

DAT WORDT NOG WAT, MET DIE REKENTOETSEN!

De rekentoetsen-3F zijn niet valide

□ Ben Wilbrink, Joost Hulshof en Henk Pfaltzgraff

Een jaar geleden hebben Wilbrink en Hulshof in dit tijdschrift betoogd dat de voorbeeld-rekentoetsen weinig van doen hebben met het rekenen dat vrijwel iedereen voor ogen staat. Inmiddels is er wetgeving en zijn de resultaten van een eerste pilot met rekentoetsen bekend. Dit artikel analyseert of de rekentoetsen-3F voor havo, mbo en eventueel vwo - waarvan Cito een voorbeeldrekentoets beschikbaar heeft gesteld - valide zijn.

De validiteit van de rekentoetsen

Het rekenen heeft de warme belangstelling van de wetgever. Van po tot vwo en mbo komen er rekentoetsen bij de afsluiting van het onderwijs (zie kader). Voor een goed begrip van de huidige situatie dient de volgende schets van wat vooraf ging; zie Wilbrink & Hulshof (2011) voor details en verwijzingen. De commissie-Meijerink gaf referentieniveaus rekenen (en Nederlandse taal). De referentieniveaus werden wet en zodoende de basis voor aan de eindexamens toe te voegen rekentoetsen, hoewel op 31 maart 2010 (Handelingen TK) men individuele toetsing riskant vond. De departementale commissie-Schmidt (2011) presenteerde de rekentoets-

wijzers (voor 3F:RTW) en Cito nam pilottoetsen af. De eerste resultaten van de pilots vielen de minister tegen (zie tabel 1 met gegevens uit de brief van de minister van 8 juni 2012), maar de resultaten waren in lijn met wat leraren er eerder al van verwachtten.

Kritieken zoals die van Wilbrink en Hulshof, gaven de minister reden om althans voor het vwo alsnog een rekentoets-3S te laten uitwerken (commissie-Van de Craats) als mogelijk alternatief voor de rekentoets-3F, met in het najaar 2012 een veldraadpleging over 3F versus 3S. Kritiek op de referentieniveaus F is algemener dan alleen of het past bij het vwo: het 'functioneel rekenen' van de F-niveaus mag geen rekenen heten.

Voor de validiteit van deze rekentoetsen zijn vier zaken van belang die hierna worden besproken: in de controverse tussen psychologie en constructivisme staan deze rekentoetsen aan de constructivistische kant, er is een terugval in basale rekenvaardigheid maar die wordt juist niet getoetst, deze rekentoetsen sluiten niet aan op vervolgonderwijs, en het gebruik van de rekenmachine is contraproductief voor rekenvaardigheid.

Realistisch rekenen tegenover basale rekenvaardigheid

De controverse in het wiskundeonderwijs betreft het realistisch rekenen zoals dat herkenbaar is in de officiële kerndoelen en referentieniveaus: het is de Nederlandse variant van de reformdidactiek in het rekenonderwijs (Schoenfeld, 2004). Het realistisch rekenen is in Nederland ingevoerd zonder een empirisch fundament

Tabel 1. Resultaten pilot rekentoets vo - maart 2012

	2F vmbo-bb	2F vmbo-kb	2F vmbo-tl	3F Havo	3F Vwo
Gemiddelde cijfer	4,3	5,4	6,4	4,9	6,7
Percentage onvoldoendes	84%	56%	28%	72%	32%

Het 'functioneel rekenen' van de F-niveaus mag geen rekenen heten



De rekentoets in het voortgezet onderwijs

De rekentoets is in de artikelen 29 en 60 van de WVO geregeld als een zelfstandig onderdeel van de eindexamens. De rekentoets kan dus niet gezien worden als een centraal examen of een schoolexamen. De bepalingen die bijvoorbeeld voor de centrale examens gelden, zijn niet automatisch op de rekentoets van toepassing. Omdat de rekentoets centraal ontwikkeld wordt en er regels gesteld worden voor de afname, is de rekentoets wel enigszins te vergelijken met een centraal examen. Er zijn echter ook belangrijke verschillen: rekenen is geen eindexamenvak en er ligt dus geen eindexamenprogramma aan de toets ten grondslag dat voorafgaand aan een examen in een eindexamenvak wordt onderwezen. Rekenen kan als onderdeel in verschillende vakken worden aangeboden. De rekentoets kent ook geen schoolspecifieke component (schoolexamen): het cijfer voor de centraal ontwikkelde rekentoets is tevens het eindcijfer.

