.

Bij het Verbindingencluster blijkt dat Vlaanderen daar een grotere bijdrage aan de totaalscore ontleent dan enig ander land, gevolgd door Finland en Nederland. Denemarken, Zweden en Duitsland verschillen niet erg voor wat de relatieve bijdrage van Verbindingen aan de totaalprestatie levert. Noorwegen en de VS vormen hier de staartgroep.

Dat in Vlaanderen de goede scores in het Reflectiecluster substantieel bijdragen aan een toppositie in het geheel, mag na al het voorgaande bijna als vanzelfsprekend worden beschouwd. Op ruime afstand volgen, met vergelijkbare bijdrage, weer Finland en Nederland. Daarna volgt een homogene groep: Denemarken, Zweden, Duitsland, Noorwegen en (op geringe afstand) de VS.

Resumerend kan gesteld worden:

- 1 Vlaanderen onderscheidt zich door een hoge score, vooral door relatieve suprematie in het Reflectiecluster, of door de relatief goede score bij de moeilijker op-
gaven. Eerder constateerden we al dat de zuiderburen het ook opmerkelijk goed
doen bij de wat formelere en abstractere opgaven.
- 2 Finland en Nederland hebben vergelijkbare profielen, met Finland net iets beter
in alle gebieden.
- 3 De oosterburen Duitsland laten zich beter vergelijken met de andere drie Scan-
dinavische landen Denemarken, Zweden en Noorwegen, waarbij de relatief
zwakke positie van Noorwegen opvalt.
- 4 De Verenigde Staten laten zich slecht vergelijken met de Europese landen uit de
noordwest hoek, wellicht met uitzondering van Vlaanderen. Het profiel is anders,
en de scores zijn relatief erg laag. Dit mag enigszins verrassend genoemd wor-
den vanwege de vermeende kleine culturele afstand.

3. Hoe werden de opgaven in Nederland gemaakt?

Truus Dekker, Monica Wijers

Het vorige hoofdstuk vergeleek de PISA-resultaten op het gebied van wiskunde van Nederlandse leerlingen met die van leerlingen uit een aantal andere landen. De resultaten van Nederland bij wiskunde waren in 2003 goed, maar een nadere analyse in dat hoofdstuk laat zien wat dat eigenlijk betekent. In dit hoofdstuk gaan we dieper in op de resultaten in Nederland voor verschillende groepen leerlingen. We gebruiken voorbeelden uit leerlingenwerk ter illustratie.

We laten daarbij de leerlingen uit het praktijkonderwijs buiten beschouwing, omdat dit een heel kleine en zeer diverse groep is. Daarnaast hebben we de resultaten van de 15-jarigen die nog niet in klas 3 of 4 van het voortgezet onderwijs zaten niet meegenomen in de nadere analyse.

Wat betekent wiskundige geletterdheid?

PISA beoogt *mathematical literacy* te meten, daaronder wordt verstaan in hoeverre de getoetste 15-jarigen in staat zijn hun wiskundige kennis en vaardigheden toe te passen in min of meer realistische situaties. Deze omschrijving zou heel goed bij het Nederlandse wiskundeonderwijs moeten passen waar immers ook veel nadruk gelegd wordt op het gebruiken van wiskundige kennis en vaardigheden binnen contexten. Toch zijn niet alle opgaven van het PISA-onderzoek goede contextopgaven. Een opgave als Timmerman, eerder besproken in hoofdstuk 2, is in feite een kale meetkundesom; de context is er met de haren bijgesleept. We noemen dat wel eens een nontext in plaats van een context. Overigens geldt dit ook vaak voor de contextopgaven in de Nederlandse schoolboeken; daarin komen veel nontexten voor. Soms was het beter geweest om maar niet net te doen alsof het om een authentieke situatie gaat, maar om de opgave gewoon binnen de context van de wiskunde zelf te plaatsen, met andere woorden, om er een 'kale' wiskunde opgave van te maken.

Verder is het natuurlijk de vraag wat wiskundige geletterdheid, wat geletterdheid op het gebied van natuurwetenschappen of zelfs wat gewoon geletterdheid eigenlijk betekent. Wij vinden dat het kunnen lezen en interpreteren van grafieken duidelijk tot 'wiskundige geletterdheid' hoort. Maar er zijn ook onderzoekers die vinden dat dit onder 'lezen van informatie' en dus onder leesvaardigheid valt. In de praktijk zijn de grenzen minder duidelijk dan uit de opdeling in verschillende vakken lijkt te volgen. Welke eisen stellen we aan wiskundige geletterdheid en vooral, aan wie stellen we die eisen? Voor een leerling van het vmbo zal wiskundig geletterd zijn iets heel anders betekenen dan voor een vwo-leerling die later natuurkunde gaat studeren. En er zijn ook zeker wiskundigen die niet erg wiskundig geletterd zijn.

Kan een leerling de tekst van de wiskundevraag wel begrijpen? En waar hoort dat dan bij? Moet een leerling die wiskundig geletterd is, ook over voldoende tekstbegrip beschikken? De eisen die we aan verschillende groepen leerlingen op dit gebied stellen voor het centraal examen zijn in ieder geval heel verschillend. De opgave Hartslag uit het vorige hoofdstuk werd slecht gemaakt maar volgens sommige deskundi-

gen is dat vooral te wijten aan de formulering van de vragen. Ook de inleidende teksten zijn moeilijk voor veel leerlingen. Aan het eind van dit hoofdstuk zullen we een voorbeeld geven van wat een docent met een dergelijke opgave in de klas zou kunnen doen.

Een ander probleem met de PISA-opgaven is dat de verschillen tussen de deelnemende landen erg groot zijn. Kijk maar eens naar vraag 3 van Jongeren steeds langer. In Noorwegen bijvoorbeeld beantwoordde ongeveer 60% van alle leerlingen de vraag fout (zie fig. 15, blz. 41), terwijl dat percentage in Vlaanderen 24% was. In landen als Brazilië en Mexico zijn die percentages ongetwijfeld veel hoger omdat die landen in het algemeen heel laag scoorden.

Daar heeft tot gevolg dat bij de opgaven van PISA voldoende eenvoudige vragen moeten zitten zodat ook zwakke leerlingen (of zwakke landen) verschillende scores kunnen halen. Je kunt het leerlingen niet aandoen ze een test voor te zetten waar ze uren aan moeten werken en waarvan je bij voorbaat weet dat ze bijna geen vraag kunnen beantwoorden. Vergelijk het bijvoorbeeld met de toetsen voor de basisvorming die we een tijdlang in Nederland hadden. Voor de vwo- en havo-leerlingen waren de opgaven vaak veel te simpel, waardoor ze een heel verkeerde indruk kregen van wat het vak inhoudt. Niks aan, dat doe je fluitend. Voor de leerlingen uit de beroepsgerichte leerwegen waren de opgaven juist veel te moeilijk. Lastig vak hoor, die wiskunde, typisch iets voor de mensen met wiskunde-knobbels, was voor hen een voor de hand liggende conclusie.

De PISA-opgaven vormen dus hoe dan ook een compromis en globaal gesproken geeft naar onze mening het niveau ervan eerder de te verwachten wiskundige geletterdheid aan van leerlingen die een beroepsgerichte opleiding volgen dan van leerlingen die mogelijk na hun examen naar hogeschool of universiteit gaan.

We kunnen, als we alleen naar Nederland kijken, binnen de opgavenverzameling opgaven onderscheiden die door de verschillende groepen leerlingen heel verschillend worden gemaakt. We laten daarvan enkele voorbeelden zien.

Verschillende strategieën en het gebruik van de Double Digit code

Bij het beoordelen van de antwoorden op de PISA-toets is een aantal antwoorden op open vragen dubbel gecodeerd. Naast een cijfer dat het behaalde aantal scorepunten voor de vraag aangeeft (meestal 0, 1, 2 of 3), is een cijfer toegevoegd dat informatie geeft over de wijze van oplossen. Dit tweede cijfer kan bijvoorbeeld verwijzen naar het al dan niet noteren van bepaalde deelstappen of tussenantwoorden of naar een bepaalde oplossingsstrategie, of naar bepaalde fouten die een leerling maakt. Code 99 geeft aan dat er niets is ingevuld of dat hetgeen opgeschreven is onleesbaar is of doorgestreept. Op die manier zou bekeken kunnen worden wat voor strategieën leerlingen gebruiken om een probleem op te lossen: maken ze een tekening? Gebruiken ze een informele manier? Worden formele oplossingsmethoden gehanteerd? Slaan leerlingen gemakkelijk een vraag over of proberen ze iets?

Het extra cijfer in de codering blijkt meestal uitsluitend te verwijzen naar een onderdeel van de vraag waarop een deelscore is behaald, zoals in het voorbeeld in figuur 1, en niet naar verschillende strategieën. Uit de codering bleek dat Nederland-

se leerlingen over het algemeen minder vaak dan leerlingen in andere OESO-landen een vraag oningevuld laten. Meestal wordt er wel iets geprobeerd.

Gedeeltelijk goed (2 punten)

Score 21:

Zoals bij score 3, maar laat na met 0,80 te vermenigvuldigen om de stappen per minuut in meter per minuut om te zetten, bijvoorbeeld zijn snelheid is 112 meter per minuut en 6,72 km/u.

- 112; 6,72km/u

Score 22:

De snelheid in meter per minuut is goed (89,6), maar de omzetting naar kilometer per uur is fout of ontbreekt.

- 89,6 m/min; 8960 km/u
- 89,6; 5376
- 89,6; 53,76
- 89,6; 0,087km/u
- 89,6; 1,49km/u

Score 23:

Correcte methode (expliciet aanwezig) met klein(e) rekenfoutje(s), andere dan bedoeld in de eerder vermelde scores 2. Geen antwoorden goed.

- $n = 140 \times 0,8 = 1120$; $1120 \times 0,8 = 896$. Hij loopt 896 m/min; 53,76km/u
- $n = 140 \times 0,8 = 116$; $116 \times 0,8 = 92,8$. 92,8 m/min -> 5,57km/u

Score 24:

Vermeldt alleen 5,4 km/u, maar niet 89,6 m/min (tussentijdse uitwerking ontbreekt).

- 5,4
- 5,376 km/u
- 5376 m/u

figuur 1 Voorbeeld van scoremodel met Double Digit codering

In het voorbeeld in figuur 1, horend bij vraag 2 van de opgave Lopen (zie fig. 16, blz. 41), is te zien dat de deelscore van 2 punten (van de 3) toegekend wordt als er een stap ontbreekt in het oplossingsproces of als er kleine rekenfoutjes zijn gemaakt. Aan de code 1, 2 of 4 (het tweede cijfer) is te zien welke stap ontbreekt, de code 3 geeft aan dat er rekenfoutjes zijn gemaakt.

Opvallend is dat de goedscore van 3 punten voor deze vraag niet een tweede code krijgt om verschillende oplossingsmethoden aan te duiden. In het correctievoorschrift staat:

Score 3 wordt gegeven als beide antwoorden goed zijn (89,6 en 5,4), ongeacht of er een uitwerking is of niet. Afrondingsfouten zijn acceptabel.

We hoopten in dit onderzoek door een nadere analyse van de Double Digit codes een beeld te kunnen geven van verschillen in strategieën tussen de verschillende leerlingengroepen. We verwachtten dat leerlingen uit het vmbo vaker informele strategieën zouden gebruiken om tot een oplossing te komen dan de havo/vwo-leerlingen. Informele strategieën zijn bijvoorbeeld: *trial en error*, proberen met getallen, een tekening

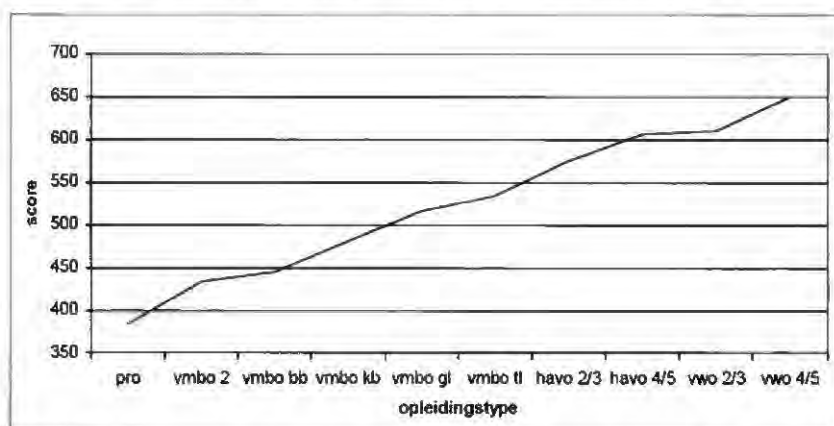
maken, en dergelijke. Ook verwachtten we dat de vmbo-leerlingen vaker een lang antwoord in dagelijkse taal zouden geven dan in meer formele wiskundetaal. We hoopten dat de *Double Digit* codering ons zou helpen dergelijke verschillen kwantitatief te beschrijven. Helaas bleek dit bij nadere analyse van de manier waarop de antwoorden zijn gecodeerd niet mogelijk te zijn.

Van de negen vragen die dubbel gecodeerd zijn, zijn er slechts twee waarbij de dubbele codering expliciet verwijst naar het (in)formele karakter van het gegeven antwoord. Dat is het geval bij een van de vragen in de vrijgegeven opgave Jongeren steeds langer. Voor deze vraag blijkt er geen significant verschil te bestaan in het type antwoord tussen vmbo-leerlingen enerzijds en havo/vwo-leerlingen anderzijds. Het grootste deel van de leerlingen in elk schooltype dat de vraag goed heeft, beantwoordt deze in dagelijkse taal.

Om beter zicht te krijgen op verschillen in antwoordgedrag hebben we van een aantal vrijgegeven opdrachten het leerlingenvoerwerk bekeken en geanalyseerd.

Hoe werden kale opgaven en formele (algebra-)opgaven in Nederland gemaakt?

Het Nederlandse schoolstelsel is een selectief stelsel. Dat blijkt ook uit de scoreverdeling op de wiskundeschaal, uitgesplitst naar de verschillende schooltypen. Naarmate het opleidingstype hoger is, verschuift ook de scoreverdeling.

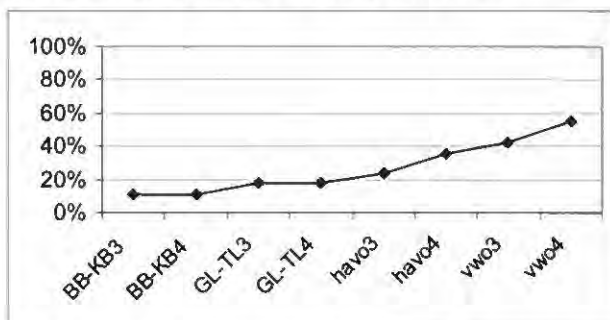


figuur 2 Gemiddelde wiskundescore per opleidingstype

De groepen pro (praktijkonderwijs), 15-jarigen in klas 2 van het vmbo en vmbo-BB/KB, in totaal ongeveer $\frac{1}{3}$ deel van de Nederlandse 15-jarigen, scoren wat lager dan de gemiddelde PISA-score van 500; $\frac{2}{3}$ deel scoort boven het gemiddelde.

Dit geldt voor de totale verzameling van de 84 PISA-opgaven, maar hoe zit dat nu met opgaven die niet zo goed passen binnen het Nederlandse wiskundeonderwijs, de kale en/of formele opgaven uit het PISA-onderzoek? In het algemeen zien de grafieken van de resultaten bij dit type opgaven er tamelijk steil uit. Met andere woorden: er is een vrij groot verschil tussen de resultaten van vmbo-BB/KB tot vwo. Kijk maar eens naar de resultaten in figuur 53 op een kale meetkundeopgave van hetzelfde

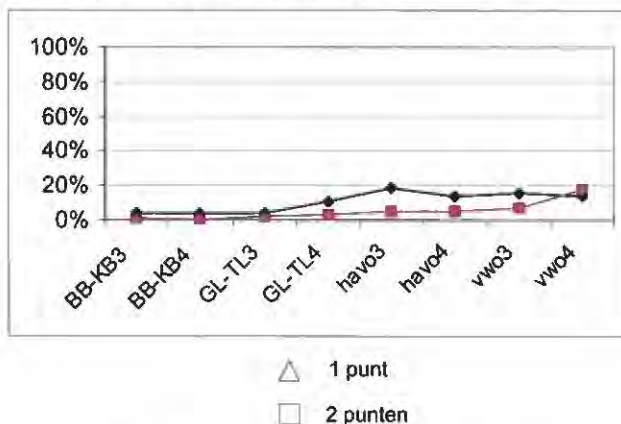
type als de eerder besproken opgave Timmerman. De – niet vrijgegeven – opgave is niet moeilijk maar de context is niet zinvol. Een nontext dus.



figuur 3 Resultaten meetkunde opgave met nontext

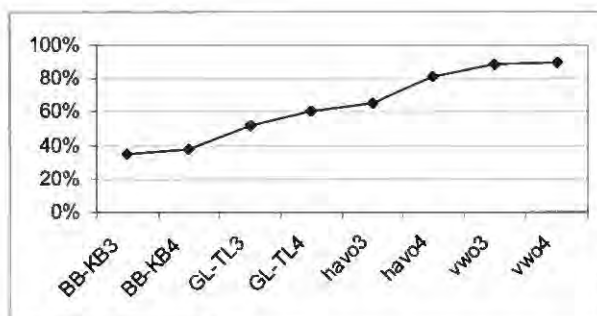
De leerlingen in het vmbo, met uitzondering van vmbo-GL/TL kunnen hier niets mee, de leerlingen uit havo en vwo scoren beter maar, gezien het niveau van de opgave, niet echt goed. De opgave staat in de rangordes voor de drie verschillende schooltypen vmbo BB/BK, vmbo GL/TL en havo/vwo overigens nagenoeg gelijk rond de 80e plaats. Het is een opgave die leerlingen uit havo en vwo, gezien de wiskundige inhoud, zeker aan zouden moeten kunnen.

De opgave, ook met een wiskundige inhoud die uit de meetkunde komt, waarbij het scoreverloop in figuur 4 hoort is niet alleen kaal maar ook heel formeel. Er is geen context. Zo'n opgave wordt door alle Nederlandse leerlingen slecht gemaakt. Het is een zeer moeilijke opgave uit het reflectiecluster.



figuur 4 Resultaten formele meetkunde-opgave

Ten slotte in figuur 5 een voorbeeld van het scoreverloop van een formele algebra-opgave. De vraag heeft betrekking op een kwadratische letterformule waarbij gevraagd wordt naar het effect van een wijziging in de getallen die in de formule voorkomen. De vraag is moeilijk, de context is zinvol.



figuur 5 Resultaten formele algebra-opgave

Bij deze resultaten valt op dat de BB/KB-leerlingen, die zelden een kwadratische formule zien en die in het algemeen alleen woordformules voorgeschoteld krijgen, hier toch in het algemeen beter op scoorden dan bij andere formele algebra-opgaven. Zou het type vraag, een meerkeuzevraag, daar iets mee te maken kunnen hebben? Dat is verder niet na te gaan.

Een opgave nader bekeken

De volgende opgave (zie fig. 6) willen we wat nader bekijken.

STEUN VOOR DE PRESIDENT

In Zedland zijn opiniepeilingen gehouden om te kijken hoeveel steun de president bij de komende verkiezingen zal krijgen. Vier krantenuitgevers hielden afzonderlijke, landelijke peilingen. De resultaten van de vier krantenpeilingen zie je hieronder:

Krant 1: 36,5% (peiling gehouden op 6 januari, op een willekeurig getrokken steekproef van 500 stemgerechtigden);

Krant 2: 41,0% (peiling gehouden op 20 januari, op een willekeurig getrokken steekproef van 500 stemgerechtigden);

Krant 3: 39,0% (peiling gehouden op 20 januari, op een willekeurig getrokken steekproef van 1000 stemgerechtigden);

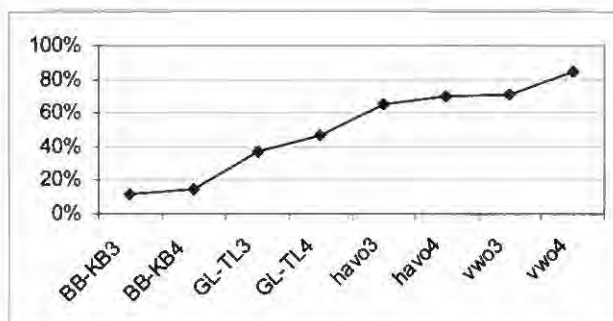
Krant 4: 44,5% (peiling gehouden op 20 januari, op 1000 lezers die belden naar de krant om hun mening te geven).

In welke krant voorspelt de opiniepeiling waarschijnlijk het beste hoeveel steun de president zal krijgen als de verkiezingen gehouden worden op 25 januari? Geef twee argumenten om je antwoord toe te lichten.

figuur 6 Opgave: Steun voor de president

Het is een opgave over een onderwerp dat het merendeel van de Nederlandse leerlingen op 15-jarige leeftijd meestal (nog) niet gezien heeft, namelijk het kiezen van een steekproef. In de bovenbouw van havo en vwo wordt overigens wel aandacht be-

steed aan de kwaliteit van een steekproef. De opgave is van het type lang antwoord vraag, dus moet er een redenering worden opgeschreven. Dat is vooral voor leerlingen uit het vmbo vaak niet eenvoudig. De opgave had een scoreniveau van 615 op de PISA-schaal. Het is een opgave uit het Verbindingencluster en het domein is Onzekerheid.



figuur 7 Resultaten in Nederland op Steun voor de president

In Nederland had gemiddeld 45% van alle leerlingen de vraag voldoende beantwoord (zie fig. 7). Zoals te verwachten was, is het goed beantwoorden van deze vraag voor BB-KB-leerlingen een probleem, slechts een enkeling weet er raad mee. De volgende leerling had in ieder geval een deel van het antwoord goed:

Leerling BB-KB-4

*Krant 2. Omdat de mensen niet hoefden op te bellen, want er zijn ook mensen die de president wel steunen maar niet naar de krant bellen.
En 20 januari is dichterbij 25 januari dus dan weten de mensen ook al beter wat ze willen.*

Een leerling uit de GL-TL-groep die een goed antwoord gaf:

Leerling GL-TL-4

Krant 3. Het is vlak voor de verkiezingen (iedereen weet een beetje op wie hij wil gaan stemmen). Ze hebben een steekproef gedaan onder veel mensen. Hoe meer mensen des te beter de steekproef.