Bron: Ontwerpbesluit tot wijziging van het Eindexamenbesluit VO, het Staatsexamenbesluit VO en het Examen- en kwalificatiebesluit beroepsopleidingen. | 23-04-2012 | OCW, EL&I | PDF document | 53 pagina's | p.32. <http://goo.gl/mMbWD>

(KNAW commissie-Lenstra), terwijl evenmin het falen was aangetoond van conventioneel rekenonderwijs dat aansluit op wiskundeonderwijs (zoals Pfaltzgraff, 2009). Confrontaties binnen commissies blijven enigszins verborgen voor de buitenwereld, maar zie de repliek van Van den Heuvel-Panhuizen et al. (2009) op analyses van de resultaten van het Periodiek Peilingonderzoek van het Onderwijsniveau (PPON) door Hickendorff et al. (2009a) met dupliek Hickendorff et al. (2009b). Vergelijking van de resultaten van PPON 1997 met die van PPON 2004 laat zien dat leerlingen op de basisvaardigheden zeer veel slechter zijn gaan scoren. Die conclusie was juist de aanleiding om rekentoetsen in te willen voeren. De rekentoetswijzer (RTW) en dus ook de voorbeeldrekentoets (VRT) voor niveau 3F gaan echter voorbij aan deze PPON-resultaten. Het signaal dat de overheid - het College voor Examens (CvE) - hiermee afgeeft aan het vo en mbo is dat basale rekervaardigheid niet van belang is als voorbereiding op de rekentoetsen en voor aansluiting op het hoger onderwijs. Een representatief item uit de VRT laat zien dat het gaat om contextopgaven waar nauwelijks aan gerekend hoeft te worden, en dat mag dan nog met een rekenmachine (zie figuur 1).

Onderzoek over redactiesommen - want daar komen die contextopgaven toch op neer - is bij uitstek een opgave voor de cognitieve psychologie. En daar is dan ook uitstekend onderzoek te vinden, zoals dat van

Desoete, Roeyers & De Clerq (2003), een voorbeeldige illustratie hoe empirisch is te onderzoeken of leerlingen met contextopgaven wel doen wat ze in de realistische didactiek worden verondersteld te doen.



Figuur 1. Vraag 3F VRT-2

Uitwerking

Berekening met de rekenmachine: $(5,2 \times 6,5 + 3,2 \times 4,6) \times 1,05 \times 44,95$

Deze opgave kan met simpeler getallen zonder gebruik van de rekenmachine (1) of ook zonder context (2):

- Variant 1: $(5 \times 6 + 3 \times 4) \times 1,10 \times 50$
- Variant 2: bereken $50 \times (5 \times 6 + 3 \times 4)$

Pfaltzgraff (2012) geeft deze uitwerkingen bij alle 60 opgaven van de VRT.

Patrick legt parket in twee kamers:

Woonkamer: 5,20 bij 6,50 meter
Slaapkamer: 3,20 bij 4,60 meter

Hij neemt 5% extra parket.

Hoeveel euro moet Patrick betalen?

€





De contextvragen toetsen abstractievermogen, naast of in plaats van rekenen

Geen rekentoets, maar een intelligentietest!

De vele contextopgaven zijn opgezet volgens een schema in de RTW dat overigens bekend is als de heuristiek van Polya (1957, p. xiv) voor het oplossen van problemen: probleem snappen - oplossingsplan maken - plan uitvoeren - resultaat evalueren. In de RTW heet het echter geen probleemoplossen, maar domeinoverstijgende vaardigheden. Het label verandert natuurlijk niets aan de aard van de zaak. Probleemoplossen staat aan de top van de cognitieve vaardigheden (Ohlsson, 2011), dus niet aan de basis, of we die nu 3F noemen of niet. De bijzondere situatie bij de meeste contextopgaven-3F is bovendien dat de probleemsituatie geen rekenkundige situatie is, maar een denkbeeldige situatie uit het dagelijks leven. Bij Polya is de probleemsituatie altijd wiskundig en wat er moet worden gesnapt is wiskunde. Bij deze VRT moet de leerling niet het rekenen, maar de dagelijkse situatie snappen. Die situatie is bovendien nooit concreet, in tegenstelling tot wat voorstanders van contexten in het rekenonderwijs beweren, maar altijd denkbeeldig. Dat kan betekenen dat deze contextvragen in feite abstractievermogen toetsen, naast of in plaats van rekenen. Evenmin als 'objectieve' keuzevragen objectief zijn (Wilbrink, 1977), zijn 'concrete' contexten concreet. Zie ook onderzoek door Depaepe, De Corte en Verschaffel (2010) over het verschil tussen rekenkundige

context of alledaagse context, althans in de manier waarop leerkrachten met contextopgaven omgaan. Depaepe et al. maken evenwel geen verbinding met de cognitieve theorie over probleemoplossen (Ohlsson, 2011).