Deze leerling uit havo/vwo heeft niet veel vertrouwen in onbevooroordeelde berichtgeving:

Leerling havo/vwo

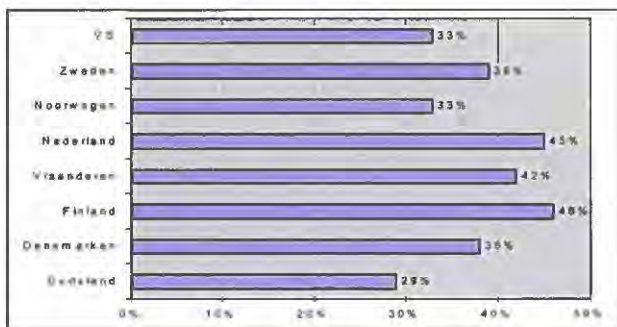
Krant 3. Ze hebben duizend mensen gevraagd en niet 500. Het zijn willekeurige mensen dus niet alleen de lezers van krant 3, dus mocht deze krant propaganda uitoefenen dan had dat geen effect op de ondervraagden.

Helemaal fout is de volgende redenering:

Leerling havo/vwo

Krant 4. Het zijn de lezers hun eigen mening; ze belden zelf op. Bij de andere was het maar een willekeurige steekproef.

De resultaten in de ons omringende landen en de VS zijn minder goed dan in Nederland (zie fig. 8), alleen de Finse leerlingen haalden een vergelijkbaar resultaat. Argumenten geven op basis van de informatie die verstrekt wordt in de opgave; daar zijn Nederlandse leerlingen relatief goed in.



figuur 8 Resultaten verschillende landen op Steun voor de president

Nederlandse leerlingen kunnen over het algemeen goed redeneren of hun mening verwoorden aan de hand van een grafiek of diagram; zowel binnen het domein van de algebra als dat van de statistiek. Aan dat onderwerp wordt overigens in de schoolboeken ook veel aandacht besteed. Goed formuleren van het antwoord is soms wel moeilijk, dat geldt vooral voor leerlingen in de lagere schooltypen en wellicht ook voor leerlingen van allochtone afkomst. Het is soms lastig om te beoordelen of een leerling, gezien zijn of haar formulering van het antwoord, een goede of gedeeltelijk goede redenering heeft gegeven. Wanneer je, zoals dat voor een docent geldt, de leerling goed kent, helpt dat bij de beoordeling. Voor een internationaal onderzoek is dat natuurlijk niet zo.

Waar hebben Nederlandse leerlingen moeite mee?

Het kunnen interpreteren van grafieken en diagrammen en het redeneren op basis van het verloop van een grafiek kwam betrekkelijk veel voor bij de PISA-opgaven. Het verklaart wellicht de goede score van Nederland op het domein Veranderingen en Relaties.

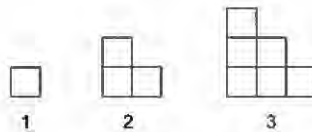
Opgaven uit hetzelfde domein waarbij een formule gebruikt of opgesteld moest worden, maken Nederlandse leerlingen veel minder goed dan de opgaven met grafieken. Toch is het kunnen opstellen van een formule en het hebben van inzicht in de structuur van een formule een belangrijke algebraïsche vaardigheid. Daar zou misschien in het wiskundeonderwijs meer aandacht aan besteed moeten en kunnen worden.

Wanneer er erg veel informatie in grote stukken tekst staat, haken Nederlandse leerlingen snel af. Dat doen ze ook wanneer er veel stappen nodig zijn om het antwoord te vinden. Bij het beoordelen van het leerlingenwerk zie je vaak dat de leerlingen wel beginnen maar stoppen wanneer het uitwerken te lang duurt of te lastig is. Ze geven betrekkelijk snel op. Duitse leerlingen zijn bijvoorbeeld beter dan Nederlandse kinderen wanneer het gaat om ingewikkelde rekenopgaven waarbij een bepaalde regel, vaak in een aantal stappen moet worden toegepast. Ze geven misschien minder snel op? Positief is in ieder geval dat Nederlandse leerlingen niet bang zijn om wel iets te proberen: ze slaan in vergelijking met andere landen weinig opgaven over.

Verder zagen we weinig kladwerk, ook niet wanneer daar uitdrukkelijk ruimte voor gegeven werd. Zijn de leerlingen gewend dat er op het papier van proefwerkopdrachten

niet geschreven mag worden? Gebruikten ze losse kladblaadjes? Het is bekend dat meerkeuzevragen als het ware uitnodigen tot het beantwoorden van de vraag 'uit het hoofd'. Uiteraard is niet na te gaan of dat ook bij de PISA-opgaven het geval was.

Bij een opgave zoals Trappatroon (zie fig. 9), waarbij een simpel patroon van vierkantjes moet worden voortgezet zou je verwachten dat veel leerlingen een tekeningetje maken van het vierde patroontje en de vierkantjes tellen. Dat is zelden of nooit te zien op de uitwerkingenbladen en als het wel gebeurt, wordt soms patroon nummer vier aan patroon nummer drie vastgetekend waarna het aantal vierkantjes dat erbij gekomen is wordt geteld. Eén leerling, die wel een tekening maakte, was een vierkantje vergeten. Slordig? Gauw tevreden? In het algemeen werden weinig tekeningen als uitleg of als kladwerk gezien bij de opgaven.



Hoeveel vierkantjes zijn er nodig voor stap vier?

figuur 9 Opgave: Trappatroon

Een voorbeeld om in de klas te gebruiken

De PISA-opgaven passen over het algemeen goed bij het Nederlandse wiskundeonderwijs. Redeneren op basis van de structuur van een formule, en formules opstellen en gebruiken scoren veel minder hoog. Daarom hier een voorbeeld van een (fictieve) klassendiscussie waarbij uitgegaan is van de opgave Harts slag uit hoofdstuk 2 (blz. 20) die, zoals gemeld, niet in de opgavenset werd opgenomen omdat maar weinig leerlingen de vragen goed konden beantwoorden.

Wat is ongeveer een normale harts slag voor iemand van jouw leeftijd?

Voor een volwassene is dat ongeveer 60 – 80 slagen per minuut. Een pasgeboren gezonde baby heeft een harts slag van meer dan 100 slagen per minuut. Voor een topsporter in rust kan de harts slag wel minder dan 30 slagen per minuut zijn.

Wat gebeurt er tijdens inspanning met je harts slag?

Tijdens inspanning gaat de harts slag omhoog, het aantal slagen per minuut neemt toe.

De aanbevolen maximale harts slag tijdens sporten is afhankelijk van de leeftijd. Hiervoor bestond jarenlang de volgende formule: *maximale harts slag* = 220 – *leeftijd*.

Wie heeft een hogere maximale harts slag, iemand uit je klas of een van de docenten?

De maximale harts slag voor een leerling is hoger. Voor iemand die ouder is wordt het getal dat van 220 wordt afgetrokken immers groter.

Hoe ziet de grafiek eruit die bij de formule hoort? Waarom weet je dat zeker?

De grafiek is een rechte lijn die begint in het punt (0, 220) en eindigt in het punt (220, 0). Je weet zeker dat het een rechte lijn is want voor elk jaar dat je naar rechts gaat (in de horizontale richting) wordt de maximale hartslag één minder.

Recent onderzoek heeft laten zien dat de formule moet worden bijgesteld. Dit is de nieuwe formule: *maximale hartslag* = $208 - (0,7 \times \text{leeftijd})$.

Een krantenartikel meldde: 'Een gevolg van het gebruik van de nieuwe in plaats van de oude formule voor de aanbevolen maximale hartslag, is dat het aantal hartslagen per minuut voor jonge mensen licht afneemt en voor ouderen licht toeneemt.'

Vanaf welke leeftijd neemt de maximale hartslag toe als gevolg van het gebruiken van de nieuwe formule?

Je kunt de grafieken tekenen (eventueel met de grafische rekenmachine) en naar het snijpunt kijken.

Je kunt naar de tabel kijken (eventueel via de grafische rekenmachine)

Je kunt de vergelijking oplossen:

$$220 - \text{leeftijd} = 208 - (0,7 \times \text{leeftijd})$$

$$220 - (0,3 \times \text{leeftijd}) = 208$$

$$(0,3 \times \text{leeftijd}) = 12$$

$$\text{leeftijd} = 40$$

Als je gaat trainen kun je beter niet de maximale hartslag aanhouden. De training is het meest effectief wanneer je hartslag 80% van de aanbevolen maximale hartslag is.

Maak een nieuwe formule die met deze gegevens rekening houdt.

effectieve hartslag =

$$\text{effectieve hartslag} = 0,8 \times (208 - (0,7 \times \text{leeftijd})) \text{ of}$$

$$\text{effectieve hartslag} = 166,4 - (0,56 \times \text{leeftijd})$$

Bespreek met leerlingen of ook

effectieve hartslag = $166 - (0,6 \times \text{leeftijd})$ een goede formule is en waarom (niet).

Stel je voor dat je in plaats van 80% het getal 60% invult in de formule die je gemaakt hebt.

Hoort de nieuwe formule dan bij intensief trainen of bij licht trainen?

De formule is geschikt voor licht trainen. De aanbevolen hartslag is dan immers lager.

4. PISA-resultaten in het vmbo

Monica Wijers, Truus Dekker

Een van de drie deelvragen van de studie waarover dit rapport handelt, was specifiek gericht op het in beeld brengen van de PISA-resultaten van de Nederlandse vmbo-leerlingen. In het onderzoeksvoorstel was de vraag als volgt geformuleerd:

Hoe scoren de Nederlandse leerlingen van vmbo tot gymnasium, en welke inhoudelijke verschillen zijn er tussen de te onderscheiden groepen? De nadruk zal liggen op de resultaten bij het vmbo, daar er op het moment een grote behoefte is naar meer inzicht in de kwaliteiten en vaardigheden van deze leerlingen; zowel onderling vergeleken, als in relatie tot havo/vwo-leerlingen van 15 jaar. Daarbij is onder andere voorzien in:

- vergelijking van scores,
- vergelijken van strategieën (dit is mogelijk door *double digit coding* van de resultaten en inzicht in het leerlingmateriaal),
- vergelijken van kenmerken van opdrachten waarop de verschillende groepen onderscheidend scoren,
- identificeren van singulariteiten.

In dit hoofdstuk gaat het om het beantwoorden van deze vraag. Eerst typeren we kort het vmbo en de vmbo-leerlingen in de PISA-steekproef.

Het vmbo

Het vmbo, voluit voorbereidend middelbaar beroepsonderwijs, is in 1999 ontstaan als nieuw schooltype, voortkomend uit een samenvoeging van mavo, vbo en sommige vormen van voortgezet speciaal onderwijs. Deze vernieuwing was erop gericht de aansluiting tussen het voormalig mavo/vbo/vso en het middelbaar beroepsonderwijs (mbo) te verbeteren. In het vmbo kiest een leerling na het tweede leerjaar een leerweg. Er zijn vier leerwegen: de theoretische, de gemengde, de kaderberoepsgerichte en de basisberoepsgerichte.¹ Naast een leerweg kiest een leerling ook een sector waarin hij/zij onderwijs volgt. Er zijn vier sectoren: techniek, zorg en welzijn, economie en landbouw.

De combinatie van sector en leerweg bepaalt welke vakken een leerling in klas 3 en 4 volgt en in welke vakken centraal examen gedaan moet worden. Naast de vier leerwegen is er leerwegondersteunend onderwijs. Dit onderwijs is bestemd voor leerlingen met achterstanden of gedragsproblemen, maar met voldoende capaciteiten om een vmbo-diploma te kunnen halen. Ten slotte is er nog het **praktijkonderwijs**, gericht op een beperkte groep leerlingen voor wie het niet mogelijk is een vmbo-diploma te halen. In tegenstelling tot de leerwegen is dit geen voorbereiding op het mbo, maar rechtstreeks op de arbeidsmarkt.

Het vmbo is dus een zeer gedifferentieerd samengesteld onderwijstype en het herbergt circa 60% van de leerlingen in het voortgezet onderwijs.

1. Bron: <http://www.vmbo.nl/> en <http://www.minocw.nl/vmbo/algemeen.html>

Vmbo-leerlingen

De PISA-opgaven zijn afgenomen bij 15-jarige leerlingen, of meer specifiek leerlingen geboren in 1987. In tabel 1 is te zien hoe in Nederland de 15-jarige leerlingen zijn verdeeld over de leerjaren en schooltypen² (bij elk type opleiding is het aantal leerlingen en het percentage van het totaal aantal leerlingen gegeven):

Tabel 1

opleidingstype	situatie van de leerlingen	aantal leerlingen	percentage
pro	praktijkonderwijs	6.237	3,4
vmbo 2	vmbo-leerlingen van 15 jaar in klas 2	3.501	1,9
vmbo BB	leerlingen op BB-niveau	28.460	15,4
vmbo KB		23.925	13,0
vmbo GL		4.656	2,5
vmbo TL		46.621	25,2
havo 2/3	havo-leerlingen in klas 2 of 3	14.119	7,6
havo 4/5	havo-leerlingen in klas 4 of 5	20.376	11,0
vwo 2/3		10.717	5,8
vwo 4/5		26.085	14,1

Voor PISA is uit deze populatie een gestratificeerde steekproef getrokken³, weergegeven in tabel 2. Hierin zijn 7 leerlingen van wie of het leerjaar of het schooltype onbekend was niet opgenomen.

Tabel 2

klas	pro		vmbo BB/KB		vmbo GL/TL		havo		vwo	
	meisje	jongen	meisje	jongen	meisje	jongen	meisje	jongen	meisje	jongen
1		2	1	1						
2	9	11	43	65	11	11	7	12	2	5
3	21	32	258	375	241	318	149	135	113	116
4	9	10	187	195	312	246	281	207	320	260
5							2		8	9
6										1

2. Bron: Resultaten PISA 2003, *Praktische kennis en vaardigheden van 15-jarigen*. (Cito, 2004), p. 43

3. Voor nadere details over de steekproeftrekking zie: Resultaten PISA 2003, *Praktische kennis en vaardigheden*. (Cito, 2004), bijlage 1. De tabel met steekproefgegevens komt uit dit rapport, p.20

De Nederlandse PISA-steekproef bestaat uit een kleine 4000 leerlingen. Daarvan laten we er in dit onderzoek circa 300 buiten beschouwing: dit zijn leerlingen van het praktijkonderwijs (deze hebben een kortere variant van de toets gemaakt) en leerlingen die niet in leerjaar 3 en 4 zitten.

Voor de nadere analyse van de PISA-resultaten in dit onderzoek hebben we de overblijvende vmbo-leerlingen in vier groepen ingedeeld:

BB/KB 3 – basis- en kaderberoepsgerichte leerweg, leerjaar 3

BB/KB 4 – basis- en kaderberoepsgerichte leerweg, leerjaar 4

GL/TL 3 – gemengde en theoretische leerweg, leerjaar 3

GL/TL 4 – gemengde en theoretische leerweg, leerjaar 4

De havo- en vwo-leerlingen zijn apart gehouden en ook hier is een onderverdeling naar leerjaar gemaakt. Tabel 3 bevat de verdeling over de genoemde categorieën:

Tabel 3

	Opleidings- type	Frequentie	% (incl. mis- sing values)	% (excl. mis- sing values)	Cumulatief %
Valid	BB-KB3	630	15,8	17,0	17,0
	BB-KB4	382	9,6	10,3	27,3
	GL-TL3	559	14,0	15,1	42,3
	GL-TL4	558	14,0	15,0	57,4
	havo3	284	7,1	7,7	65,0
	havo4	488	12,2	13,2	78,2
	vwo3	229	5,7	6,2	84,4
	vwo4	580	14,5	15,6	100,0
	Total	3710	92,9	100,0	
Missing value	overige	282	7,1		
Totaal		3992	100,0		

De 2129 vmbo-leerlingen in de PISA-steekproef kunnen ook langs andere lijnen verdeeld worden, bijvoorbeeld in combinatie met sector, zoals in tabel 4 te zien is.

Het blijkt duidelijk dat de leerwegen niet gelijkmatig over de sectoren zijn verdeeld. We kiezen daarom bij de analyses als eerste ingang de leerweg, of de combinaties van leerwegen. De sectoren kunnen niet onderling vergeleken worden zonder daarbij de ongelijke verdeling over de leerwegen te betrekken. Dit hebben we buiten beschouwing gelaten. Een aparte analyse op combinatie van leerweg en sector levert te kleine groepen om zinvolle analyses te kunnen maken.

Eén punt verdient hier nog speciale aandacht: De 15-jarige vmbo-leerlingen in Nederland hebben niet meer allemaal wiskunde. Wiskunde is alleen een verplicht vak in de sectoren Techniek en Landbouw (groen). Voor leerlingen in de twee andere sectoren, Economie en Zorg en Welzijn, is wiskunde niet verplicht. Wel kunnen leerlingen dit

vak in de vrije keuzeruimte kiezen. Soms stelt een school – vaak om organisatorische redenen – wiskunde toch verplicht voor alle leerlingen in (een van) deze sectoren. Leerlingen in de theoretische leerweg, die vaak een afdeling vormt van een havo/vwo-school, kiezen vaak wiskunde om door te kunnen stromen naar de havo, waarvoor wiskunde verplicht is. Deze keuze kan los staan van de sector. De leerlingen in de theoretische leerweg volgen geen beroepsgericht programma maar hebben een pakket van uitsluitend algemeen vormende vakken.

Tabel 4

Leerweg			sector					Totaal
			Techniek	Landbouw	Economie	Zorg&Wiz	Onbekend	
BB	Klas	3	142	13	70	100	36	361
		4	72	10	45	63	4	194
KB	Klas	3	84	18	84	72	11	269
		4	55	14	57	58	4	188
GL	Klas	3	15	5	24	10	2	56
		4	12	7	16	16	2	53
TL	Klas	3	95	6	195	135	72	503
		4	102	14	199	163	27	505

De classificatie van leerlingen zonder wiskunde is lastig. Dit moet worden afgeleid uit informatie uit de vragenlijst die leerlingen invullen. Er is geen aparte vraag opgenomen over het wel of niet hebben van wiskunde; wel zijn er vragen over het aantal uur wiskunde per week en het aantal medeleerlingen bij wiskunde. De invulling van de vragenlijst op dit punt is niet altijd even eenduidig en betrouwbaar. Er zijn bijvoorbeeld leerlingen die melden in een normale schoolweek geen wiskunde te hebben gehad, maar ze vermelden wel een aantal medeleerlingen dat in de wiskundeklas zit. Als een leerling op beide vragen geen antwoord heeft gegeven, maar wel in de sector techniek of landbouw zit, waar wiskunde een verplicht vak is, dan is de leerling in tabel 5 toch meegeteld als een leerling die wiskunde heeft. Deze classificatie is daarom verre van zeker, en geeft slechts een globale indicatie. Ongeveer 10% van de vmbo-leerlingen uit de steekproef lijkt geen wiskunde meer te hebben.

Tabel 5

Wiskunde	sector					Totaal
	Techniek	Landbouw	Economie	Zorg&Wiz	Onbekend	
ja	577	87	629	503	140	1936
nee	0	0	61	114	18	193
Totaal	577	87	690	617	158	2129

De verdeling van deze leerlingen over de leerwegen binnen de sectoren is zeer divers en is vanwege het geringe aantal niet verder in kaart gebracht. Er wordt in de analyses ook geen rekening mee gehouden.

Een ander punt van aandacht in dit verband zijn de leerlingen met LWOO-indicatie (zie tabel 6). Een leerling krijgt deze indicatie – en komt daarmee in aanmerking voor leerwegondersteunend onderwijs – als er op een aantal kenmerken problemen of achterstanden zijn.⁴ Deze leerlingen vinden we vooral terug in het derde leerjaar BB. In de groep BB3 heeft circa een derde deel van de leerlingen deze indicatie. Bij interpretatie van de scores dient hiermee rekening gehouden te worden.

Tabel 6

Leerweg			LWOO		Totaal
			nee	ja	
BB	Klas	3	248	113	361
		4	172	22	194
KB	Klas	3	263	6	269
		4	187	1	188
GL	Klas	3	54	2	56
		4	51	2	53
TL	Klas	3	500	3	503
		4	505	0	505

De meeste Nederlandse 15-jarige vmbo-leerlingen zitten onvertraagd in leerjaar 3 of 4, ze hebben dan dus 9 of 10 jaar onderwijs gevolgd. In de beroepsgerichte leerwegen zit ruim 30% in klas 4 en een kleine 20% onvertraagd in klas 3. In totaal is ongeveer 50% vertraagd, meestal zitten deze leerlingen in klas 3. Met vertraging is in de analyse van de PISA-resultaten geen rekening gehouden.

Prestaties vmbo-leerlingen

De set wiskundeopgaven van PISA bestaat uit 84 opgaven. Deze opgavenset moet zodanig zijn samengesteld dat ook de 'zwakste' leerlingen uit de deelnemende landen een deel ervan goed kunnen maken. Anderzijds moeten er ook opgaven zijn die interessant en uitdagend zijn voor de beste leerlingen. Voldoen aan deze twee uitersten is lastig. Voor de betere leerlingen in Nederland (met name de vwo-leerlingen) is een groot deel van de opgaven te makkelijk. Hierdoor is het minder goed mogelijk om op grond van de PISA-resultaten een gedifferentieerd beeld te vormen van de wiskundige geletterdheid van deze betere leerlingen. Eigenlijk zijn er te weinig uitdagende opgaven waarmee de beste van de betere leerlingen te onderscheiden zijn. De opgavenset van PISA lijkt voor de Nederlandse leerlingen beter toegesneden op de

4. Het gaat dan om: IQ, leerachterstanden en sociaal emotionele ontwikkeling.

zwak tot gemiddeld scorende leerlingen. Globaal gezegd is dit het hele vmbo. De wiskundige geletterdheid in de PISA-opgaven past beter bij wat we vragen van de vmbo-leerlingen. Voor deze leerlingengroep ontstaat een goed gedifferentieerd beeld van hun mogelijkheden op het gebied van wiskundige geletterdheid.

Algemene resultaten

Tabel 7 laat zien hoe de leerlingen in de verschillende leerjaren en opleidingstypen scoren op wiskunde. Het PISA-gemiddelde is op 500 gezet, de eenvoudigste opgave op het gebied van wiskundige geletterdheid heeft een drempelwaarde van 262, de moeilijkste opgave van 801. De drempelwaarde is een indicatie voor de moeilijkheid van een opgave.⁵

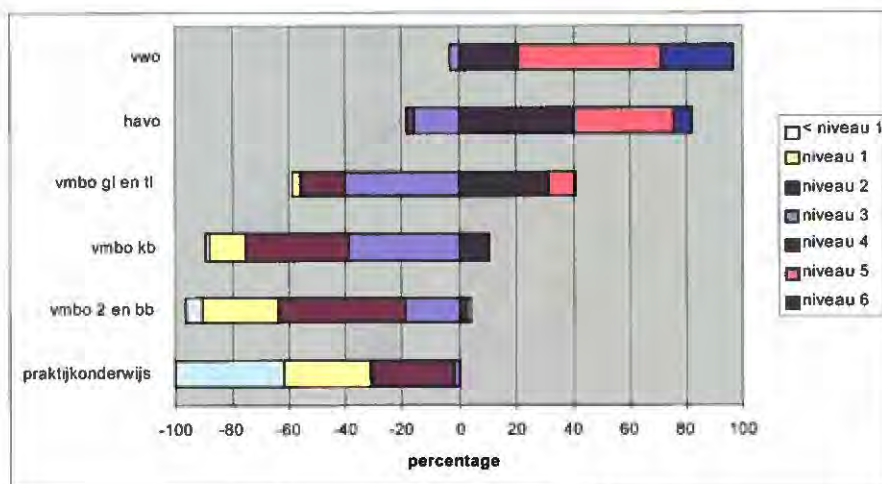
Tabel 7

opleidingstype	domeinen
	<i>wiskunde</i>
pro	384
vmbo2 en BB	444
vmbo KB	481
vmbo GL/TL	532
havo	594
vwo	638
Nederland totaal	538

De leerlingen van het praktijkonderwijs en de beroepsgerichte leerwegen BB en KB scoren onder het PISA-gemiddelde van 500, het gaat om circa een derde deel van de Nederlandse leerlingen. De GL/TL-leerlingen scoren rond het landelijk gemiddelde. Op zich is dit scorepatroon niet onverwacht: naarmate het schooltype 'hoger' is, is de gemiddelde wiskundescore op PISA dat ook. Het Nederlandse onderwijsstelsel is een selectief stelsel, zoals ook in vorige hoofdstukken is vermeld.

De PISA-opgaven worden in zes zogenaamde vaardigheidsniveaus geplaatst. Als we de zes vaardigheidsniveaus van PISA erbij betrekken zien we het volgende beeld (zie fig. 1).

5. Voor nadere uitleg over de drempelwaarden en de vaardigheidsniveaus van PISA verwijzen we naar hoofdstuk 1 en naar het rapport *Resultaten PISA 2003, Praktische kennis en vaardigheden*. (Cito, 2004).



figuur 1 Niveauverdeling wiskunde per opleidingstype

Hoewel de OESO het criterium legt bij scores boven niveau 1, wordt in het Nederlandse PISA-rapport⁶ terecht opgemerkt dat we in een samenleving als de onze, die zich wil profileren als kenniseconomie, hogere eisen moeten stellen aan het algemene niveau van wiskundige geletterdheid van de burgers. Eigenlijk is niveau 4 daarbij het minimumniveau. Op dat niveau kan iemand zelfstandig met vragen en problemen van kwantitatieve aard omgaan, en daarbij onder meer keuzes maken uit geschikte representaties en modellen, veronderstellingen doen, eigen interpretaties en redeneringen gebruiken om uitleg te construeren en te communiceren. Als we deze 'eis' leggen naast het beeld dat uit figuur 1 naar voren komt, zien we dat van de BB- en KB-leerlingen minder dan 10% dit niveau haalt en van de GL/TL-leerlingen maar ongeveer 40%. Dit is een zorgelijke situatie.

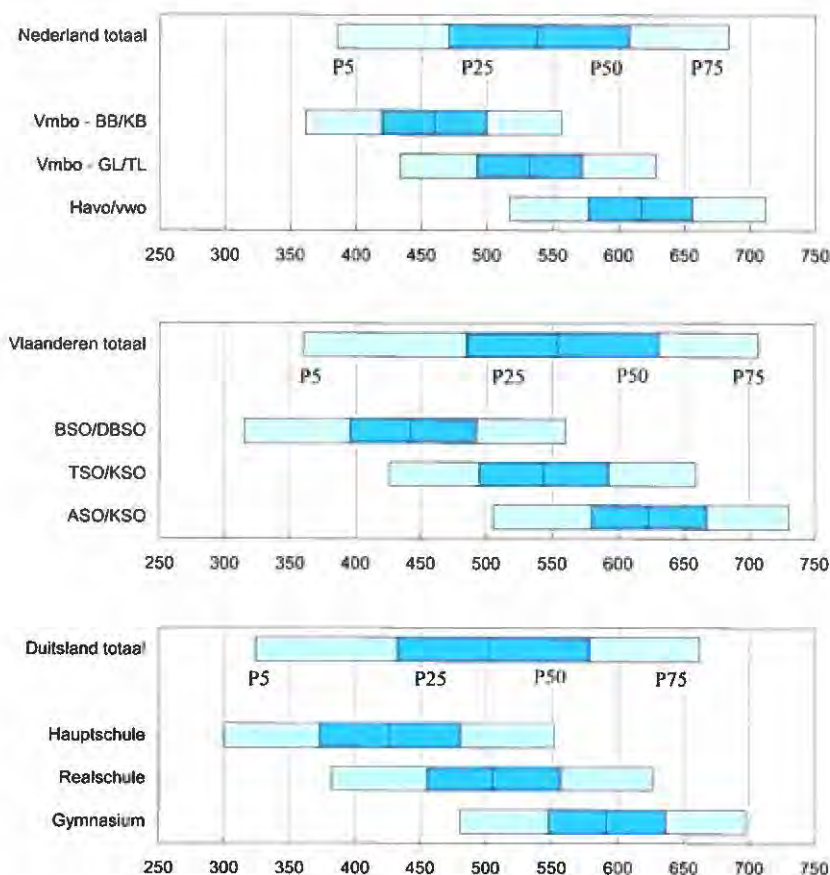
Resultaten van vmbo-leerlingen internationaal vergeleken

PISA is een internationaal onderzoek, het is mogelijk om ook voor specifieke leerlingpopulaties de resultaten te vergelijken. We hebben in figuur 2 de prestaties van de Nederlandse vmbo-leerlingen vergeleken met die van vergelijkbare leerlingengroepen uit onze buurlanden Vlaanderen en Duitsland.⁷

Opvallend is het verschil in spreiding van de resultaten; deze is in Nederland het kleinst voor alle leerlingengroepen. In Nederland doen de zwakste leerlingen het vergeleken met Vlaanderen en Duitsland relatief goed. We merken nog op dat het speciaal onderwijs in alle landen buiten beschouwing is gelaten. De hoge gemiddelde score van Vlaanderen, wordt met name gehaald door de prestaties van de betere leerlingen. Zowel de betere leerlingen van de middengroep (vergelijkbaar met onze vmbo-GT/TL-leerlingen) als die van de beste groep (ons havo/vwo) doen het in Vlaanderen beter dan in Nederland en Duitsland.

6. Pagina 32-33.

7. Zie hoofdstuk 1 voor informatie over de schooltypen in Duitsland en Vlaanderen en de leerlingenverdeling



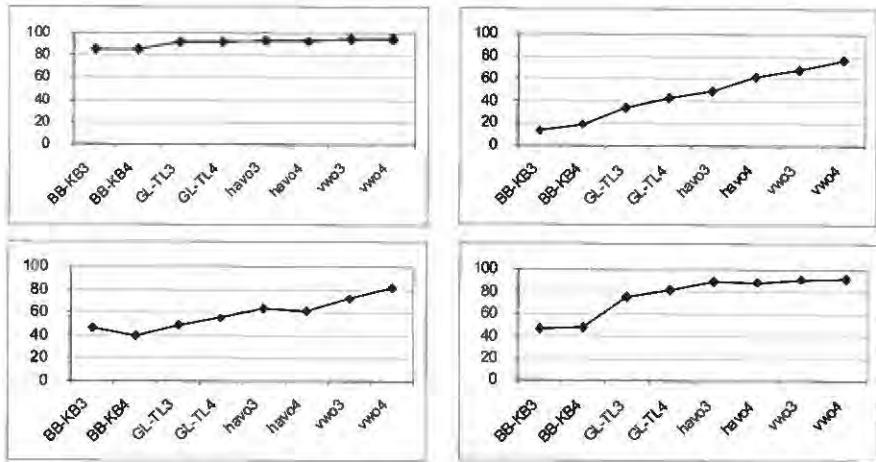
figuur 2 Resultaten per schooltype in verschillende landen

Onderscheidende items voor vmbo-leerlingen

Net als in hoofdstuk 2 bekijken we de best en slechtst gemaakte opgaven. Als we de rangorde in opgaven voor Nederland splitsen in drie verschillende rangordes per opleidingstype, te weten vmbo BB/KB, vmbo GL/TL en havo/vwo, kunnen we de prestaties van de vmbo-leerlingen wat beter in beeld krijgen.

We bespreken een aantal opgaven waarop de vmbo-leerlingen opvallend scoren. We hebben om deze opgaven te vinden zowel gekeken naar de rangordes van de opgaven voor de verschillende groepen leerlingen als naar de scorepatronen van de opgaven.

De bovenste twee scorepatronen in figuur 3 laten een te verwachten beeld zien: naarmate het schooltype en het leerjaar hoger is, is het aantal leerlingen dat een vraag correct beantwoordt hoger. De grafieken zijn alleen bedoeld om een globaal



figuur 3 Percentages goedscores voor de verschillende groepen leerlingen

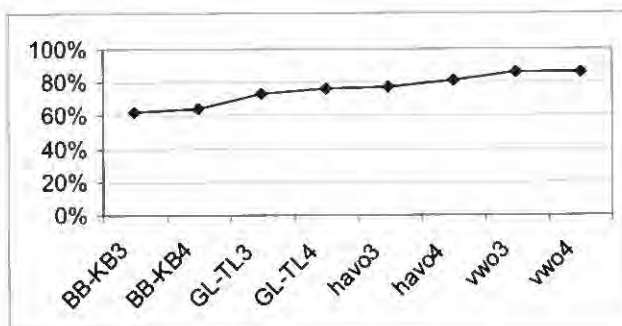
beeld te geven van de trend van BB naar vwo. Enige voorzichtigheid bij de interpretatie van deze grafieken is op zijn plaats. Het is namelijk niet te bepalen of havo 3 een 'hogere' schooltype is dan vmbo GL-TL-4. Waar we meer in detail naar resultaten kijken vermelden we de gescoorde percentages erbij.

De derde grafiek in figuur 3 laat een wat onregelmatig patroon zien, met name bij dergelijke patronen is het bekijken van de onderliggende percentages van belang. De laatste grafiek is een voorbeeld waarin opvalt dat de BB/KB-leerlingen duidelijk lager scoren dan verwacht. Op grond van deze grafiek zou je bijvoorbeeld kunnen vermoeden dat het onderwerp van deze opgave niet voorkomt in het curriculum van deze leerlingen. Dit scorepatroon vraagt in ieder geval om een nadere analyse. We laten hieronder een aantal voorbeelden zien van opgaven waarop vmbo-leerlingen op de een of andere manier opvallend scoren en die we daarom nader hebben geanalyseerd.

Voorbeeld 1: opgave over afstand op een kaart schatten

Het eerste voorbeeld van een opgave waarin de vmbo-leerlingen opvallend scoren is een opgave waarin de afstand op een kaart moet worden geschat met behulp van een schaallijn. Deze opgave komt uit het cluster Verbindingen en zit in vaardigheidsniveau 3. De scores van de verschillende leerlingengroepen op deze opgave liggen redelijk dicht bij elkaar zoals te zien is in de grafiek in figuur 4. Van de leerlingen uit vmbo BB/KB-3 maakt ruim 60% de opgave goed, voor vwo 4 is dat bijna 90%.

Bij de vmbo-BB/KB-leerlingen staat deze opgave op de 12e plaats in de rangorde van de *p*-waarden van goed naar slecht gemaakte opgaven. In de rangorde voor alle Nederlandse leerlingen staat deze 23e, voor vmbo-GLT/TL-leerlingen staat deze opgave op de 22e en voor havo/vwo-leerlingen pas op de 38e plaats. Dat is lager dan de 28e plaats die deze opgave heeft als we kijken naar de drempelwaarde (498) die valt binnen vaardigheidsniveau 3. De vmbo BB/KB-leerlingen maken deze opgave dus relatief goed. Misschien geloven de havo/vwo-leerlingen niet dat de opgave zo makkelijk

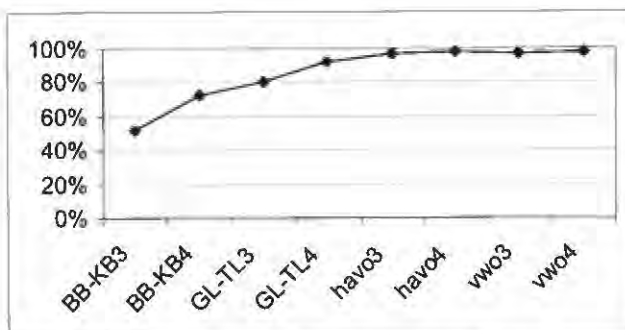


figuur 4 Scoreverloop voorbeeld 1

is als deze voor hen lijkt en zoeken ze er meer achter, waardoor ze fouten maken. Voor de vmbo-ers, met name uit de beroepsgerichte leerwegen, is de opgave waarschijnlijk herkenbaar. We kunnen niet met zekerheid vaststellen of er een relatie is met de sector, want slechts 35% van de leerlingen uit vmbo BB/KB die met dit PISA-onderzoek meedeelde zat in de sector Techniek, waar werken met schaal veel voorkomt. Hoewel deze leerlingen het grootste aandeel vormen, is het aandeel leerlingen uit de sectoren Economie en Zorg en Welzijn, waar wiskunde geen verplicht vak is, samen ruim 50%.

Voorbeeld 2: opgave over uitleggen hoe het gemiddelde wordt berekend

Het omgekeerde – vmbo BB/KB-leerlingen presteren duidelijk minder dan verwacht – is bijvoorbeeld het geval bij een Reproductie-opgave waarin moet worden uitgelegd hoe een gemiddelde is bepaald. Het betreft ook hier een opgave uit vaardigheidsniveau 3. Deze opgave staat in de vmbo BB/KB *p*-waarde rangorde op de 21e plaats – wat redelijk overeenkomt met de 24e plaats in de volgorde als we alleen naar de drempelwaarden kijken. In de algehele Nederlandse *p*-waarde ranglijst staat deze opgave op de 12e plaats. Dit is ook zo voor GL/TL, voor havo/vwo staat deze opgave zelfs op 2.



figuur 5 Scoreverloop voorbeeld 2

Zoals in de grafiek (zie fig. 5) van de scores is te zien, heeft van de leerlingen uit klas 3 vmbo BB en KB slechts 52% deze opgave goed. Voor de 4e klas leerlingen BB

en KB is dat 72% en voor de GL/TL-leerlingen ligt de score tussen 80 en 90%. Wellicht hebben leerlingen in de derde klas BB/KB nog weinig statistiek gehad of vinden ze het type opdracht waarin om een formulering wordt gevraagd moeilijk.

Voorbeeld 3: opgaven over werken met een letterformule

Een derde voorbeeld van een opgave waarbij de verschillen in rangorde groot zijn, is de opgave De beste auto, waarin leerlingen met een letterformule moeten werken (zie fig. 6).

DE BESTE AUTO

Een autoblad gebruikt een beoordelingssysteem voor nieuwe auto's en kent de prijs van 'Auto van het Jaar' toe aan de auto met de hoogste totaalscore. Er zijn vijf nieuwe auto's getest. In de onderstaande tabel staan de toegekende waarderingscijfers.

Auto	Veiligheid (V)	Brandstof-verbruik (B)	Ontwerp (O)	Interieur (I)
Ca	3	1	2	3
M2	2	2	2	2
Sp	3	1	3	2
N1	1	3	3	3
KK	3	2	3	2

De waarderingscijfers hebben de volgende betekenis:

3 punten = Uitstekend

2 punten = Goed

1 punt = Redelijk

Vraag 1: DE BESTE AUTO

Om de totaalscore van een auto te berekenen past het autoblad de volgende formule toe, die een gewogen som is van de afzonderlijke waarderingscijfers:

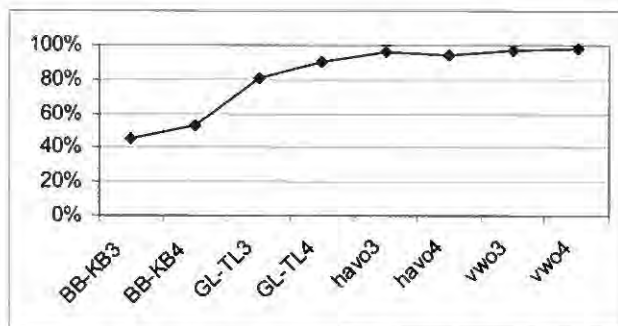
$$\text{Totaalscore} = (3 \times V) + B + O + I$$

Bereken de totaalscore voor auto 'Ca'. Noteer je antwoord hieronder.

Totaalscore voor 'Ca':

figuur 6 Opgave: De beste auto, deelvraag 1

Deze opgave uit het Reproductiecluster (vaardigheidsniveau 2) staat in de Nederlandse rangorde naar *p*-waarde op de 19e plaats; dit is lager dan de 12e plaats die gebaseerd is op de drempelwaarden. Ter vergelijking: in Duitsland, Vlaanderen en de USA staat deze opgave hoger (respectievelijk 11e, 13e, 7e), in Noorwegen (33e) en Zweden (37e) veel lager. De lage plaats voor Nederland hangt samen met de lage plaats die deze opgave in de vmbo BB/KB rangorde heeft, namelijk de 29e. Voor de GL/TL-leerlingen staat deze opgave op de 10e plaats en voor havo/vwo zelfs op de derde.



figuur 7 Scores op De beste auto vraag 1

In de grafiek in figuur 7 is duidelijk te zien dat de scores voor de leerlingen uit BB/KB veel lager zijn dan die van de andere leerlingen. De score van de vmbo-leerlingen uit de beroepsgerichte leerwegen ligt rond de 50%, terwijl de leerlingen uit de gemengde en theoretische leerweg 80 à 90% goed scoren. Dit is in vergelijking met andere opdrachten op dit vaardigheidsniveau een groot verschil.

Vmbo-leerlingen uit met name de basisberoepsgerichte leerweg scoren in het algemeen op opgaven met (letter)formules aanzienlijk slechter dan volgens het gemiddelde scorepatroon te verwachten zou zijn, vooral als die formules gepresenteerd worden in combinatie met een wat abstracte of kale en weinig betekenisvolle context. Dit kan natuurlijk te maken hebben met het feit dat de leerlingen in de basisberoepsgerichte leerweg (BB) alleen met woordformules hoeven te werken. Het zelf opstellen van eenvoudige (lineaire) formules lukt de meeste vmbo-leerlingen helemaal niet. Dit onderwerp krijgt ook weinig of geen aandacht in het curriculum van de vmbo-leerlingen.

Bij de tweede deelvraag van de opgave De Beste Auto, een vraag uit het cluster Reflectie in vaardigheidsniveau 5, moeten de leerlingen zelf een formule maken (zie fig. 8).

Vraag 2: DE BESTE AUTO

De fabrikant van de auto 'Ca' vond de formule voor de totaalscore oneerlijk.

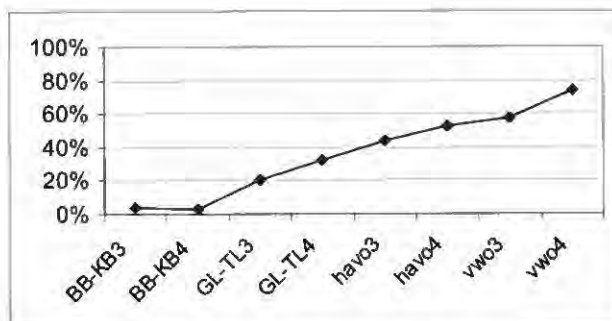
Geef een formule om de totaalscore zo te berekenen dat auto 'Ca' er als winnaar uitkomt.

Jouw formule moet elk van de vier variabelen bevatten. Geef jouw formule door positieve getallen in te vullen op de vier stippelijntjes in de onderstaande formule.

Totaalscore = V + B + O + I

figuur 8 Opgave: De beste auto, deelvraag 2

Deze vraag wordt door minder dan 5% van de leerlingen uit de beroepsgerichte leerwegen goed gedaan, bij de andere vmbo-leerlingen ligt de goedscore rond de 25%.



figuur 9 Scores op De beste auto, deelvraag 2

In de vmbo BB/KB-ranglijst naar *p*-waarde hoort deze tweede deelvraag bij de 5 slechtst gemaakte vragen. Ook in de totale ranglijst voor Nederland staat deze opgave laag, namelijk op de 76e plaats van de 84. Dit komt overeen met de plaats in de rangorde van alle opgaven wanneer deze naar 'moeilijkheid', aangegeven door de zogenaamde drempelwaarde, gesorteerd worden

Voorbeeld 4: opgaven omtrent het gemiddelde

Een ander type opgave uit het Reproductiecluster waarop met name de leerlingen uit de beroepsgerichte leerwegen lager dan verwacht scoren, zijn opgaven waarin iets met een gemiddelde moet worden gedaan. Soms betreft dit de berekening van een gemiddelde, maar ook op opgaven waarin met een gegeven gemiddelde moet worden gerekend of geredeneerd, of waarin moet worden uitgelegd hoe een gemiddelde valt te berekenen, scoren de BB- en KB-leerlingen slecht.