Overigens zijn de talige plagen die contexten kunnen treffen (Evers-Vermeul & Land, 2012), ook in de VRT terug te vinden. Plus fouten tegen het Nederlands - Patrick legt al parket dat hij nog moet kopen! - die ook niet in het voordeel van de leerlingen werken. De leerling die uitlegt dat deze vraag geen antwoord heeft, omdat Patrick het parket al aan het leggen is, heeft gelijk. Scriptieonderzoek van Van der Weegh (2005) laat zien dat taalproblemen de uitkomsten van rekentoetsen sterk kunnen beïnvloeden. Ondanks alle beperkingen van een enkel scriptieonderzoek (deze kreeg wel de NVO-scriptieprijs 2006) zijn die uitkomsten verontrustend.

Rekenen is een onderdeel van de wiskunde en in beginsel gaat het dan om een strak systeem van betekenissen, afspraken en regels. Een vraag als VRT-48 - Uit hoeveel blokjes bestaat dit bouwsel - is dus geen wiskunde: onzichtbare blokjes zijn niet telbaar, we weten niet eens of ze er zijn. Wiskunde is geen giswerk, wiskunde is precies. Natuurlijk zou de leerling veronderstellingen in kunnen voeren om bij VRT-48 tot een gesloten probleem te komen, maar bij deze digitaal af te nemen VRT telt alleen de uitkomst. Veel opgaven in de VRT bestaan uit een stapeling van deelopgaven. Het valt niet in te zien wat, op het niveau 3F, hiervan de bedoeling is anders dan het moeilijker maken van de opgaven om zo beter tussen leerlingen te differentiëren. Maar zo gaat de toets verschillen in intellectuele capaciteiten meten, zoals verschillen in capaciteit van het werkgeheugen. Zo levert de parketvraag onnodig een hogere mentale belasting op, wat niet bijdraagt aan zijn validiteit als rekenvraag.

Geen aansluiting van dit rekenen-3F op wiskunde

Opvallend afwezig in de VRT zijn breuken, zoals er ook geen vermenigvuldigen en delen wordt gevraagd. Terwijl breukenvaardigheid toch een voorwaarde is

Figuur 2. Vraag 3F VRT-48

Bij de rekentoetsen behoort het gebruik van een rekenmachine te worden verboden



voor vruchtbare deelname aan wiskundeonderwijs. Over die voorwaardelijkheid is nu ook empirisch onderzoek: Siegler et al. (2012) laten met data uit cohortstudies zien dat vaardigheid in breuken en in delen voorspellend is voor latere prestaties in wiskunde, ook na correctie voor onderliggende verschillen in intelligentie en gezinsachtergrond. Pas op: het gaat hier om vaardigheid in breukrekenen, niet om 'functioneel rekenen' met breuken.

Hier blijkt weer het belang van empirisch onderzoek naar de juistheid van stellige rekenopvattingen, in dit geval dat breuken van marginaal belang zouden zijn. De referentieniveaus zijn bedoeld om doorlopende leerlijnen te creëren, maar hier is een voorbeeld van een doodlopende leerlijn van po naar vo en nog eens van vo naar ho. Leerlingen zonder breukvaardigheid zullen afgeschrikt worden door wiskunde en dat vak mogelijk mijden. Dat veel leerlingen breukrekenen later toch niet meer nodig hebben, is dan een 'selffulfilling prophecy'. Verwaarlozing van het breukenonderwijs leidt zo tot kostbare inperking van de onderwijskansen van veel leerlingen, ook van de betere leerlingen. Wilbrink & Hulshof (2011) wezen er al op dat dit rekenen op niveau-3F niet het rekenen is dat het ho vraagt, en dus misplaatst is in eindexamens havo, mbo en vwo. Het onderzoek van Siegler et al. toont dat ook de aansluiting van rekenen op wiskunde in het vo zelf in het geding is.

De rekenmachine kan rekenen, maar de leerling ook?

Bij verreweg de meeste rekenopgaven is gebruik van de rekenmachine toegestaan. Bij veel opgaven, zoals bij som VRT-2 in figuur 1, zijn de getallen zo gekozen dat er niet vlot mee valt te rekenen, anders dan door de rekenmachine te gebruiken. Een dwingende noodzaak voor deze getallenkeuze ontbreekt. Het kan namelijk heel goed anders, zoals Pfaltzgraff in de uitwerking demonstreert. Deze opgaven worden onnodig gedegradeerd van rekenen tot knoppen-drukken.