Een voorbeeld van een dergelijke opgave is te zien in de figuur 10:

NATUURKUNDEPROEFWERKEN

Vraag 1: NATUURKUNDEPROEFWERKEN

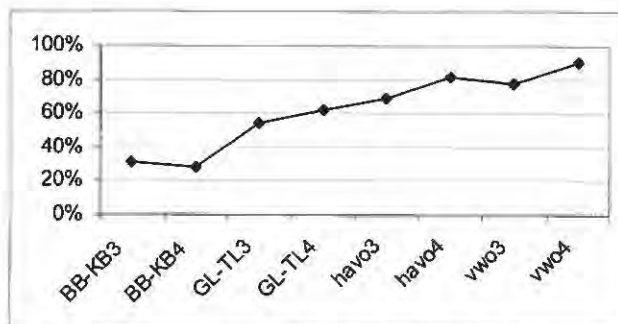
Op de school van Mei-Lin geeft de natuurkundeleraar proefwerken waarvoor een maximale puntenscore kan worden gehaald van 100. Mei-Lin haalde voor haar eerste vier natuurkundeproefwerken een gemiddelde van 60 punten. Voor het vijfde proefwerk kreeg ze 80 punten.

Wat is het gemiddelde resultaat van Mei-Lin voor natuurkunde na de vijf proefwerken?

Gemiddelde:

figuur 10 Opgave: Natuurkundeproefwerken

Deze opgave Natuurkundeproefwerken (Reproductie, vaardigheidsniveau 4) maakt ongeveer 30% van de leerlingen uit de beroepsgerichte leerwegen goed. Voor de gemengde en theoretische leerweg is dat zo'n 60%. Voor havo en vwo leerlingen ligt dat percentage boven de 80%.



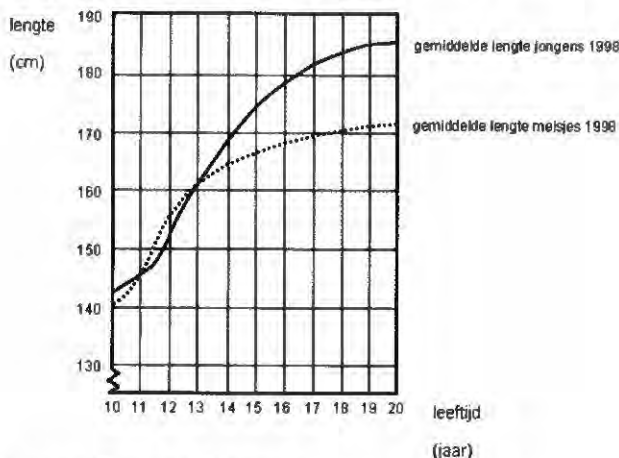
figuur 11 Scores op de Natuurkundeproefwerken

De fout die vmbo-leerlingen het meest maken, is dat zij als antwoord het niet-gewogen gemiddelde van 60 en 80 geven, dus 70. Ruim 20% van de vmbo-leerlingen maakt deze fout. We kunnen helaas niet aangeven uit welke leerweg deze leerlingen komen.

Een andere opgave waarin het gemiddelde een rol speelt en waarop de BB/KB-leerlingen erg laag scoren is de eerste deelvraag van de opgave Jongeren steeds langer (zie fig. 12).

JONGEREN STEEDS LANGER

In deze grafiek wordt de gemiddelde lengte van zowel jongens als meisjes in Nederland in 1998 weergegeven.

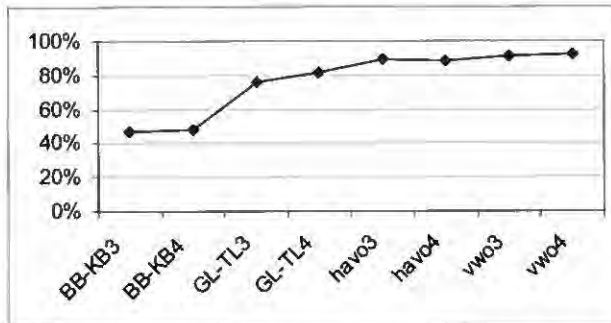


Vraag 1: JONGEREN STEEDS LANGER

Sinds 1980 is de gemiddelde lengte van 20-jarige meisjes met 2,3 cm toegenomen tot 170,6 cm. Wat was de gemiddelde lengte van 20-jarige meisjes in 1980?

figuur 12 Opgave: Jongeren steeds langer

Op deze vraag geeft minder dan 50% van de vmbo BB/KB-leerlingen het juiste antwoord, terwijl dit voor de andere groepen leerlingen boven de 75% ligt.



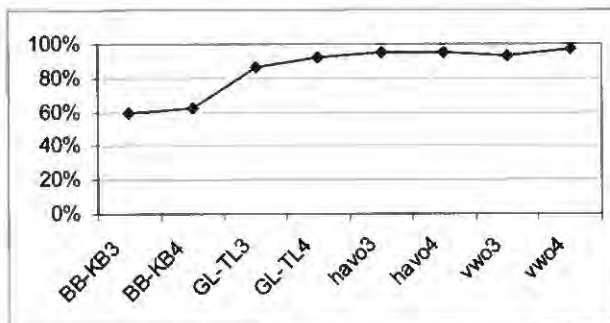
figuur 13 : Scores op Jongeren steeds langer

Dit kan te maken hebben met het begrip 'gemiddelde'. Het kan ook zijn dat leerlingen uit BB/KB de opgave slecht maken doordat ze moeite hebben met het lezen van de tekst en de correcte vertaling ervan naar de wiskundig gezien eenvoudige opdracht: $170,6 - 2,3$. Een groot deel van de leerlingen heeft $170,6 + 2,3$ gedaan, mogelijk vanwege het woord 'toegenomen', dat valt te associëren met 'optellen'.

Uit het feit dat op bijna alle opgaven uit het Reproductiecluster met het begrip 'gemiddelde' de BB/KB-leerlingen relatief slechter scoren, trekken we voorzichtig de conclusie dat deze leerlingen het begrip 'gemiddelde' – en mogelijk ook het woord – op een basaal niveau niet begrijpen.

Voorbeeld 5: opgave over rekenen in een complexe situatie

Dat BB/KB-leerlingen moeite hebben met rekenen in complexe situaties blijkt uit de lage resultaten op de vragen bij de opgave Wisselkoers, die ook in hoofdstuk 2 is besproken.



figuur 14 Scores op opgave 2 van Wisselkoers

Bij de eerste twee vragen van deze opgave die beide in het reproductiecluster zitten, scoren de BB/KB-leerlingen aanzienlijk lager dan de overige leerlingen: zo'n 60 à 70%,

tegen 90% voor de andere groepen (zie de scores op vraag 2 in fig. 14). Ook in de rangorden naar *p*-waarden vallen deze vragen op: voor vmbo BB/KB staan de twee vragen op plaats 11 en 17, terwijl ze in de rangordes voor vmbo GL/TL en havo/vwo rond de 6e plaats staan en in de ordening naar drempelwaarde op plaats 3 en 11.

Leerlingenwerk

De scores op clusterniveau (Reproductie, Verbindingen en Reflectie), op domeinniveau (Veranderingen en Relaties, Onzekerheid, Hoeveelheid en Vorm en Ruimte)⁸ en op de individuele opgaven⁹, geven een globaal zicht op de prestaties van de onderscheiden groepen vmbo-leerlingen. Om meer in detail iets over de wiskundige gecijferdheid van vmbo-leerlingen te kunnen zeggen, is leerlingenwerk bekeken. Steekproefsgewijs is voor een aantal opgaven leerlingenwerk bekeken en geanalyseerd. We beschrijven de bevindingen gegroepeerd onder een aantal kopjes en laten voorbeelden van leerlingenwerk zien. We leggen daarbij de nadruk op het werk van de vmbo-leerlingen, maar zullen in een aantal gevallen ook iets opmerken over de andere leerlingen. Ook in het vorige hoofdstuk zijn voorbeelden uit leerlingenwerk te vinden.

Wat schrijven de leerlingen op?

Nederlandse leerlingen slaan in vergelijking met leerlingen uit andere landen niet snel opgaven over. Ze schrijven bij open vragen vaak iets op. Dat is niet altijd een echt antwoord op de vraag. Soms is het een opmerking als:

Nou ik weet het niet want ik weet helemaal niet hoe ik dit uit moet rekenen.

Dergelijke antwoorden, die de score van een onjuist antwoord meekrijgen en niet de aanduiding van een open gelaten vraag, vertekenen het beeld van de leerlingen die niets overslaan natuurlijk wel een beetje.

Vmbo-leerlingen lijken iets meer over te slaan dan havo/vwo-leerlingen en ze slaan vaker grote opgaven met veel tekst over. Het lijkt overigens niet zo te zijn dat wanneer ze één van de deelvragen niet hebben beantwoord ze de volgende ook overslaan. Nederlandse leerlingen zijn wel gewend aan opgaven met deelvragen.

Verder valt op dat de Nederlandse leerlingen in alle schooltypen snel tevreden lijken te zijn over hun eerste antwoord. Ook wanneer we vermoeden dat een opgave binnen de capaciteiten van de leerling valt, zien we soms toch dat een leerling na een wat vluchtige eerste aanzet, niet verder gaat. Het lijkt erop dat veel leerlingen er vanuit gaan dat als het niet meteen lukt, of als de eerste stap niet meteen een eindantwoord geeft, het ook niet meer zal gaan lukken. We zien ook nauwelijks kladwerk of tussenstappen in een berekening en al helemaal geen schetsjes of tekeningen, ook niet als de opgave daartoe duidelijk aanleiding geeft (zie voorbeeld in hoofdstuk 3).

Met name de vmbo-leerlingen geven soms slechts een deel van het antwoord, bijvoorbeeld alleen het getal 11 in plaats van het complete interval van 11 – 13; één voorbeeld in plaats van alle mogelijkheden. Dit kan te maken hebben met taalproble-

8. Zie hiervoor het Nationale rapport: *Resultaten PISA 2003, Praktische kennis en vaardigheden*. (Cito, 2004).

9. Zie hiervoor de publicatie met de vrijgegeven opgaven.

men bij het lezen van de vraag, het kan zijn dat de leerlingen snel tevreden zijn met het feit dat ze een antwoord hebben gevonden, het kan ook een onjuiste veronderstelling zijn dat er bij wiskundeopgaven maar één goed antwoord is.

Ook blijkt uit analyses van leerlingenwerk dat met name de vmbo-leerlingen vaak maar wat doen met de getallen of gegevens uit een opgave. Ze schrijven gegevens over, voeren er een willekeurige bewerking mee uit, maar lijken geen doelgericht oplossingsproces uit te voeren. Vmbo-leerlingen geven mede daardoor meer verschillende antwoorden dan havo/vwo-leerlingen. We illustreren dit aan de hand van de opgave in figuur 15, die ook elders in deze publicatie is opgenomen.

NATUURKUNDEPROEFWERKEN

Vraag 1: NATUURKUNDEPROEFWERKEN

Op de school van Mei-Lin geeft de natuurkundeleraar proefwerken waarvoor een maximale puntenscore kan worden gehaald van 100. Mei-Lin haalde voor haar eerste vier natuurkundeproefwerken een gemiddelde van 60 punten. Voor het vijfde proefwerk kreeg ze 80 punten.

Wat is het gemiddelde resultaat van Mei-Lin voor natuurkunde na de vijf proefwerken?

Gemiddelde:

figuur 15 Opgave: Natuurkundeproefwerken

In onderstaande tabellen 8 en 9 staan de meest voorkomende antwoorden op de opgave uit figuur 15 met de frequentie. Het goede antwoord is vetgedrukt.

Tabel 8 vmbo

Antwoord	6,4	7	7,5	28	64	65	70	80			totaal
Frequentie	23	17	4	24	307	9	148	14			709
%	3	2	1	3	43	1	21	2			76%

Tabel 9 havo/vwo

Antwoord	6,4	7	7,5	28	64	65	70	80			totaal
Frequentie	2			7	211	1	14				255
%	1			3	83		5				92%

Het goede antwoord 64 is zowel bij vmbo als bij havo/vwo het relatief meest voorkomende antwoord: bij vmbo scoort 43% het juiste antwoord en bij havo/vwo doet maar liefst 83% het goed. In de tabellen 4.8 en 4.9 valt te zien dat vmbo-leerlingen meer verschillende antwoorden geven. Een aantal van de alternatieve antwoorden is goed

verklaarbaar. Zo is het veel voorkomende antwoord 70 naar alle waarschijnlijkheid het resultaat van het niet-gewogen middelen van de waarden 60 en 80.

Het antwoord 28 dat bij beide groepen leerlingen even vaak voorkomt, is waarschijnlijk de uitkomst van de berekening $\frac{60 + 80}{5}$. Deze berekening past bij de volgende redenering: er zijn in totaal vijf proefwerken geweest met slechts twee verschillende scores, te weten 60 en 80, die scores tel je op en deel je door het aantal proefwerken, en dat is vijf, zo krijg je als antwoord 28.

Voor het antwoord 6,4 komt wellicht als verklaring in aanmerking dat leerlingen te ver doorgegaan zijn met rekenwerk. Na de berekening van de op zich correcte waarde 64 heeft een leerling wellicht abusievelijk bedacht dat deze waarde nog moest worden omgezet in een cijfer tussen 0 en 10 (de in Nederland gebruikelijke cijfers bij proefwerken). Zo valt ook het antwoord '7' te verklaren.

De andere antwoorden laten zich moeilijk herleiden tot resultaten van in de context wiskundig zinvolle bewerkingen. Het vermoeden is dat de leerlingen die deze overige antwoorden geven zomaar iets gedaan of gegokt hebben.

Leerlingen kijken vaak ook niet of het antwoord wel mogelijk is binnen de gegeven context. Ze vertrouwen op de juistheid van hun berekeningen en accepteren antwoorden als 1,59 cm of zelfs van 0,63 cm voor de hoogte van een trap en 2 meter en 2 cm als paslengte. Hoewel het mogelijk is dat je je eigen kleine foutjes over het hoofd ziet, zouden dergelijke antwoorden toch tot nadenken moeten stemmen.

Aan de andere kant zien we dat de Nederlandse leerlingen in kort antwoord vragen soms wel een korte uitleg geven, ook als daar niet om wordt gevraagd. Ze hebben op school geleerd dit te doen. Zo'n korte uitleg, bijvoorbeeld 'het is 9 uur minder dan in Sydney, dus 10 uur 's ochtends', vormt dan een onderdeel van het antwoord terwijl een schetsje of tussenberekening alleen voor de leerling zelf zou zijn – een middel om het eigen werk te organiseren. Dit laatste doen leerlingen, met name in het vmbo, weinig.

Passief taalgebruik: lezen van de opgave

Wanneer leerlingen vragen voorgelegd krijgen binnen een min of meer realistische situatie is de eerste eis het kunnen vertalen of mathematiseren van de vraag. Dat is voor sommige leerlingen een probleem wanneer hun kennis van de Nederlandse taal om wat voor reden dan ook onvoldoende is. Wanneer je niet weet wat tegenover elkaar liggende zijvlakken zijn, kun je de vraag die daarover gesteld wordt niet goed begrijpen.

Een ander voorbeeld waarbij een beperkte taalvaardigheid mogelijk een rol speelt zijn vragen waarbij leerlingen het goede antwoord moeten omcirkelen. Er is een aantal vmbo-leerlingen dat consequent, door hun hele toetsboekje heen, de verkeerde mogelijkheid kiest, alsof ze omcirkelen verwarren met doorstrepen.

Actief taalgebruik: open vragen beantwoorden

Leerlingen, en vooral vmbo-leerlingen, hebben moeite om een antwoord te formuleren. Vooral wanneer ze verklaringen of argumenten moeten geven, blijken ze op grote taalproblemen te stuiten. Er worden veel slordig geformuleerde antwoorden gegeven, wat het moeilijk maakt om te beoordelen of de leerling de vraag goed beantwoordde. Waarschijnlijk herkennen docenten dit wel uit het werk van hun eigen leerlingen. Een welwillende docent die zijn leerlingen kent, zal een slecht geformuleerd antwoord wellicht nog accepteren; voor degenen die de PISA-antwoorden hebben gescoord ligt dat lastiger. Uit de toetsboekjes bleek ook dat deze antwoorden moeilijk te scoren waren. We geven hier een aantal voorbeelden van dergelijke slordig geformuleerde en niet eenvoudig te beoordelen antwoorden. Het zijn antwoorden op de volgende vraag uit de opgave Jongeren steeds langer:

'Leg uit hoe je aan de grafiek kunt zien dat het gemiddelde groeitempo van meisjes langzamer wordt na hun 12de jaar.'

Een vmbo-leerling geeft het volgende antwoord:

Je ziet dat ze niet meer omhooggaan maar een beetje dezelfde lengte.

Waarschijnlijk heeft deze leerling het wel begrepen en bedoelt hij/zij te zeggen dat de grafiek niet meer omhoog gaat (stijgt), maar zo'n beetje op dezelfde lengte (hoogte) blijft lopen.

Een nog onduidelijker formulering is deze:

Omdat het bol begint te staan, de lijn, ze gaan eerst recht omhoog en dan trekt ie bol.

Het zou kunnen dat de nadruk op zelfstandig werken in de les en zelf antwoorden nakijken, wat in het Nederlands onderwijs op grote schaal voorkomt, tot gevolg heeft dat leerlingen vaak weinig werk van anderen zien en veelal niet weten welke redeneringen helemaal goed zijn, welke bijna goed en welke eisen aan de formulering worden gesteld.

Uit onderzoek (Van den Boer, 2003) is gebleken dat met name allochtone leerlingen, die relatief vaak in het vmbo zitten, meer problemen hebben met taal bij wiskunde dan vaak wordt aangenomen. Deze leerlingen blijken meestal erg vaardig in het verhullen van hun taalproblemen, waardoor ze al snel buiten het zicht van de docent blijven.

Nagenoeg alle Nederlandse leerlingen die bovenstaande vraag goed beantwoordden, hebben dit gedaan in alledaagse taal. Die dagelijkse taal wordt gebruikt wanneer ze refereren aan de situatie zelf, maar ook wanneer ze bijvoorbeeld over de grafiek schrijven:

Dan gaat de streep niet meer zo snel omhoog.

in plaats van:

Dan stijgt de grafiek niet meer zo snel.

Er zijn maar weinig Nederlandse leerlingen die opschrijven dat de grafiek minder steil wordt en al helemaal niet dat de hellingshoek kleiner wordt of dat de snelheid van verandering vermindert. Wiskundige termen als 'hellingshoek' of 'afnemende stijging' komen nagenoeg niet voor. De dubbele codering van de antwoorden maakt in dit geval onderscheid tussen correcte antwoorden in dagelijkse taal (code 11) en correcte

antwoorden in wiskundige taal (code 12). Het vergelijken van de dubbel gecodeerde antwoorden laat zien dat havo/vwo-leerlingen iets vaker wiskundige termen in hun antwoord gebruiken dan vmbo-leerlingen, ongeveer 6% tegen een krappe 2%.

Kennis van de context gebruiken

Ook beantwoorden de leerlingen soms een andere vraag dan de gestelde. Leerlingen geven hun antwoorden soms uitsluitend binnen de context. Dit verschijnsel zullen veel docenten herkennen. Er wordt een goed antwoord gegeven maar helaas, niet op de vraag die gesteld werd. Zoals in het volgende voorbeeld, waarin een leerling op de vraag: 'Hoe kun je aan de grafiek zien dat het gemiddelde groeitempo van meisjes langzamer wordt na hun 12de jaar?' antwoordde:

De puberteit begint bij meisjes eerder en is daarom ook eerder afgelopen.

De vmbo-leerlingen betrekken de situatie ook regelmatig op zichzelf en relateren hun antwoorden aan hun eigen situatie of hun kennis erover. Ze schrijven dan ook vaker in de eerste persoon:

Nee, het is maar 7 meer, ik vind dat niet enorm.

Ik denk dat het dom is dat te zeggen, je weet niet eens wat 505-520 betekent, dus daarom ben ik het er niet mee eens.

Verschillen tussen vmbo en havo/vwo-leerlingen

Aan de hand van de opgave over afval (zie fig. 16) illustreren we enkele verschillen in antwoorden tussen vmbo-leerlingen enerzijds en havo/vwo-leerlingen anderzijds.

AFVAL

Voor een werkstuk over het milieu verzamelen leerlingen gegevens over de afbraaktijd van verschillende soorten weggegooid afval.

Soort afval	Afbraaktijd
Bananenschil	1 - 3 jaar
Sinaasappelschil	1 - 3 jaar
Kartonnen dozen	0,5 jaar
Kauwgum	20 - 25 jaar
Kranten	een paar dagen
Plastic bekertjes	meer dan 100 jaar

Een leerling wil de resultaten weergeven in een staafdiagram.

Geef één argument waarom een staafdiagram niet geschikt is om deze gegevens weer te geven.

figuur 16 Opgave: Afval

De volgende antwoorden zijn gegeven door vmbo-leerlingen. Een groot aantal van deze antwoorden komt in min of meer dezelfde bewoordingen vaker dan een keer voor.

Allereerst een serie als onjuist beoordeelde antwoorden:

Het heeft niet veel nut.

Dan wordt het aantal niet goed zichtbaar.

Dat is niet duidelijk genoeg.

Omdat het allemaal niet past en dan klopt het ook niet.

Omdat het heel erg is wat er gebeurt.

Het zijn geen precieze cijfers.

Omdat de afbraaktijd verschillend is.

Dat kan niet omdat je moet selecteren van tijd tot tijd.

Je weet geen exacte tijden.

Omdat het niet helemaal zeker is dat de spullen zo lang blijven liggen want volgens mij wordt er maar gegokt.

In deze antwoorden zien we een aantal van de eerder genoemde aspecten: alledaags taalgebruik, slordige of onduidelijke formuleringen, eigen (vermeende) kennis over de context, geen antwoord geven op de vraag, het gebruik van de eerste persoon.

Hieronder een aantal als goed beoordeelde antwoorden van vmbo-leerlingen:

Omdat je de krant niet kunt uitrekenen en plastic bekertjes is meer dan honderd dus kan niet.

Onlogisch, heeft geen zin, want dan heb je een heel grote staafdiagram.

Omdat niet overal dezelfde afbraaktijd wordt genomen kun je het niet goed vergelijken en ook niet weergeven in een staafdiagram.

Dan wordt het een rommeltje. De een begint bij 0,5 jaar en de andere bij meer dan 100 jaar.

Ook in deze goede antwoorden zien we het alledaagse taalgebruik. Geen enkele leerling gebruikt wiskundige termen zoals as, stapgrootte, staaflengte en dergelijke.

Ter vergelijking ook enkele antwoorden van havo/vwo-leerlingen. Eerst twee onjuiste antwoorden:

Omdat je niet weet hoeveel afval er is per soort.

Het wordt te ingewikkeld omdat sommige dingen in dagen staan.

En twee juiste:

Dan moet je een verticale as nemen die tot 100 gaat met een lage stapgrootte, want 'een paar dagen' moet erin af te lezen zijn.

Omdat het niet precies is, onder andere de kranten en de plastic bekertjes.

Deze antwoorden zijn iets netter geformuleerd en in het eerste antwoord zijn wiskundige termen gebruikt. Antwoorden van havo/vwo-leerlingen onderscheiden zich met name op deze twee aspecten van de antwoorden van vmbo-leerlingen.

Resultaten en conclusies

De PISA-items op zich geven een goed beeld van de prestaties van verschillende groepen vmbo-leerlingen. De toetsopgaven zijn zodanig dat ze een gedifferentieerd beeld laten zien.

Met name op het domein Veranderingen en Relaties scoren Nederlandse leerlingen goed, ook de vmbo-leerlingen. In het algemeen scoren Nederlandse leerlingen goed op opgaven waarbij een grafische voorstelling afgelezen of geïnterpreteerd moet worden. Redeneervragen bij grafieken en diagrammen scoren goed, evenals vragen uit het domein Onzekerheid. Nederlandse leerlingen hebben wat meer moeite met kalere opdrachten, met complex rekenwerk en met het opvolgen van algoritmes. Voor de vmbo-leerlingen geldt dit in sterkere mate dan voor leerlingen van havo/vwo. De Duitse leerlingen doen deze opgaven juist opvallend goed.

Het blijkt duidelijk dat de vmbo-leerlingen achterblijven op de formelere en abstractere opgaven. Dit geldt niet alleen voor de opgaven met formules, maar ook voor bijvoorbeeld meetkundeopgaven die abstract van aard zijn of die een gekunstelde context hebben. Waar havo/vwo-leerlingen er nog in lijken te slagen door deze contexten heen te kijken, lukt dat de vmbo-leerlingen niet. We vermoeden dat door het gebrek aan relevantie en authenticiteit van een dergelijk probleem de motivatie om het op te lossen afneemt, en dat dit gecombineerd met de moeilijkheidsgraad leidt tot de zeer lage score bij de vmbo-leerlingen.

Het lijkt vooral de vmbo-leerlingen te ontbreken aan wiskundig zelfvertrouwen en een onderzoekende wiskundige houding. De leerlingen lijken erg gericht te zijn op het geven van een antwoord. Ze geven het na een mislukte of half gelukte eerste poging snel op in plaats van te reflecteren op het antwoord of de methode en nog eens iets anders te proberen. Ook maken ze opvallend weinig schetsjes of tussenberekeningen. Met opgaven die wat leeswerk betreft complex en lang zijn, hebben met name de vmbo-leerlingen grote moeite. Dit zijn de typen opgaven die deze leerlingen nog wel eens overslaan. Een ander punt dat opvalt is de grote moeite die vmbo-leerlingen hebben met het verwoorden van argumenten en redeneringen. Dit is niet alleen een 'gewoon' taalprobleem, ook het onvoldoende in staat zijn wiskundige terminologie op de juiste manier te gebruiken speelt hierbij een rol.

Al met al zijn de scores van de Nederlandse vmbo-leerlingen niet slecht in internationaal perspectief. Vooral de zwakste groep(en) doen het relatief nog goed. Toch is er reden tot zorg vanwege het geringe deel van deze leerlingen dat vaardigheidsniveau 4 haalt. Dat is het niveau waarbij een wiskundige onderzoekende houding nodig is. Dat het merendeel van de vmbo-leerlingen dat niet haalt geeft redenen om het wiskundeonderwijs aan deze leerlingen nog eens nader te bekijken.

Resumerend kan gesteld worden:

- 1 Nederlandse vmbo-leerlingen in de beroepsgerichte leerwegen (de zwakste groep) doen het in internationale vergelijking met onze buurlanden relatief goed.
- 2 Enige zorg over de wiskundige geletterdheid van de Nederlandse vmbo-leerlingen is gerechtvaardigd omdat slechts weinigen vaardigheidsniveau 4 bereiken.
- 3 Nederlandse vmbo-leerlingen zijn goed in het omgaan met grafieken maar hebben moeite met meer abstracte, formele en kale opgaven. Duitse en Vlaamse leerlingen zijn hierin duidelijk beter.
- 4 Nederlandse vmbo-leerlingen lijken gebrek aan wiskundig doorzettingsvermogen te hebben.
- 5 Met name de verwoording van redeneringen en argumenten en daarbij het gebruiken van wiskundetaal is lastig voor veel vmbo-leerlingen.

Aanbevelingen voor het wiskundeonderwijs op het vmbo

PISA beoogt de wiskundige geletterdheid van leerlingen te toetsen. Wiskundig geletterd zijn is van belang voor functioneren in maatschappij, school en beroep; daar zullen de meeste mensen het wel over eens zijn. Natuurlijk is dat niet het enige aspect dat daarvoor van belang is. Hoewel we niet alle PISA-opdrachten even relevant vinden voor de wiskundige geletterdheid van de Nederlandse vmbo-leerlingen willen we op grond van de analyses toch enkele aanbevelingen voor het wiskundeonderwijs in het vmbo doen.

We bepleiten voor de meerderheid van de vmbo-leerlingen niet om te streven naar een hoger beheersingsniveau van de formele en abstracte wiskunde. Voor deze groep leerlingen is het ontwikkelen van een algemeen hoger niveau van wiskundige geletterdheid van meer belang. Daarbij hoort wel de vergroting van het zelfvertrouwen op het gebied van wiskunde en de ontwikkeling van een wiskundige houding.

Op de eerste plaats is het daarvoor van belang om meer inhoudelijke interactie te realiseren over de wijze van oplossen en over de kwaliteit van antwoorden en redeneringen. Het is aan te bevelen leerlingen elkaars werk te laten bekijken. Door hierover te praten, leren leerlingen hun eigen en andermans antwoorden op waarde te schatten.

Dit vereist deels een ander type opdrachten dan die in de meeste schoolboeken staan. In de Nederlandse wiskundeboeken komen veel opdrachten voor die bestaan uit deelvragen waarop een kort antwoord voldoet. Er komt zelden een echte complexe probleemstelling voor en er worden weinig redeneringen gevraagd. Er wordt niet echt gewerkt aan de wiskundige geletterdheid zoals die door PISA is gedefinieerd, en zoals die voor het halen van niveau 4 gewenst is.

Uit het feit dat de Nederlandse leerlingen in het algemeen, en vmbo-leerlingen in het bijzonder, slecht scoren op de vragen uit het cluster Reflectie en op het gehele domein probleemoplossen, concluderen we dat er behoefte is aan authentieke probleemstellingen waarin dit type hogere orde vaardigheden toegepast moeten worden.

Er liggen kansen voor betekenisvolle opdrachten in de directe schoolomgeving van de leerlingen door bijvoorbeeld aan te sluiten bij de sectorspecifieke beroepsgerichte

vakken en daarin de wiskunde te gebruiken om een probleem op te lossen. Voor de theoretische leerweg zou er beter kunnen worden voorbereid op de wiskunde van de havo.

Er zou minder nadruk kunnen worden gelegd op het geven van een antwoord alleen. Leerlingen zouden moeten leren dat een eerste poging niet altijd meteen goed hoeft te zijn, dat het maken van schetsjes en het gebruik van kladpapier iets is wat helpt (en mag).

Er zou voor de vmbo-leerlingen meer aandacht gegeven moeten worden aan de taal, zowel aan het lezen, spreken en begrijpen als aan het schrijven, bijvoorbeeld bij het verwoorden en opschrijven van het oplossingsproces. Daarbij zouden meer verbindingen tussen alledaagse taal en wiskundige begrippen moeten worden gelegd.

Bij al het bovenstaande geldt dat er meer gelegenheid zou moeten zijn voor onderlinge uitwisseling (interactie) en reflectie. Er zou wat minder nadruk kunnen liggen op het individueel maken en nakijken van sommen, want dat lijkt toch wat anders te zijn dan wiskundige geletterdheid. Elkaars oplossingen en verwoordingen bekijken, bespreken en beoordelen kan een goede aanvulling in die richting zijn.

We willen ten slotte nog opmerken dat de PISA-opgaven, hoewel de contexten regelmatig betrekking hebben op authentieke situaties, op leerlingen toch in essentie als schoolse toetsvragen zullen overkomen. Om de wiskundige geletterdheid van vmbo-leerlingen in met name de beroepsgerichte leerwegen in de volle breedte te kunnen vaststellen zal er ook gekeken moeten worden naar hoe die leerlingen hun wiskundige kennis en vaardigheden inzetten in authentieke situaties, zoals die bijvoorbeeld in de beroepsgerichte programma's in de opleiding voorkomen. Een dergelijke *setting* ligt buiten de mogelijkheden van PISA, maar kan wel in het onderwijs zelf gerealiseerd worden.

5. Over probleemoplossen en wiskunde

Kees Lagerwaard, Ger Limpens

Inleiding

In de uitgangspunten van het PISA-onderzoek is vastgelegd dat er behalve aandacht voor de drie kerngebieden leesvaardigheid, wiskundige geletterdheid en natuurwetenschappen, ook aandacht zal zijn voor een per cyclus variërend onderzoeksgebied. In PISA 2003 was dat het niet-curriculum gebonden onderwerp probleemoplossen. Bij wiskundige geletterdheid neemt de vaardigheid probleemoplossen een vrij belangrijke plaats in. Daarom is het interessant om te zien wat de overeenkomsten en verschillen zijn tussen het algemene domein probleemoplossen en wat we binnen wiskunde onder probleemoplosvaardigheden verstaan.

Wat wordt binnen PISA bedoeld met het onderwerp probleemoplossen?

De behoefte om probleemoplossen als domein te introduceren in het PISA-onderzoek is ingegeven door het feit dat het oplossen van problemen in het algemeen een belangrijk educatief doel is in het schoolprogramma in ieder land. Het onderwijsveld en beleidsmakers vinden het van groot belang dat leerlingen de vaardigheden verwerven die nodig zijn om problemen in werkelijke situaties op te lossen. Dat betekent allereerst dat leerlingen de hun aangeboden informatie begrijpen. Leerlingen moeten vervolgens uit de geboden informatie de belangrijkste elementen kunnen halen die het probleem representeren en kunnen begrijpen hoe die elementen samenhangen. Ze moeten daarna zelf een concrete voorstelling kunnen maken van het probleem. Ten slotte moeten ze het probleem kunnen oplossen en hun oplossing kunnen beoordelen, beargumenteren en overdragen aan anderen. In de ogen van de OESO is probleemoplossen een vaardigheid die een basis verschaft om effectief te kunnen deelnemen aan het maatschappelijk verkeer en om persoonlijke activiteiten te kunnen ontplooiën.

Probleemoplossen is een vaardigheid die dwars door het schoolcurriculum heen loopt. In allerlei schoolvakken komt deze vaardigheid aan bod, maar steeds binnen de context van dat vak en dus in combinatie met vakspecifieke kennis. Hier moet dan ook het verschil worden gezocht tussen opgaven uit het domein probleemoplossen en opgaven uit de andere domeinen wiskunde, natuurwetenschappen en leesvaardigheid waar ook een beroep wordt gedaan op probleemoplosvaardigheden. Een uitgebreide verhandeling over het domein is te vinden in de internationale publicatie hierover (OECD, 2003a).

Gegeven de PISA-context is besloten bij de ontwikkeling van opgaven de nadruk te leggen op drie soorten situaties waarin leerlingen hun vaardigheden op dit gebied moeten tonen. Dit zijn situaties waarin ze:

- beslissingen moeten nemen onder zekere voorwaarden;
- systemen moeten beoordelen en ontwerpen voor een specifieke toepassing;
- een oplossing moeten bedenken voor een slecht functionerend systeem of voorwerp op basis van een serie symptomen (*trouble shooting*).

Voor de opgaven heeft dat geresulteerd in een drietal typen opgaven: de besluitvormingsitems, de systeemanalyse- en ontwerpitems en de *trouble shooting*-items.

Bij de toetsing van de vaardigheid probleemoplossen gaat het vooral om het toetsen van de processen die ten grondslag liggen aan het oplossen van problemen. Daarbij heeft men proberen te vermijden dat er veel inhoudskennis nodig zou zijn. In de bij deze publicatie horende bundel met PISA-opgaven zijn alle opgaven probleemoplossen opgenomen. De cognitieve vaardigheden die bij deze opgaven van het domein probleemoplossen bij uitstek getoetst worden, zijn analytisch denken en redeneren. In de meeste opgaven moeten de leerlingen informatie systematiseren en analyseren. Vervolgens moeten ze oplossingen vinden die aan de gegeven voorwaarden voldoen. Heel vaak moet men in het dagelijks leven problemen oplossen waarbij men niet veel inhoudelijke kennis nodig heeft, maar waarbij men in staat moet zijn te redeneren. Men moet een probleem systematisch kunnen benaderen door elimineren, opsommen en indelen. Dit zijn kortom de te toetsen vaardigheden bij opgaven probleemoplossen.

In Nederland probleemoplossen slechter dan wiskunde

Het is opmerkelijk dat de score in Nederland voor probleemoplossen aanzienlijk lager ligt dan de score van wiskunde. Het scoreverschil tussen deze twee domeinen is, vergeleken met de andere OESO-landen, het grootst in Nederland. In het internationale rapport geeft men de hoge kwaliteit van het wiskundeonderwijs in Nederland als verklaring. In de internationale rangorde bevindt Nederland zich wat betreft wiskunde op plaats 4 en bij probleemoplossen op plaats 12. Hierbij dient te worden opgemerkt dat de schalen voor wiskunde en probleemoplossen onafhankelijk van elkaar tot stand zijn gekomen. Dat betekent onder andere dat de waarde 538 op de PISA-wiskundeschaal niet zonder meer mag worden vergeleken met het getal 520 op de schaal van probleemoplossen.

Anderzijds zijn de twee schalen wel ontworpen op basis van de prestaties van alle deelnemende landen. Het gemiddelde van alle scores bij wiskunde is op 500 gezet en de standaardafwijking op deze wiskundeschaal is 100. De scores van de Nederlandse leerlingen zijn dus gerelateerd aan het 'mondiale gemiddelde'. Precies hetzelfde is gedaan met de scores bij probleemoplossen.

Je kunt dus zeggen dat gelijke scores bij wiskunde en probleemoplossen op deze 'wereldschalen' in feite ook vergelijkbare prestaties zijn, mondiaal gezien. Vanuit dat perspectief kunnen we de score 538 bij wiskunde van Nederland vergelijken met de score 520 bij probleemoplossen. Maar wat misschien veelzeggender is, is het verschil in rangorde op beide schalen. In de tabellen 1 en 2 is goed te zien dat de positie van Nederland bij probleemoplossen een heel andere is dan die bij wiskunde: bij wiskunde is Nederland daadwerkelijk een koploper, terwijl Nederland bij probleemoplossen 'slechts' een goede middenmoter is.

De constatering, gecombineerd met het feit dat binnen het domein van wiskundige geletterdheid probleemoplossen een belangrijke plaats inneemt, is aanleiding om nader te onderzoeken waarom Nederland bij wiskunde bij de koplopers hoort maar bij probleemoplossen slechts in de middengroep eindigt.

Tabel 1 Gemiddelde score op de wiskundeschaal

	Gemiddelde
*Hong Kong - China	550
Finland	544
Zuid-Korea	542
Nederland	538
*Liechtenstein	536
Japan	534
Canada	532
België	529
*Macau - China	527
Zwitserland	527
Australië	524
Nieuw-Zeeland	523
Tsjechië	516
IJsland	515
Denemarken	514
Frankrijk	511
Zweden	509
Oostenrijk	506
Ierland	503
Duitsland	503
Slowakije	498
Noorwegen	495
Luxemburg	493
Polen	490
Hongarije	490
Spanje	485
*Letland	483
Verenigde Staten	483
*Russische Federatie	468
Portugal	466
Italië	466
Griekenland	445
*Servië	437
Turkije	423
*Uruguay	422
*Thailand	417
Mexico	385
*Indonesië	360
*Tunesië	359
*Brazilië	356

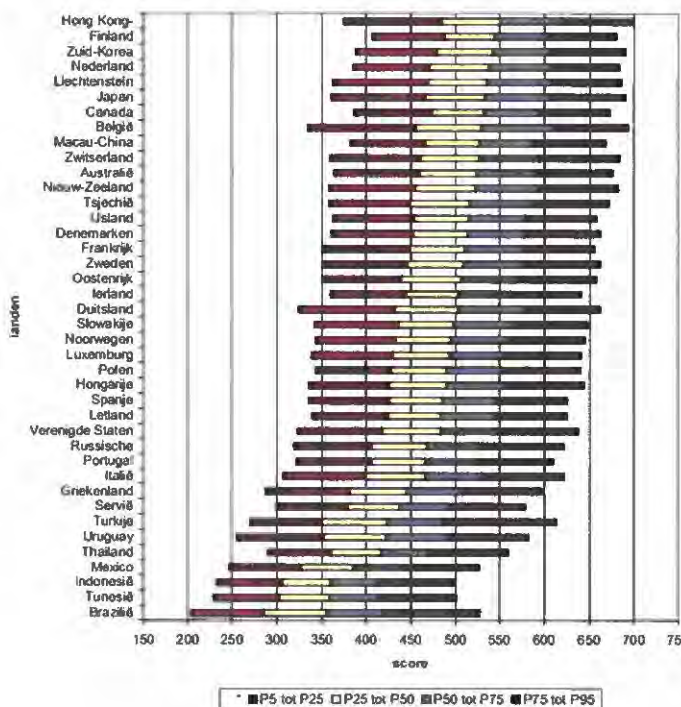
Tabel 2 Gemiddelde score op de schaal van probleemoplossen

	Gemiddelde
Korea	550
Finland	548
*Hong Kong - China	548
Japan	547
Nieuw-Zeeland	533
*Macau - China	532
Australië	530
Canada	529
*Liechtenstein	529
België	525
Zwitserland	521
Nederland	520
Frankrijk	519
Denemarken	517
Tsjechië	516
Duitsland	513
Zweden	509
Oostenrijk	506
IJsland	505
Hongarije	501
Ierland	498
Luxemburg	494
Slowakije	492
Noorwegen	490
Polen	487
*Letland	483
Spanje	482
*Russische Federatie	479
Verenigde Staten	477
Portugal	470
Italië	469
Griekenland	448
*Thailand	425
*Servië	420
*Uruguay	411
Turkije	408
Mexico	384
*Brazilië	371
*Indonesië	361
*Tunesië	345

* Partnerlanden van OESO

(Het Verenigd Koninkrijk heeft niet volledig aan de steekproeven voldaan, waardoor de resultaten niet officieel zijn erkend.)

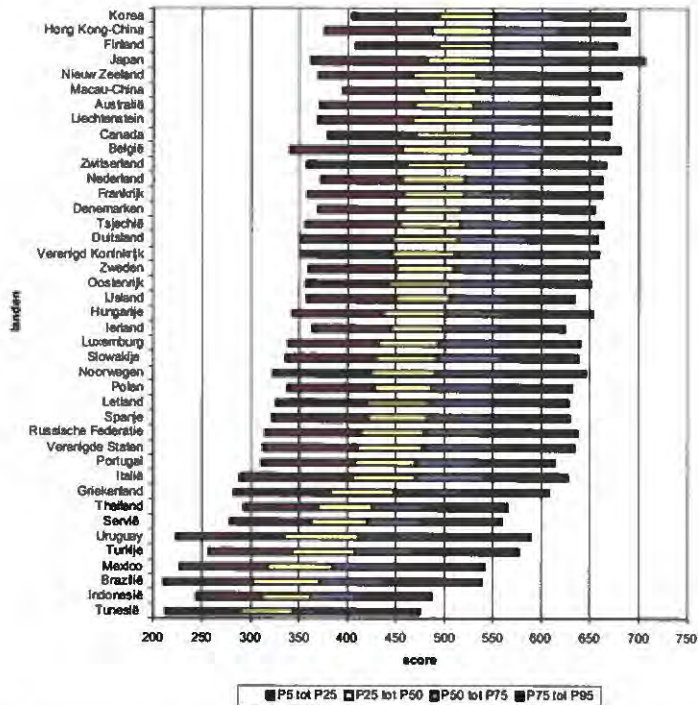
De scoreverdeling totaal binnen Nederland bij probleemoplossen en bij wiskunde



figuur 1 Scoreverdeling op de wiskunde-schaal in de OESO- en partnerlanden

Als we in de scoreverdeling in figuur 1 (ook opgenomen in het Nederlandse PISA 2003-rapport) kijken naar Nederland en we laten de 5% zwakste en de 5% beste leerlingen buiten beschouwing, dan zien we bij wiskunde dat de afstand tussen P5 en P95 een kleine 300 is op de al eerder beschreven 500-1000 schaal. De scores lopen van ongeveer 375 tot ongeveer 675.