Volgens de minister ligt de bevoegdheid om gebruik van een rekenmachine toe te staan bij eindexamens bij het CvE. Het CvE zegt (CvE, 2012, p. 3-4): "Zonder

rekenmachine is het niet meer mogelijk vragen in een realistische context te stellen en wordt het moeilijk om tot variatie in opgaven te komen (het aantal 'kloppende' wiskundeopgaven met nette getallen is heel beperkt)." De keuze van de getallen in de VRT, zie het voorgaande, logenstrafte deze uitspraak.

Wel of niet gebruik van een rekenmachine toestaan maakt natuurlijk verschil in de uitkomsten. Bij individuele uitkomsten weet men niet eens of en zo ja hoe vaak de leerling naar de rekenmachine heeft gegrepen en dus niet in hoeverre er inderdaad is gerekend. Zolang er geen experimenteel bewijs is dat 'het niet uitmaakt', is aan te nemen dat het wel degelijk uitmaakt. In *Mathematics Framework for California Public Schools* (Lundin (ed), 2006, p.254) wordt Loveless (2004) aangehaald over het gebruik van de rekenmachine: "Regarding the differences in the computational accuracy of U.S. students with and without a calculator, Loveless wrote: "These differences are enormous - the difference between signaling mastery and signaling incompetence." He concludes: "These finding[s] suggest that making calculators available on a test of computation skills can make the difference between concluding that students have acquired certain skills - and concluding that they haven't. On each of these items, at least 40% of the nation's nine year olds computed correctly with or without a calculator provided. For most of the remaining students, calculators are the difference in whether they compute correctly or get the calculation wrong."

Een motie in de Tweede Kamer vraagt om gebruik van de rekenmachine in het onderwijs terug te dringen. In een advies aan de minister hierover, bepleit SLO juist meer gebruik van de rekenmachine (Buijs & Tolboom, 2012). Critici, onder wie Lenstra (KNAW-commissie rekenonderwijs), Van de Craats (Resonansgroep wiskunde), Tijms (Stichting Goed Rekenonderwijs), steken hun mening niet onder stoelen of banken. Zo concludeert Van de Craats: 'Bij de rekentoetsen behoort het gebruik van een rekenmachine te worden verboden' (Buijs & Tolboom, 2012, p.69). De rekenmachine toelaten in de rekentoetsen staat gelijk aan een uitnodiging aan het onderwijs en de leerlingen om rekenvaardigheid bij te houden met de rekenmachine. Een contradictio in



De scholen zijn verantwoordelijk voor goed rekenonderwijs

terminis.

Evenals bij taalverzorging is het bij het rekenen zo dat de uit het po meegekregen basisvaardigheden in het vo onvoldoende worden onderhouden en vermeerderd. Zonder oefening zal dit zo blijven; oefenen met een rekenmachine is geen oefenen in rekenen. Een rekentoets met gebruik van een rekenmachine, en zonder enige rekenprocedure te vragen, is onderdeel van de problemen in het Nederlandse rekenonderwijs, niet van de oplossing.

Verschillen havo - vwo op dezelfde toets-3F: intelligentie of vaardigheid?

Het is aan Cito en CvE om onomstotelijk duidelijk te maken dat deze VRT toetst of de rekenvaardigheid op peil is en zeker niet toetst op individuele verschillen in intelligentie of taalvaardigheid. In dit verband is het opmerkelijk dat de pilots in maart 2012 aanzienlijke verschillen tonen in resultaten van havisten en vwo-ers op dezelfde rekentoets-3F. Zo'n resultaat doet vrezen dat gerichte voorbereiding op deze rekentoets niet optimaal mogelijk is. Als de rekentoets transparant is, moeten gemiddelde en standaarddeviatie op de toets voor havo en vwo gelijk zijn, omdat verschillen in intelligentie zijn geabsorbeerd in verschillende voorbereidingstijden. Zo niet, dan is er in de mate waarin er verschillen zijn sprake van een intelligentietest. Een empirische toets op die voorbereidbaarheid is dan ook dat de resultaten op de rekentoets-3F bij de examens havo gelijk zijn aan die bij de examens vwo. Of: bij een gelijktijdig afgenomen intelligentietest moet de hypothese verworpen kunnen worden dat de correlatie tussen rekentoets-3F havo en intelligentietest positief is, zeg $> 0,6$.