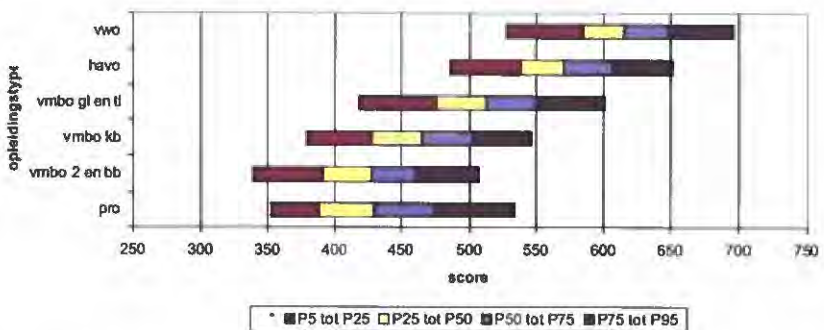
Ook bij probleemoplossen (zie fig. 2), eveneens uit het Nederlandse PISA 2003-rapport, vinden we een afstand van circa 300 punten tussen P5 en P95. Binnen de 4 kwartielen is de verdeling vrijwel dezelfde. Ook de P5 en de P95 liggen bij probleemoplossen dicht bij die van wiskunde, te weten van 370 tot 660. Hier zien we dus vrijwel hetzelfde patroon, maar een beetje lager. Maar het resulteert wel in een 18 punten lager gemiddelde: bij probleemoplossen 520 en bij wiskunde 538. En de verdelingen bij wiskunde en probleemoplossen van de middelste 90% (de breedte van de score-intervallen van de diverse groepen waarin P5, P25, P50, P75 en P95 die 90% verdelen) zijn heel goed vergelijkbaar (zie beide figuren): de opbouw van het scoreverloop bij wiskunde en probleemoplossen is, met andere woorden, nagenoeg gelijk.



figuur 2 Scoreverdeling op de probleemoplossen-schaal in de OESO- en partnerlanden

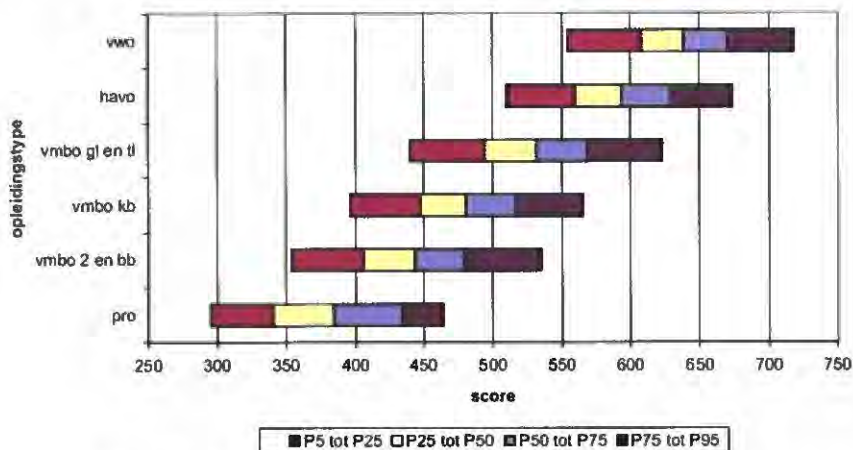
De scoreverdeling binnen de schooltypen in Nederland bij probleemoplossen en bij wiskunde

Over onderstaande figuur 3 merken we ten eerste op dat de groep leerlingen die hier aangeduid wordt met 'pro' (het praktijkonderwijs) een groep is die in de PISA-steekproef te klein is om verantwoorde conclusies op te baseren: het betrof hier slechts een dertigtal leerlingen. Alle andere opleidingstypes vertonen een regelmatig stijgende vaardigheid wat betreft probleemoplossen naarmate het onderwijstype hoger is.



figuur 3 Scoreverdeling op de schaal van probleemoplossen per opleidingstype

Precies hetzelfde patroon zien we als we in onderstaande figuur 4 (figuur 2.11 uit het Nederlandse rapport) kijken naar de vaardigheidsverdelingen per opleidingstype bij wiskunde.

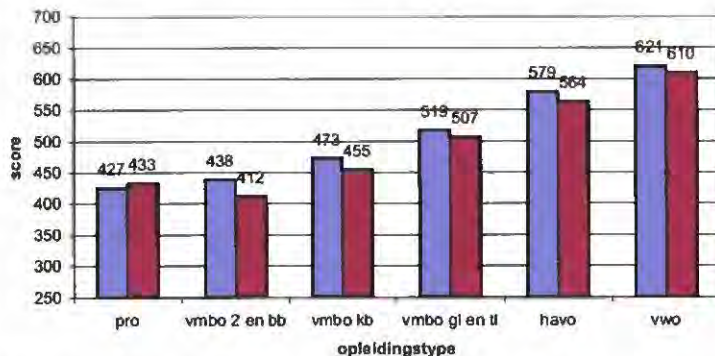


figuur 4 Scoreverdeling op de wiskundeschaal per opleidingstype

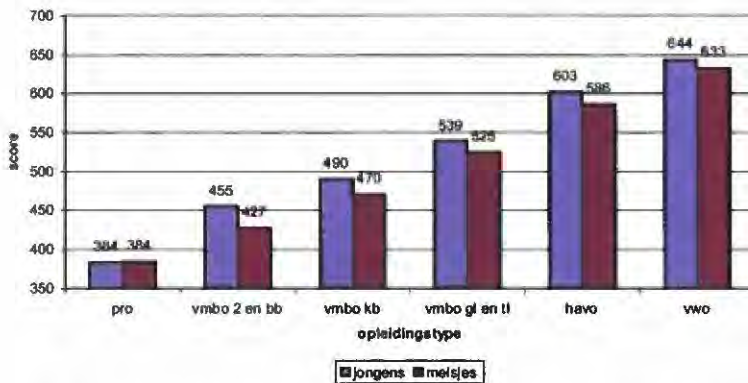
Als we het geheel bekijken ten opzichte van het mondiale scorepatroon, dan kunnen we zeggen dat ook binnen elk schooltype de scores bij wiskunde en probleemoplossen een volstrekt vergelijkbare opbouw hebben. Alleen is de score bij probleemoplossen consistent zo'n 15 tot 20 punten lager dan bij wiskunde.

Over de schooltypen heen zien we dat leerlingen in het PISA-onderzoek een zelfde aantal punten lager scoorden bij de opdrachten probleemoplossen dan op de wiskunde-items. Mochten we een verschil willen signaleren, dan zouden we kunnen zeggen dat het scoreverschil wat toeneemt naarmate het schooltype hoger wordt. De wiskundescore van de BB leerlingen steekt het minst boven hun probleemoplosscore uit, terwijl dat verschil op vwo-niveau het grootst is.

Jongens/meisjes bij probleemoplossen en wiskunde



figuur 5 Gemiddelde score van jongens en meisjes op de schaal van probleemoplossen

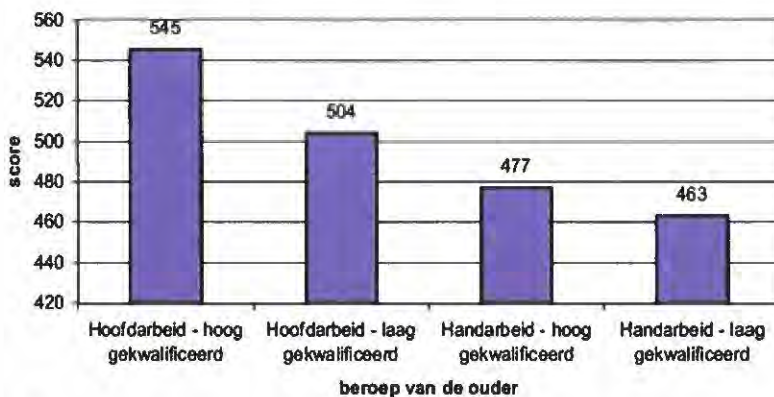


figuur 6 Gemiddelde score van jongens en meisjes op de wiskundeschaal

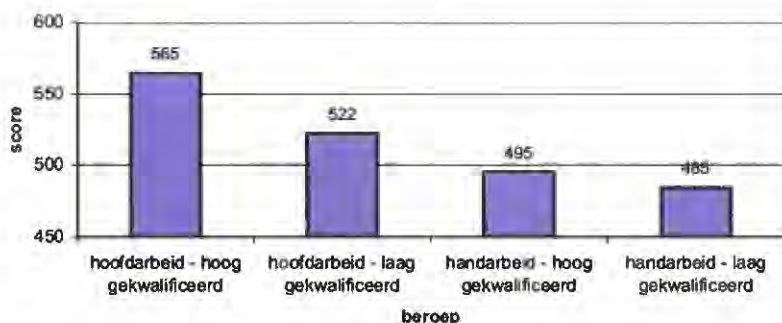
Als we binnen de schooltypen kijken naar de verschillen in score van jongens en meisjes zien we (zie bovenstaande figuren 5 en 6, de figuren 5.5 respectievelijk 2.15 uit het Nederlandse rapport) opnieuw dat vergelijkbare beeld bij probleemoplossen en wiskunde.

In alle schooltypen vmbo BB, KB, GL/TL, havo en vwo zijn de scores van de jongens gemiddeld steeds iets hoger. Het verschil is het grootst binnen vmbo2+BB, namelijk 26 bij probleemoplossen en 28 bij wiskunde. Het verschil is bij KB 18, respectievelijk 20, bij GL/TL vinden we 12 en 14, bij havo 15 en 17 en bij vwo twee keer 11. Dus ook bij de verschillen in score tussen jongens en meisjes zijn de uitkomsten bij probleemoplossen en wiskunde vergelijkbaar met de kanttekening dat de verschillen bij probleemoplossen een heel klein beetje kleiner zijn.

Beroep van de ouder(s) en scores bij probleemoplossen en wiskunde



figuur 7 Gemiddelde score op de schaal van probleemoplossen in relatie tot het beroep van de ouder



figuur 8 Wiskundescore in relatie tot het beroep van de ouder

Het verhaal wordt eentonig. Als de leerlingen worden ingedeeld naar beroep(sgroepen) van de ouders – zie figuur 7 en 8 (figuren 5.8 respectievelijk 2.21 uit het Nederlandse rapport) – zien we opnieuw voor de scores bij probleemoplossen en wiskunde een consistent verschil van 15 tot 20 punten.

Het voorstellen van enkele probleemoplos-items

Voordat we proberen de inhoudelijke relatie tussen items van probleemoplossen en wiskunde-items iets meer te duiden, beschouwen we enkele voorbeelden van probleemoplos-items. Zie hierna in figuren 9, 10 en 11 de voorbeelden: Vakantie (vraag 1 en vraag 2), Vakantiekamp (vraag 1) en Irrigatie (vraag 1). Overigens treft u in de bundel PISA-opgaven alle items van het onderzoeksgebied probleemoplossen aan. Omdat dit gebied in de toekomst geen onderdeel uit zal maken van verdere PISA-studies zijn alle bijbehorende items voor publicatie vrijgegeven. In de bundel treft u ook enkele gegevens per item aan over de wijze waarop Nederlandse leerlingen deze vragen beantwoord hebben.

Vakantie

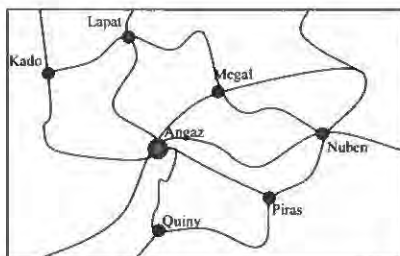
De opgave Vakantie (zie fig. 9) bevat twee vragen, een kort antwoord vraag en een zogeheten lang antwoord vraag. Beide vragen vallen onder de eerder omschreven categorie besluitvormingsitems. Bij dit type opgave moet een leerling beslissingen nemen onder bepaalde, gegeven voorwaarden. De eerste vraag is min of meer bedoeld als verkennende activiteit: aan de hand van kaart en tabel moet vastgesteld worden a) welke verbinding de kortste weg oplevert tussen het tweetal gegeven plaatsen en b) hoe lang die kortste verbinding is. De op de kaart voor de hand liggende verbinding blijkt ook na berekening de gezochte verbinding. Bij de tweede activiteit moet vastgesteld worden waar gedurende een zekere periode overnacht moet worden, gegeven zekere beperkingen. De reeds gedeeltelijk ingevulde tabel geeft al enige zekerheid, waarna de rest via een rechtlijnige redenering (waarmee niet gezegd is dat dit een eenvoudige redenering zou zijn) volgt. Over alle schooltypes heen gaf in Nederland 54% van de leerlingen een score 1-antwoord bij vraag 1. Vraag 2 is een vraag die een deelscore kent: 42% van de Nederlandse leerlingen gaf een score 2-antwoord en slechts 4% van de Nederlandse leerlingen werd met score 1 beloond. In de bij deze publicatie horende opgavenbundel is terug te vinden hoe de Nederlandse leerlingen per schooltype scoorden bij deze opgave. De opgave Vakantie is daarmee toch een redelijk moeilijke opgave.

VAKANTIE

Dit probleem gaat over het vinden van de beste route voor een vakantie.

In figuur 1 en 2 zie je een kaart van het gebied en de afstanden tussen de dorpen.

Figuur 1: Een wegenkaart van de wegen tussen de dorpen



Figuur 2: De kortste afstand tussen de dorpen in kilometers.

Angaz							
Kado	550						
Lapat	500	300					
Megal	300	850	550				
Nuben	500		1000	450			
Piras	300	850	800	600	250		
Quiny	350	900	850	650	800	550	
	Angaz	Kado	Lapat	Megal	Nuben	Piras	Quiny

Vraag 1: VAKANTIE

Bereken de kortste afstand over de weg tussen Nuben en Kado.

Vraag 2: VAKANTIE

Anna woont in Angaz. Ze wil **Kado** en **Lapat** bezoeken. Ze kan op één dag maar 300 kilometer reizen, maar ze kan de reis opsplitsen door te overnachten op campings tussen de dorpen.

Anna overnacht **twee nachten** in elk dorp, zodat ze in elk dorp een hele dag heeft om rond te kijken.

Stippel Anna's reisroute uit door in de volgende tabel in te vullen waar ze telkens de nacht doorbrengt.

Dag	Overnachting
1	Camping tussen Angaz en Kado
2	
3	
4	
5	
6	
7	Angaz

figuur 9 Opgave: Vakantie

Vakantiekamp

De opgave Vakantiekamp (zie fig. 10) is een opgave (met slechts één vraag, een lang antwoord vraag) uit de categorie Systeemanalyse en -ontwerp, waarbij de leerling een systeem moet beoordelen en/of ontwerpen voor een specifieke toepassing. In dit geval betreft het de indeling van een zevental slaapzalen, uitgaande van een aantal 'regels', aantallen jongens, meisjes, volwassenen en bedden op de diverse slaapzalen. Dit is een behoorlijk complex probleem. Een leerling zal zich, bij het beantwoorden van deze vraag, in eerste instantie moeten realiseren dat het probleem slechts stap voor stap kan worden opgelost en zal een keuze moeten maken welk aspect de eerste aandacht behoeft. Er zijn – en dat maakt het voor leerlingen zeker niet eenvoudiger – meerdere juiste antwoorden mogelijk. Een leerling die daarvan bewust wordt tijdens de beantwoording, zal dat vermoedelijk als een extra complicerende factor ervaren.

De opgave Vakantiekamp kende ook een beoordeling op deelscoreniveau. Van de Nederlandse leerlingen scoorde 21% de maximale score 2. En 41% kreeg een score 1. Opvallend hierbij is wellicht dat het percentage score 1-leerlingen nagenoeg onafhankelijk bleek te zijn van het schooltype. Zie daarvoor ook de opgavenbundel.

Irrigatie

De opgave Irrigatie (zie fig. 11) heeft een drietal vragen. In dit deel van de publicatie is slechts de eerste vraag van de context Irrigatie opgenomen. De gehele opgave Irrigatie valt in de categorie *trouble shooting*. Hierin moet een leerling een oplossing bedenken voor een slecht functionerend systeem of voorwerp op basis van een serie symptomen. De eerste vraag van deze context is een min of meer verkennende activiteit. In de bij dit rapport horende opgavenbundel treft u ook de andere vragen die onder deze context vallen aan. In vraag 1 moeten alle mogelijke routes die bij een bepaalde set instellingen horen getekend worden. Wat bij deze opgave bij leerlingen overigens nog wel tot een probleem kan leiden, is het feit dat in de figuur alle mogelijke sluisdeuren als gesloten getekend lijken te zijn terwijl de verdere opgave duidelijk maakt dat dit niet altijd het geval is. De eerste vraag van Irrigatie werd redelijk goed gemaakt: in Nederland behaalde circa 74% van de populatie de maximale score 1. Van de vwo 4-leerlingen in de steekproef gaf ruim 97% een juist antwoord op deze vraag. Zie voor meer details de opgavenbundel.

VAKANTIEKAMP

Het Jongerenwerk in Zed organiseert een vakantiecamp van 5 dagen. Er hebben zich 46 kinderen (26 meisjes en 20 jongens) ingeschreven voor het kamp, en acht volwassenen (vier mannen en vier vrouwen) hebben zich als vrijwilliger aangemeld om het kamp te organiseren en de kinderen te begeleiden.

figuur 1: Volwassenen

Mevr. Grietjes
Mevr. Chantal
Mevr. Merel
Dhr. Nelissen
Dhr. Stevens
Mevr. Kelly
Dhr. Willems
Dhr. Peters

figuur 2: Slaapzalen

Naam	Aantal bedden
Rood	12
Blauw	8
Groen	8
Paars	8
Oranje	8
Geel	6
Wit	6

Regels voor de slaapzalen:

- 1 Jongens en meisjes slapen op gescheiden slaapzalen.
- 2 Op elke slaapzaal slaapt ten minste één volwassene.
- 3 De volwassene(n) op de slaapzaal zijn van hetzelfde geslacht als de kinderen.

Vraag: VAKANTIEKAMP**Slaapzaalindeling**

Vul de tabel op de volgende pagina in om de 46 kinderen en de 8 volwassenen te verdelen over de slaapzalen. Zorg ervoor dat je alle regels daarbij toepast.

Naam	Aantal jongens	Aantal meisjes	Naam van volwassene(n)
Rood			
Blauw			
Groen			
Paars			
Oranje			
Geel			
Wit			

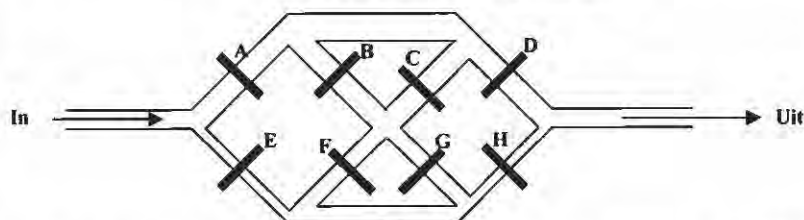
figuur 10 Opgave: Vakantiekamp

IRRIGATIE

In de figuur hieronder zie je een overzicht van een systeem met irrigatiekanalen voor de bevoeiing van landbouwgrond. De sluisdeuren A tot en met H kunnen worden geopend en gesloten om het water te laten stromen naar de plaatsen waar het nodig is. Als een sluisdeur dicht is, kan er geen water doorheen.

Dit probleem gaat over het zoeken van een sluisdeur die vastzit, waardoor er geen water door het kanalsysteem stroomt.

Figuur 1: Een systeem met irrigatiekanalen



Michael merkte dat het water niet altijd loopt waar het naartoe zou moeten lopen. Hij denkt dat een van de sluisdeuren vastzit en dus niet opengaat als men de deur op 'open' wil zetten.

Vraag 1: IRRIGATIE

Michael gebruikt de instellingen uit tabel 1 om de sluisdeuren te testen.

Tabel 1: Instellingen van de sluisdeuren

A	B	C	D	E	F	G	H
Open	Dicht	Open	Open	Dicht	Open	Dicht	Open

Ga uit van de sluisdeurinstellingen uit tabel 1 en teken in afbeelding 1 alle mogelijke routes die het water kan volgen, aangenomen dat alle sluisdeuren werken volgens de instellingen.

figuur 11 Opgave: Irrigatie

Inhoudelijke relatie wiskunde – probleemoplossen

Bovenstaande voorbeelden geven al aan dat er, ondanks het feit dat er slechts elementaire wiskundige technieken gehanteerd hoeven te worden bij probleemoplossings-items, toch een sterke verwantschap tussen probleemoplossen en wiskunde in PISA is. De items, hetzij wiskunde, hetzij probleemoplossen doen namelijk regelmatig een beroep op de vaardigheid 'analytisch redeneren'.

Factor-analyse en enkele wiskunde- en probleemoplos-items nader beschouwd

Een factoranalyse (hierbij wordt met statistische technieken naar achterliggende factoren gezocht die correlaties tussen gemeten variabelen kunnen verklaren) bracht aan het licht dat het hiervoor genoemde analytisch redeneren veel probleemoplos-vragen verbindt met een aanzienlijk aantal wiskundevragen. Het internationale rapport over de uitkomsten bij probleemoplossen in PISA 2003 besteedt aandacht aan de relatie wiskunde en probleemoplossen. Op basis van dit rapport kunnen de volgende zaken gemeld worden:

- De sterkste verbinding werd aangetroffen tussen het probleemoplos-item Vakantie vraag 2, het wiskunde-item Skateboard vraag 2 en het wiskunde-item Dobbels-
stenen maken vraag 2.
- Andere vragen die hoog scoorden als het gaat om dit aspect zijn de drie probleem-oplos-vragen van de context Irrigatie, de probleemoplos-items Opzet van een cursus vraag 1, vraag 2 van Bibliotheekstelsel, Vakantie vraag 1 en Energiebehoefte en de wiskunde-items Berovingen, Skateboard vraag 1 en de twee vragen uit Lopen.

Bovengenoemde probleemoplos-opgaven doen een sterk beroep op de vaardigheden analyseren, redeneren, organiseren, controleren/verifiëren en evalueren. Deze vaardigheden treffen we ook aan in het *framework* dat de basis vormde bij de totstandkoming van de wiskunde-items van PISA.

Ook de probleemoplos-items Aansluitingen, Design by numbers vraag 3 en Vakantiekamp vraag 1 correleren in hoge mate met de wiskunde-opgaven Toetsresultaten, Skateboard vraag 3, Export vraag 2 en Timmerman. Nadere bestudering van deze items leidt tot enkele kanttekeningen. Vraag 2 van de probleemoplos-opgave Vakantie is een opgave van hoog niveau: de leerling moet de situatie doorgronden en vervolgens rekening houden met alle gegevens en randvoorwaarden.

Analyse van de resultaten bracht aan het licht dat, net als bij wiskunde, ook hier de score van Vlaanderen iets hoger is dan die van Nederland. Verder blijft de score van Duitsland een aardig stukje achter. We vinden dat deze opgave ook wel een plaats zou kunnen hebben in de wiskundeset. Wiskunde-technisch gezien is er in feite weinig aan de hand. Je moet een tabel kunnen aflezen en vaststellen of een afgelezen waarde al dan niet groter is dan 300.

De wiskunde-opgave Skateboard vraag 2, vraagt naar de hoeveelheid verschillende boards die je kunt samenstellen. Daartoe moet de leerling een schema lezen en doorgronden, de beschreven situatie begrijpen en tenslotte een telprobleem oplossen (het rekenen zelf $3 \times 2 \times 2 \times 1$ is elementair). Het is een meerkeuzevraag, dus het denkgedrag van de leerling is niet aan een uitwerking af te lezen. Ook vraag 1 van deze opgave doet een beroep op begrip van en inzicht in de aangeboden informatie. Dit item is, vergeleken met Vakantie vraag 2, wel wat makkelijker.

De wiskunde-opgave Dobbels-
stenen maken vraag 2 vraagt te beslissen welke uitslagen van een dobbelsteen van een echte dobbelsteen zijn (waarbij gegeven is dat de som van ieder tweetal tegenoverliggende ogen steeds zeven is). De moeilijkheid zit hier in de voorstelling die een leerling zich van de echte kubus moet maken om te beslissen welke vlakken tegenoverliggend zijn op de dobbelsteen. Afgezien van dit abstracte ruimtelijke inzicht worden hier geen andere wiskundige vaardigheden vereist. Het lijkt een heel ander probleem dan de twee hierboven.

Het wiskunde-item Berovingen vereist het inzicht dat hier 'statistisch gelogen' wordt. Het gaat om het inzicht dat de suggestieve presentatie in feite probeert een groot verschil te tonen waar het slechts om een toename van 508 naar 515 gaat.

Die andere genoemde probleemoplosvragen vertonen geen consistent beeld in hun relatie tot wiskunde. Zo is bij Irrigatie vraag 1 de score in Nederland iets hoger dan die in Vlaanderen en Duitsland veel zwakker. Bij vraag 2 van deze contexten wijken deze 3 landen echter bijna niet van elkaar af. En in vraag 3 – waar in een tabel steeds 'open' of 'dicht' moet worden ingevuld – doet Nederland het veel slechter dan gemiddeld. En dat terwijl alle drie vragen flink scoren op de factor analyseren.

Het herkennen van wiskunde-elementen in items probleemoplossen

Tabel 3 Overzicht items probleemoplossen

<i>naam unit</i>	<i>vraagtype</i>
Besluitvormingsitems	
Energiebehoefte vraag 1	Kort antwoord
Energiebehoefte vraag 2	Lang antwoord
Naar de film vraag 1	Meerkeuze
Naar de film vraag 2	Meerkeuze
Vakantie vraag 1	Kort antwoord
Vakantie vraag 2	Lang antwoord
Aansluitingen	Lang antwoord
Systeemanalyse- en ontwerpitems	
Bibliotheeksysteem vraag 1	Kort antwoord
Bibliotheeksysteem vraag 2	Lang antwoord
Design by numbers vraag 1	Meerkeuze
Design by numbers vraag 2	Meerkeuze
Design by numbers vraag 3	Lang antwoord
Opzet van een cursus	Lang antwoord
Vakantiekamp	Lang antwoord
Trouble shooting-items	
Irrigatie vraag 1	Lang antwoord
Irrigatie vraag 2	Meerkeuze
Irrigatie vraag 3	Lang antwoord
Diepvriezer vraag 1	Meerkeuze
Diepvriezer vraag 2	Meerkeuze

Bovenstaande uitkomsten over de samenhang tussen wiskunde- en probleemoplosvragen negerend, hebben we door alleen maar naar de vragen zelf te kijken geprobeerd uit de probleemoplos-items die items te selecteren die in de ogen van de bij dit rapport betrokken experts een duidelijke verwantschap met wiskunde vertonen. De complete lijst met items, ingedeeld in de drie typen, staat in onderstaande tabel.

Niet alle hierboven opgenomen items waren in de ogen van de betrokkenen direct wiskundig van aard. Selectie op dit aspect leverde de lijst op zoals gepresenteerd in tabel 4. In deze tabel treft men behalve de naam van de betreffende items ook informatie aan over het wiskundedomein waar het betreffende item in onze ogen het meest thuis hoort.

Tabel 4 Wiskunde-achtige probleemoplos-items

<i>naam unit</i>	<i>toelichting/opmerking</i>	<i>domein</i>
Besluitvormingsitems		
Energiebehoefte vraag 1	aflezen in combinatietabellen + sommeren	Hoeveelheid
Energiebehoefte vraag 2	aflezen in combinatietabellen, sommeren, vergelijken met gegeven voorwaarden, conclusie trekken	Hoeveelheid
Vakantie vraag 1	routelengtes narekenen aan de hand van kaartje en tabel en kortste route kiezen	Hoeveelheid
Vakantie vraag 2	reisschema opstellen uitgaande van 3 voorwaarden waarvan 1 expliciet numeriek	Hoeveelheid
Aansluitingen	optimale route in termen van tijd en kosten op basis van kaartje en elementaire kwantitatieve info	Hoeveelheid
Systeemanalyse- en ontwerpitems		
Design by Numbers vraag 1	correct computercommando koppelen aan gegeven figuur	Veranderingen en Relaties
Design by Numbers vraag 2	ietwat complexer computercommando koppelen aan gegeven structuur	Veranderingen en Relaties
Design by Numbers vraag 3	zelf commando's binnen verstrekte commandostructuur verzinnen voor gegeven figuur	Veranderingen en Relaties
Vakantiekamp	slaapkamerooster opstellen op basis van o.a. enkele numerieke voorwaarden	Hoeveelheid

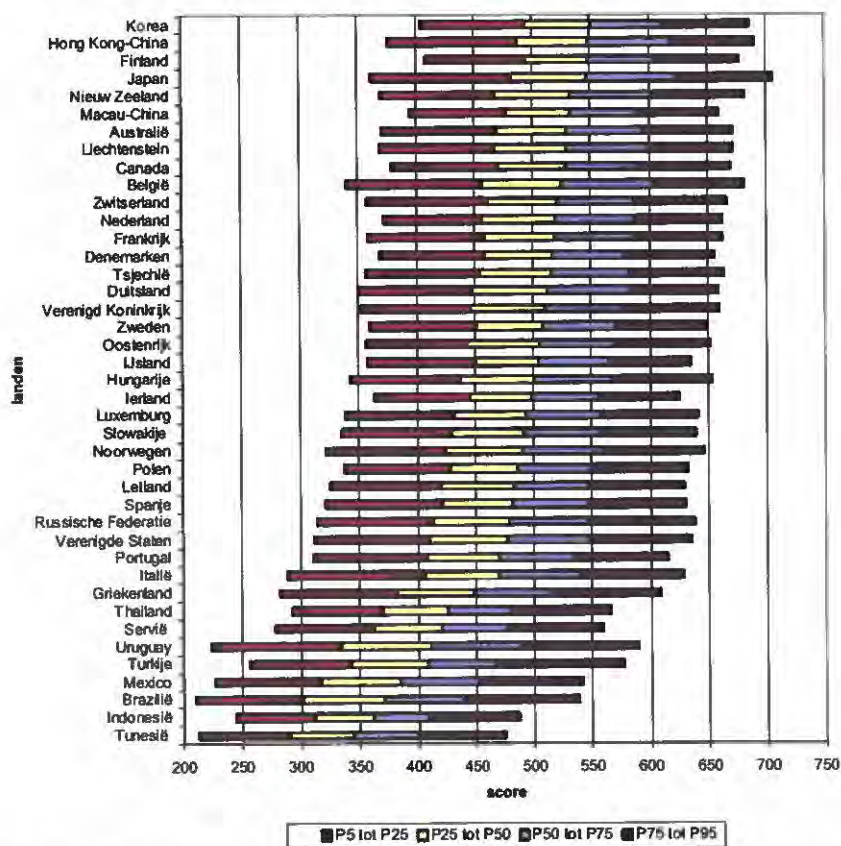
De niet in deze tabel opgenomen items zijn in onze ogen duidelijk minder wiskundig van aard. Opvallend is dat er uit de categorie *trouble shooting* geen enkele vraag door de auteurs als wiskundig werd geclassificeerd. Dat neemt echter niet weg dat ieder van ons ervan overtuigd is dat vaardigheden die in de opgaven van de categorie *trouble shooting* voor een correct antwoord nodig zijn, ook een rol kunnen spelen bij wiskunde. Bij deze classificatie stond ons het Nederlandse wiskundecurriculum voor 15-jarigen voor ogen en constateerden we dat items als die van de categorie *trouble shooting* daarin niet onder te brengen zijn.

Een internationale vergelijking van scores op meer of minder wiskundig getinte items probleemoplossen

We zijn gaan onderzoeken wat die verwantschap met wiskunde van de items in tabel 4 betekent voor de wijze waarop Nederlandse kinderen in verhouding tot de kinderen in de landen (Duitsland, Denemarken, Finland, Vlaanderen, Noorwegen, Zweden en de VS) die in deze studie vergeleken worden met Nederland de betreffende items uitvoeren.

In eerste instantie vergeleken we de Nederlandse scores op deze en de andere probleem-oplos-items met de scores van de Deense kinderen.

De reden voor een directe vergelijking met Denemarken is gelegen in het feit dat Nederland en Denemarken in de internationale vergelijking bij probleemoplossen heel erg dicht bij elkaar in de buurt blijken te liggen. Zie in dat verband weer de figuur afkomstig uit het Nederlandse PISA 2003-rapport die ook hieronder (figuur 12) staat afgebeeld (dit is dezelfde grafiek als figuur 2).



figuur 12 Scoreverdeling op de schaal van probleemoplossen in de OESO- en partnerlanden

We zien in de vorige figuur dat Nederland, op plaats 12, en Denemarken, op plaats 14, heel erg veel overeenkomen: het verschil tussen het gemiddelde van elk van beide landen is nog geen 5 punten op de PISA-schaal en ook de breedte van de middelste 90 procent gemeten op de PISA-schaal is nagenoeg gelijk bij beide landen. Op grond van deze gegevens lijkt de veronderstelling gerechtvaardigd dat Deense en Nederlandse kinderen vergelijkbaar vaardig zijn als het gaat om probleemoplossen. Als we de probleemoplos-items echter clusteren op basis van de hierboven gemaakte tweedeling, zien we een opvallend verschil:

Tabel 5 Denemarken en Nederland vergeleken wat betreft al dan niet wiskundige probleemoplos-items

	DK>NL	DK<NL	totaal
wiskundige probleemoplos-items	2	7	9
niet-wiskundige probleemoplos-items	6	4	10
totaal	8	11	19

Deze tabel is het product van de analyse bij welke items Nederlandse kinderen beter/slechter scoorden dan de Deense kinderen door te kijken naar het percentage leerlingen dat bij het betreffende item geen punten (score 0 of score 9) behaalde. Zo scoorde bijvoorbeeld, zo blijkt uit de internationale analyse, op Design by Numbers vraag 1 in Denemarken 47,9% van de leerlingen geen punten terwijl 46,5% van de Nederlandse kinderen hier geen punten behaalde. Design by Numbers vraag 1 werd als een wiskunde-achtig probleemoplos-item beschouwd en zodoende is dit item een van de zeven items in bovenstaande tabel waarbij Nederland beter scoorde dan Denemarken. Tabel 6 geeft iets meer in detail informatie over deze vergelijking.

Tabel 6 Details rond vergelijking Denemarken en Nederland

naam item	wiskunde?	%-nulscore Denemarken	%-nulscore Nederland
Bibliotheeksysteem vraag 1		23,5	26,2
Bibliotheeksysteem vraag 2		84,5	73,5
Design by Numbers vraag 1	wiskunde	47,9	46,5
Design by Numbers vraag 2	wiskunde	49,3	48,3
Design by Numbers vraag 3	wiskunde	58,6	54,5
Opzet van een cursus		55,4	60,1
Aansluitingen	wiskunde	53,6	53,0
Vakantiekamp	wiskunde	33,7	38,5
Diepvriezer vraag 1		46,6	44,0
Diepvriezer vraag 2		43,1	49,1
Energiebehoefte vraag 1	wiskunde	10,3	11,7
Energiebehoefte vraag 2	wiskunde	56,8	53,3
Naar de film vraag 1		12,4	17,9
Naar de film vraag 2		25,1	28,4
Vakantie vraag 1	wiskunde	54,5	45,6
Vakantie vraag 2	wiskunde	54,6	54,2
Irrigatie vraag 1		32,0	26,0
Irrigatie vraag 2		51,5	40,7
Irrigatie vraag 3		58,6	44,2

In tabel 6 zagen we dat Nederland en Denemarken bij probleemoplossen eenzelfde gemiddelde score hadden, maar op items verschillend scoorden. We kunnen de Nederlandse scores ook met een wat groter aantal landen vergelijken. Daarbij hebben we per vraag gekeken naar de plaats van Nederland tussen Duitsland, Denemarken, Finland, Vlaanderen, Noorwegen, Zweden en de VS als we bij ieder van de probleemoplos-items een rangorde bepalen op grond van de gemiddelde scores. Hierbij baseerden we ons ook weer op dezelfde percentages als bij de vergelijking met Denemarken. Daarna hebben we de verzameling items weer gesplitst in wiskundige probleemoplos-items en niet-wiskundige probleemoplos-items.

Tabel 7 Vergelijking Nederland – nabije buitenland op al dan niet wiskundige probleemoplos-items

naam item	wiskunde?	%-geen score Nederland	rangorde Nederland
Design by Numbers vraag 1	wiskunde	46,5	3
Design by Numbers vraag 2	wiskunde	48,3	2
Design by Numbers vraag 3	wiskunde	54,5	4
Aansluitingen	wiskunde	53,0	3
Vakantiekamp	wiskunde	38,5	5
Energiebehoefte vraag 1	wiskunde	11,7	3
Energiebehoefte vraag 2	wiskunde	53,3	3
Vakantie vraag 1	wiskunde	45,6	3
Vakantie vraag 2	wiskunde	54,2	3
Bibliotheeksysteem vraag 1		26,2	6
Bibliotheeksysteem vraag 2		73,5	2
Opzet van een cursus		60,1	3
Diepvriezer vraag 1		44,0	6
Diepvriezer vraag 2		49,1	5
Naar de film vraag 1		17,9	6
Naar de film vraag 2		28,4	3
Irrigatie vraag 1		26,0	1
Irrigatie vraag 2		40,7	3
Irrigatie vraag 3		44,2	5

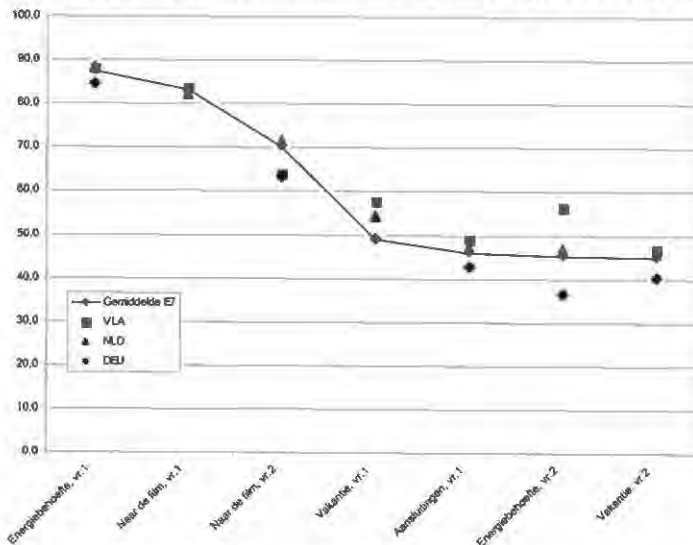
We zien in tabel 7 dat Nederland bij de wiskundig getinte probleemoplos-vragen een vrij stabiele positie inneemt rond de derde plaats. Bij de niet-wiskundige probleemoplos-vragen is het beeld sterk wisselend, waarbij nogal eens een van de laagste posities wordt ingenomen.

Een en ander leidt tot de veronderstelling dat Nederlandse leerlingen weliswaar behoorlijk goed worden opgeleid als het gaat om de wiskundige vaardigheden die bij PISA aan de orde komen, maar aanzienlijk minder goed uit de verf komen als het andere, minder wiskundige probleemoplossende vaardigheden betreft. Hierover verder speculerend is het misschien niet zo gek om te veronderstellen dat dit laatste in feite een tekortkoming van het Nederlandse onderwijs betekent. Nederlandse kinderen kunnen de wiskundeopgaven behoorlijk goed beantwoorden omdat ze daar een heel gerichte opleiding in genoten hebben, veel gericht vermoedelijk dan het merendeel van de 15-jarigen in andere landen. Maar de daadwerkelijke, abstractere probleem aanpak wordt juist door Nederlandse kinderen veel slechter beheerst. Terwijl we vermoedelijk juist willen bereiken, door probleemgericht te werken aan wiskundige geletterdheid, dat dit een automatische transfer naar minder wiskundig gerelateerde problemen als gevolg zou hebben. Dat blijkt, zo vertellen ons bovenstaande PISA 2003-gegevens, niet of nauwelijks het geval te zijn.

Vergelijking Nederland – Duitsland/Vlaanderen

In de figuren 13, 14 en 15 zijn scores op de probleemoplos-items afgebeeld, onderverdeeld naar de drie typen. In elke grafiek zijn de bijbehorende vragen afgebeeld in opklimmende moeilijkheid. Daarbij is uitsluitend gerekend met de scores in zeven Europese landen, waaronder Nederland (E7). In de figuren zijn bij elke vraag ook de gemiddelde scores afgebeeld van de leerlingen uit Vlaanderen, Duitsland en Nederland. We weten dat ook bij probleemoplossen Vlaanderen van deze drie het best presteerde. Nederland deed het iets beter dan Duitsland, maar het verschil is klein, zeker als je dat vergelijkt met het verschil tussen deze twee landen bij wiskunde.

Laten we eens wat beter naar de scores per vraag kijken. Bij de opgaven die vallen onder Besluitvorming (zie daarvoor ook tabel 3) scoren de Nederlandse leerlingen niet zo slecht (zie fig. 13). Daar zit niet de voornaamste oorzaak van de lage score voor Nederland.

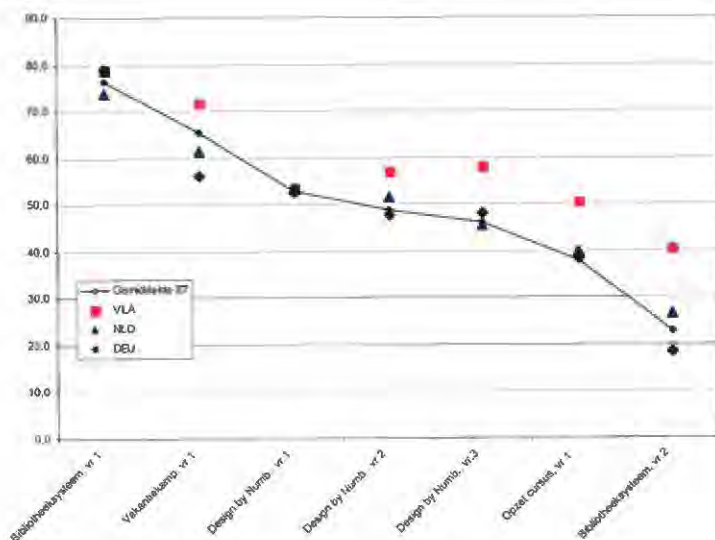


figuur 13 Scores Besluitvorming items

Systeemanalyse- en ontwerpitems liggen de Nederlandse leerlingen wat minder (zie fig. 14). Bij Bibliotheekstelsysteem worden de regels gepresenteerd in een stroomschema en in tekst. De eerste vraag kan met die uitleenregels vrij snel worden beantwoord. Toch scoren de Nederlandse leerlingen hier ver onder de maat. De tweede vraag over dit Bibliotheekstelsysteem is de moeilijkste van alle probleemoplos-vragen. Ook bij deze vraag, waarin een vrij omvangrijk stroomschema moet worden ontworpen, doet Nederland het in vergelijking met Vlaanderen niet goed.

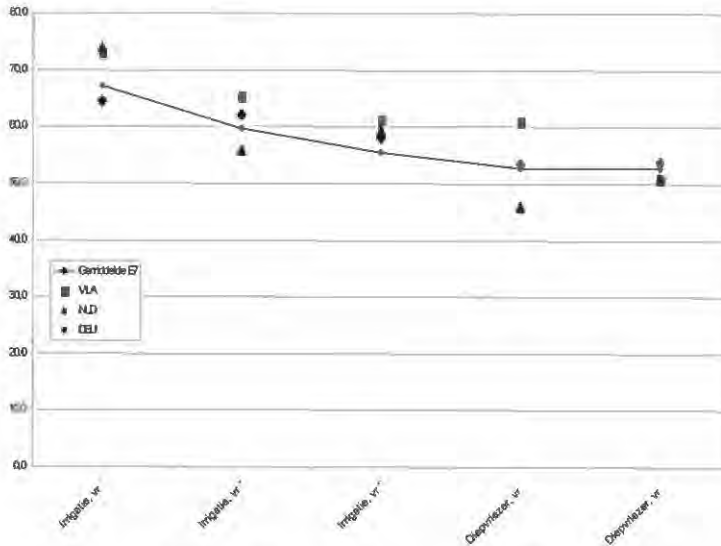
Ook de opgave Opzet van een cursus valt onder Besluitvorming. Alweer moet op basis van een aantal regels en voorwaarden een schema worden ontworpen. Opnieuw blijft de score van Nederland flink achter.

Deze vragen zijn tamelijk tijdrovend en vereisen geen rekenactiviteit of wiskundige bewerkingen. Bij de opgave Vakantiekamp moet ook flink georganiseerd worden, maar hier speelt ook tellen/rekenen een rol. Ook hier doen Nederlandse leerlingen het niet goed.



figuur 14 Scores systeemanalyse en -ontwerp items

Onder de noemer trouble shooting vallen de opgaven Irrigatie en Diepvriezer. Ook deze opgaven zijn niet kwantitatief van aard, en passen dan ook eigenlijk niet bij de wiskundeopgaven. Anderzijds worden er wel denkactiviteiten vereist zoals redeneren en concluderen, die we in het PISA-onderzoek ook bij de wiskundige geletterdheid onderbrengen. Het lijkt wel of de niet aan wiskunde verwante setting een nadelig effect heeft op de oplosvaardigheid van de Nederlandse leerlingen. Bij Irrigatie gaan de eerste twee vragen redelijk goed en gaat het pas bij de derde vraag echt mis. Bij Diepvriezer blijven de Nederlandse scores al bij de eerste vraag echt achter (zie fig. 15). Er zijn echter ook wel lichtpuntjes. Zo steekt Nederland zowel Vlaanderen als Duitsland duidelijk de loef af bij de tweede vraag van Naar de film (zie fig. 13). Terwijl dit ook een vraag is waar een conclusie moet worden getrokken op basis van een veelheid aan gegevens. De score bij de eerste vraag van deze opgave blijft iets achter; wellicht omdat de eerste vraag zoveel meer werk is?



figuur 15 Scores trouble shooting items

Bij vragen waarin de verwantschap met wiskunde duidelijker is, doen de Nederlandse leerlingen het meestal wat beter. Dat zien we bijvoorbeeld bij Energiebehoefte, met name bij de eenvoudige eerste vraag. Maar ook zo'n vraag had best een plaats mogen hebben tussen de wiskundeopgaven.

Toch blijft het beeld bij probleemoplossen, ook bij de meer wiskundig getinte opgaven minder fraai dan bij wiskunde. Zijn onze leerlingen misschien toch niet zo flexibel dat ze hun redeneer- en probleemoplosvaardigheden ook in minder herkenbare situaties kunnen inzetten?

Vergelijking Nederland – Duitsland/Vlaanderen, rekening houdend met schooltype

We hebben geprobeerd meer zicht te krijgen op de prestaties bij probleemoplossen door ook op een geheel andere wijze naar de verschillende PISA-gegevens te kijken. We hebben daartoe Nederland, Duitsland en Vlaanderen bekeken op drie schooltypen I, II en III. Onder schooltype I vallen in Nederland de leerlingen van vmbo BB/KB, schooltype II betreft vmbo GL/TL en schooltype III staat voor havo/vwo. De leerlingen uit Duitsland en Vlaanderen zijn op vergelijkbare wijze op hun onderwijsniveau geordend. Per schooltype is daarna gekeken naar drie verschillende opgaven wiskunde en drie opgaven probleemoplossen, uitgezocht op vergelijkbare PISA-score. We zagen in het merendeel van de gevallen dat het Nederlandse wiskundeonderwijs een grotere positieve invloed heeft naarmate het schoolniveau hoger is. In de volgende tabellen (tabel 8, 9 en 10) zijn specifiekere gegevens te vinden.

Tabel 8

item	Type I	Type II	Type III		PISA-score
M144q01 Verven van een kubus	50,6	65,9	87,6	VLA	497
	77,7	83,6	91,9	NL	
	49,6	60,1	76,4	DEU	
X603q01 Irrigatie	46,6	72,3	85,9	VLA	497
	54,4	72,8	91,0	NL	
	39,9	65,4	86,6	DEU	

Tabel 9

item	Type I	Type II	Type III		PISA-score
520q02 Skateboard	12,8	43,0	74,8	VLA	570
	19,3	53,9	80,7	NL	
	21,6	40,0	70,2	DEU	
X602q01 Vakantie	29,8	52,5	74,3	VLA	570
	31,9	53,8	72,0	NL	
	27,0	48,2	70,6	DEU	

Tabel 10

item	Type I	Type II	Type III		PISA-score
M704q02 Steun voor de president	5,3	39,5	65,5	VLA	657
	3,7	26,3	59,5	NL	
	7,4	18,9	44,9	DEU	
X402q02 Bibliotheek	5,0	30,6	63,3	VLA	658
	3,0	15,2	52,6	NL	
	1,8	11,0	42,0	DEU	

In de eerste van bovenstaande tabellen (tabel 8) zien we gegevens rond een tweetal gemiddeld moeilijke items van wiskunde respectievelijk probleemoplossen. Beide opgaven scoren op de PISA-schaal met een waarde van 497 heel dicht bij het OESO-gemiddelde, zijnde 500. We zien dat van de leerlingen van onderwijstype I (vmbo BB/KB en de internationale pendanten daarvan) in Duitsland 49,6% de betreffende opgave M144q01 goed beantwoordt, waar in Nederland 77,7% van dit type onderwijs de vraag goed weet te beantwoorden. Duidelijk is dat Nederlandse type I-leerlingen hier een stuk beter scoren: het verschil bedraagt $77,7 - 49,6 = 28,1\%$ in het 'voordeel' van

Nederland. Het verschil bij deze vraag tussen Nederland en Duitsland bij type III-leerlingen (havo/vwo of daarmee vergelijkbaar) is $91,9 - 76,4 = 15,5$. Als we kijken naar de verschillen tussen dezelfde categorieën bij de probleemoplos-opgave zien we het volgende: bij type I is dat verschil $54,4 - 39,9 = 14,5$ en bij type III: $91,0 - 86,6 = 4,4$.

Bij tabel 9 kijken we eveneens naar de verschillen:

Het wiskunde-item M520q02 levert het volgende: verschil type I: $19,3 - 21,6 = -2,3$ en verschil type III: $80,7 - 70,2 = 10,5$.

Bij het probleemoplos-item X602q01 vinden we: verschil type I: $31,9 - 27,0 = 4,9$ en verschil type III: $72,0 - 70,6 = 1,4$

En bij tabel 10 levert dit op:

M704q02: verschil type I: $-3,7$; verschil type III: $14,6$

X402q02: verschil type I: $1,2$; verschil type III: $10,6$.

Over het algemeen geven ook deze uitkomsten een bevestiging van het reeds eerder signaleerde verschijnsel: met name de beter/hoger opgeleide Nederlandse kinderen profiteren van het Nederlandse wiskundeonderwijs. Hoe hoger het opleidingsniveau van de leerling, hoe meer die leerling zich in positieve zin wiskundig onderscheidt van zijn Duitse collega-leerling op een vergelijkbaar type onderwijs. Maar omgekeerd kun je ook zeggen dat hoe basaler de leerling is opgeleid, hoe geringer deze leerling zich in wiskundige zin weet te onderscheiden van vergelijkbare leerlingen uit andere landen. Ook bij vergelijking van de gegevens uit bovenstaande tabellen tussen Vlaamse en Nederlandse type I-, II- en III-leerlingen wordt dit beeld bevestigd.

Conclusies/implicaties

Omdat de scores bij zowel wiskunde als probleemoplossen zijn gezet op een gemiddelde van 500 met een standaardafwijking van 100, kunnen we stellen dat gelijke scores in beide onderwerpen ook een zelfde – relatieve – prestatie weergeven.

Nederland heeft bij wiskunde een gemiddelde score van 538 en bij probleemoplossen 520. In het rapport over probleemoplossen wordt gesteld dat een relatief hogere gemiddelde score voor wiskunde dan voor probleemoplossen aangeeft dat in dit land het wiskundeonderwijs behoorlijk effectief is geweest (in het perspectief van de wiskunde die door PISA wordt gemeten).

Dat Nederland beter scoort in wiskunde (of anders gezegd: achterblijft bij probleemoplossen) kan dus op deze manier worden verklaard. Dat kan als een compliment worden opgevat voor het wiskundeonderwijs in Nederland. Maar het is wellicht ook verklaarbaar als je kijkt naar de mate waarin het Nederlandse wiskundeonderwijs is afgestemd op PISA-items. We durven wel de uitspraak aan dat er geen land ter wereld is waar het wiskundeonderwijs zo goed is afgestemd op de vaardigheden die in het PISA-framework worden onderscheiden en die gezamenlijk de wiskundige geletterdheid beschrijven. Mocht men in andere landen niet tevreden zijn met de PISA-score bij wiskunde en men zou die score middels een verandering in het wiskundeonderwijs teweeg willen brengen, dan zou men er verstandig aan doen goed te kijken naar het Nederlandse wiskunde curriculum in de eerste drie klassen van het voortgezet onderwijs. Wanneer men daar het een en ander (of alles) van zou overnemen voor het onderwijs in eigen land, dan zouden de PISA-scores ongetwijfeld vrij snel omhoog

gaan. Landen als Hong Kong, China, Zuid-Korea en Japan hebben een vrij klassiek wiskundeprogramma en hun PISA-scores zijn heel hoog. Nederland zou door deze landen zeer waarschijnlijk flink worden afgetroefd op PISA-score wiskunde wanneer de curricula een grotere gelijkenis met de inhoud van het Nederlandse wiskundeonderwijs zouden vertonen.

Het rapport over probleemoplossen zegt dat de toets voor PISA-probleemoplossen los van schoolvakken staat. Er is in geen van de landen apart onderwijs in probleemoplossen. In schoolcurricula kom je aspecten van probleemoplossen tegen, maar dan ingebed in vakken als wiskunde en *science*. De prestatieverschillen tussen leerlingen bij probleemoplossen worden voor een niet gering deel verklaard door verschillen in wiskundige redeneervaardigheid. Wanneer de verschillen in score bij wiskunde en probleemoplossen aanzienlijk zijn, is het goed dit nader te bekijken.

Misschien is voor Nederlandse kinderen bij probleemoplossen dat analytisch redeneren zonder tegelijkertijd een beroep te doen op (uitgebreide) wiskundige technieken wel lastiger dan het redeneren in combinatie met de inzet van wiskundig gereedschap. Zou dat komen omdat er bij probleemoplossen op hoger niveau een beroep wordt gedaan op die analytische vaardigheid? En dat daarmee de relatieve zwakte van de leerlingen in echte algemene probleemoplosvaardigheden aan het licht komt?

Naar aanleiding van de diverse wijzen waarop we in dit hoofdstuk probleemoplossen en wiskunde met elkaar vergeleken hebben, veronderstellen we nog wat sterker dan voorheen dat de Nederlandse kinderen alleen maar zo hoog op de wiskundige geletterdheid ranglijst staan omdat ze, in tegenstelling tot hun leeftijdgenoten in de andere PISA-landen, bij wiskunde opgaven kregen voorgelegd die hen bekend (of vertrouwd) voorkomen. Dat maakt het voor hen meer een reproductieve bezigheid, terwijl het voor leerlingen uit andere landen veel meer nieuwe, onbekende, onverwachte, verrassende problemen waren die daardoor eigenlijk op voorhand al wat moeilijker voor hen waren. We moeten dus meer respect hebben voor de (scores van de) Japanse leerlingen dan voor de Nederlandse, ondanks het feit dat de gemiddelde score van Japan met 534 enkele punten lager was dan de Nederlandse.

Bovendien denken we argumenten te hebben voor de bewering dat met name de beter/hoger opgeleide Nederlandse kinderen profiteren van het Nederlandse wiskundeonderwijs. Hoe hoger het opleidingsniveau van de leerling hoe meer die leerling zich in positieve zin wiskundig onderscheidt van zijn Duitse collega-leerling op een vergelijkbaar type onderwijs.

Toch is het niet zo dat we het gevoel hebben het opvallende verschil in de positie van Nederland op de internationale rangordellijsten van wiskunde en probleemoplossen nu daadwerkelijk doorgrond te hebben. In dit hoofdstuk hebben we gezien dat een sluitende verklaring voor dit fenomeen tot op heden door ons niet gegeven kan worden.

6. Opmerkingen en Aanbevelingen naar aanleiding van dit Rapport

NEDERLAND INTERNATIONAAL

Vlaanderen onderscheidt zich door een hoge score, de hoogste van alle 'landen' die meedoen aan PISA. Vlaanderen dankt deze score vooral aan de opgaven uit het Reflectiecluster, en aan het feit dat er beter gescoord wordt bij de wat formelere en abstracte opgaven.

Nederland en de officiële winnaar Finland hebben een vergelijkbaar profiel: de score berust vooral op goede resultaten in Reproductiecluster. Finland doet het alleen overal net iets beter.

Opvallend is dat Nederland relatief achterblijft bij problemen die meer mathematisering vereisen en ook meer stappen vertonen in het oplossingsproces. Dit valt op te maken uit de resultaten van Nederland bij Probleemoplossen, en door die te relateren aan de deelresultaten bij Wiskunde. Met name middenmoter Duitsland zit Nederland hier dicht op de hielen.

Dat Probleemoplossen in Nederland een relatief zwak ontwikkeld gebied is blijkt ook uit een vergelijking van de scores voor Wiskunde en Probleemoplossen: in geen van de deelnemende landen is het verschil zo groot als in Nederland.

Duitsland heeft een vergelijkbaar profiel als de Scandinavische landen Noorwegen, Zweden en Denemarken. Finland (zoals eerder opgemerkt) en Nederland hebben een vergelijkbaar profiel dat zich daarvan onderscheidt.

De Verenigde Staten laten zich slecht met de Europese noordwesthoek vergelijken. De Verenigde Staten blijven op alle fronten achter, waarbij opvalt dat met name de prestatie in het Reproductiecluster achter blijft.

Nederlandse vmbo-leerlingen in de beroepsgerichte leerwegen (de zwakste groep), doen het relatief goed in vergelijking met Vlaanderen en Duitsland.

Dat Nederland goed scoort bij Wiskunde kan deels verklaard worden uit het feit dat het curriculum in Nederland meer past bij de Wiskundige geletterdheid uit PISA dan dat van veel andere landen. Alhoewel daarbij opgemerkt kan worden dat Nederland ook goed scoort bij het TIMSS onderzoek dat juist curriculum georiënteerd is.

Het peil van wiskundige geletterdheid in Nederland lijkt te dalen. Dit moet met enig voorbehoud worden gezegd aangezien de scores in 2000 en 2003 om diverse redenen (omvang en representativiteit van de Nederlandse steekproef in 2000, veranderingen in het toetsdesign) niet zomaar te vergelijken zijn. De uitkomsten van PISA 2006 zullen ons daarover betrouwbaarder informeren.

NEDERLAND ALGEMEEN

Nederlandse leerlingen blinken uit in het lezen en interpreteren van grafieken en andere grafische voorstellingen. Ook het redeneren op basis van die gegevens gaat de leerlingen goed af.

Nederlandse leerlingen blijven achter wanneer er formules gemaakt of geïnterpreteerd moesten worden of, in het algemeen, wanneer de benodigde wiskunde wat formeler wordt.

Bij problemen met veel informatie, met name in de vorm van tekst, hebben Nederlandse leerlingen relatief veel moeite. Vooral als er meer stappen nodig zijn, en meer mathematisering. Dit is niet het gevolg van het niet aandurven van het probleem: Nederlandse leerlingen beginnen aan bijna iedere opgave, in tegenstelling tot leerlingen uit andere landen.

NEDERLAND vmbo

De Nederlandse vmbo-leerlingen in de beroepsgerichte leerwegen doen het relatief, in vergelijking met het buitenland, goed.

De wiskundige geletterdheid van de Nederlandse vmbo-leerling volgens de PISA-norm geeft aanleiding tot bezorgdheid: slechts weinigen halen het niveau 4, dat toch gezien mag worden als het doel van wiskundige geletterdheid voor een leerling van 15 jaar oud.

Naast het goed omgaan met grafieken en het achterblijven bij wat formelere wiskunde (zie eerder) valt bij vmbo-leerlingen op dat het verwoorden van oplossingen niet eenvoudig is.

AANBEVELINGEN

Er dient *meer* aandacht te komen voor de meer *formele* en *abstracte* aspecten van de wiskunde, waarbij de stand van zaken in Vlaanderen – en niet in Finland – een voorbeeld kan zijn. Dit geldt vooral voor de havo en vwo leerling.

Er dient *meer* aandacht te komen voor authentiek *probleemoplossen* (binnen de wiskunde); over de gehele breedte van het Nederlandse onderwijs blijven de leerlingen achter: meerstapsproblemen en mathematiseren leveren de leerling te veel moeilijkheden op. Hierbij zijn Korea en Finland goede voorbeelden.

Er dient meer *uitdagend en authentiek probleem georiënteerd* wiskundeonderwijs te komen op het vmbo. Dit omdat het niet verantwoord is binnen de Nederlandse context dat zoveel vmbo-leerlingen het geletterdheidsniveau 4 niet halen. Daarbij dient met name ook *het taalaspect* van deze wiskunde aan de orde te komen: het lezen en interpreteren van een tekst en visuele informatie en het verwoorden en in beeld brengen van de oplossing.

Er dient nader onderzoek te komen naar het feit dat het *niveau* van het Nederlandse wiskundeonderwijs *lijkt* te dalen als we de gangbare internationale onderzoeken PISA en TIMSS als maatlat hanteren. Daarbij lijkt het verstandig het niet te laten bij een verwijzing naar de teloorgang van de staartdeling.

Bijlage

Overzicht van de codering; vrijgegeven opgaven na PISA 2003

Item code	Naam van de opgave
M124Q01	P2000 Lopen vraag 1
M124Q03	P2000 Lopen vraag 2
M145Q1	P2000 Dobbelstenen
M150Q01	P2000 Jongeren steeds langer, vraag 1
M150Q03	P2000 Jongeren steeds langer, vraag 2
M150Q02	P2000 Jongeren steeds langer, vraag 3
M179Q01	P2000 Roofovervallen
M266Q01	P2000 Timmerman
M402Q01	Chatten via internet, vraag 1
M402Q02	Chatten via internet, vraag 2
M413Q01	Wisselkoers, vraag 1
M413Q02	Wisselkoers, vraag 2
M413Q03	Wisselkoers, vraag 3
M438Q01	Export, vraag 1
M438Q02	Export, vraag 2
M467Q01	Gekleurde snoepjes
M468Q01	Natuurkundeproefwerk
M484Q01	Boekenrekjes
M505Q01	Afval
M509Q01	Aardbeving
M510Q01	Keuzes
M513Q01	Toetsresultaten
M520Q01	Skateboard, vraag 1
M520Q02	Skateboard, vraag 2
M520Q03	Skateboard, vraag 3
M547Q01	Trap
M555Q02	Dobbelstenen maken, vraag 2
M702Q01	Steun voor de president

M704Q01	De beste auto, vraag 1
M704Q02	De beste auto, vraag 2
M806Q01	Trappatroon
Alleen field test	
M471Q01	Voorjaarskermis
M515Q01	Kinderschoenen
M521Q01	Tafeltennis-toernooi
M525Q01	C02-uitstoot verminderen, vraag 1
M525Q02	C02-uitstoot verminderen, vraag 2
M525Q03	C02-uitstoot verminderen, vraag 3
M543Q03	Ruimtevlucht
M555Q01	Dobbelstenen maken, vraag 1

Literatuur

- Baye, A., I. Demonty, A. Fagnant, A. Matoul en Chr. Monseur (2003). *Quels défis pour notre système éducatif*. Luik: Universiteit van Luik.
- Boer, C.J.E.M van den (2003). *Als je begrijpt wat ik bedoel - Een zoektocht naar verklaringen voor achterblijvende prestaties van alloctone leerlingen in het wiskundeonderwijs*. Utrecht CD-B press
- Dekker, T. (2005). *De Nederlandse PISA-resultaten nader bekeken*. Nieuwe Wiskrant jrg 25 (2), pp. 5-8.
- Dekker, T en M. Wijers: *Wiskundeniveau van vijftienjarigen mag nog wel wat hoger*. Het Onderwijsblad nr 22, p. (december 2005).
- Lagerwaard, K. en G. Limpens: PISA 2003', en
- Koch, J. en C. Neuman (2005). *Olympiade ohne Verlierer*. Der Spiegel 29 pp. 46-47
- Lange de J. (2005): *Wiskunde om gecijferd van te worden, deel 1*. Nieuwe Wiskrant jrg 24 (3) pp. 42-48.
- Lange de, J. (2005): *Wiskunde om gecijferd van te worden, deel 11*. Nieuwe Wiskrant jrg 25 (2), pp. 9-14.
- Meyer, Inge de, Jan Pauly en Luc Van de Poele (2004). *Leren voor de problemen van morgen - de eerste resultaten van PISA2003*. Gent: Universiteit Gent.
- Meyer, Inge de en Luc Van de Poele (2004). *Leerstijlen en onderwijsstrategieën in PISA2000*. Gent: Universiteit Gent .
- Wijers, M. M. (2005). *Mathematical literacy in PISA: 'the Netherlands' point of view*. Presentatie EARLI 2005: Nicosia.
- Wijnstra, J.M. (2001). *Bruikbare kennis en vaardigheden voor jonge mensen*. Nederlandse uitkomsten van het OESO Programme for International Student Assessment (PISA) op het gebied van begrijpend en studerend lezen, wiskunde en de natuurwetenschappelijke vakken in het jaar 2000. Arnhem: Citogroep.
- Cito groep en Johan M. Wijnstra (2000). *Bruikbare kennis en vaardigheden voor jonge mensen*. Arnhem: Citogroep.
- Cito groep (2004). *Resultaten PISA-2003 - praktische kennis en vaardigheden van 15 jarigen*. Arnhem: Cito groep
- OECD (1999). *Measuring Student Knowledge and Skills - A new framework for Assessment*. OECD programme for International Student Assessment. Paris: OECD
- OECD (2003). *Pisa 2003 Data Analysis Manuel- SPSS® Users*. Paris: OECD
- OECD (2003). *The Pisa 2003 Assessment Framework, Mathematics, reading, science and problem solving knowledge and skills*. Paris: OECD
- OECD (2004). *Learning for Tomorrow's World - First results from PISA 2003*. Paris: OECD