Afsluiting

Het gaat om rekentoetsen die van hoge kwaliteit moeten zijn, want de rekentoets wordt een nieuw onderdeel van de eindexamens, van de kernvakkenregel bovendien, met precies één herkansingsmogelijkheid. Dit lijkt nog niet tot het grote publiek doorgedrongen. Het complexe veld overziende is het duidelijk dat vele instellingen een stukje bijdragen, maar geen verantwoordelijkheid voor het geheel nemen. De verantwoordelijke minister heeft de zorg voor rekentoetsen die inderdaad rekenen toetsen, en kan deze gebrui-

ken om het onderwijs te evalueren, in plaats van er de leerlingen individueel op af te rekenen. Zo dacht de Tweede Kamer er toch over toen in 2010 deze discussie op de rol stond (Handelingen, 2010). Het doel van alle aandacht voor rekenen is immers dat de scholen - niet de leerlingen - verantwoordelijkheid nemen voor goed rekenonderwijs en onderhoud van rekenvaardigheden.

Literatuur

- Buijs, K. & Tolboom, J. (red.) (april 2012). *De rol van de rekenmachine in po, s(b)o en vo. Advies SL* <http://goo.gl/p3fhY>
- Cito (2012). *Voorbeeldrekentoets VO 3F* <http://goo.gl/qLEzD>
- CvE (2012). *Advies. Brief 28-3-2012*. <http://goo.gl/04Mfv>
- Depaepe, F., Corte, E. de & Verschaffel, L. (2010). Teachers' approaches towards word problem solving: Elaborating or restricting the problem context. *Teaching and Teacher Education* 26, 152-160.
- Desoete, A., Roeyers, H. & De Clercq, A. (2003). Can offline metacognition enhance mathematical problem solving? *Journal of Educational Psychology* 95, 188-200.
- Evers-Vermeul, J. & Land, J. (2012). Slecht gelezen of slecht geleerd? 2. *Examens, Tijdschrift voor de toets praktijk*, 9(2), 25-29.
- Handelingen Tweede Kamer der Staten-Generaal. Vergaderjaar 2009-2010. Vergaderingnummer 70. <http://goo.gl/YI5Pt> en <http://goo.gl/zxXKH>
- Hickendorff, M. et al. (2009a). Solution strategies and achievement in Dutch complex arithmetic: latent variable modeling of change. *Psychometrika* 74(2), 331-350. <http://goo.gl/EXtMJ>
- Hickendorff, M. et al. (2009b). How to measure and explain achievement change in large-scale assessments: a rejoinder. *Psychometrika* 74(2), 367-374. <http://goo.gl/6eSP8>
- Hickendorff, M. (2011). *Explanatory latent variable modelling of mathematical ability in primary school*. Proefschrift Leiden. <http://goo.gl/zBfq>
- Loveless, T. (2004). *Computation skills, calculators, and achievement gaps: An analysis of NAEP items*. Brown Center on Education Policy, The Brookings Institution. <http://goo.gl/zSI8J>

Voor de rekentoets is er op het eindexamen slechts één herkansingsmogelijkheid



- Lundin, J. et al. (eds.) (2006), *Mathematics Framework for California Public Schools. Kindergarten Through Grade Twelve*. Sacramento: California Department of Education. <http://goo.gl/K3b0d>
- OCW (2012). Brief 8 juni 2012. <http://goo.gl/J9XU5>
- Ohlsson, S. (2011). *Deep learning*. Cambridge University Press.
- Pfaltzgraff, H. (2009). *Spijkerboek 1: Rekenen*. Epsilon Uitgaven.
- Pfaltzgraff, H. (2012). *Uitwerking en commentaar bij de voorbeeldrekentoets 3F van het Cito*. <http://goo.gl/eRJ7n>
- Polya, G. (1945/1957 2nd). *How to solve it*. Princeton University Press. <http://goo.gl/6i3AB>
- Schoenfeld, A.H. (2004). The math wars. *Educational Policy*, 18, 253-286. <http://goo.gl/43Faj>
- Siegler, R.S., et al. (online first 14 June 2012). Early predictors of high school mathematics achievement. *Psychological Science*. <http://goo.gl/OBGYF>
- Van den Heuvel-Panhuizen, M., et al. (2009). Large-scale assessments of change in student achievement: Dutch primary school students' results on written division in 1997 and 2004 as an example. *Psychometrika*, 74(2), 351-365. <http://goo.gl/hdup0>
- Weegh, J. van der (2005). *Schooltaalvaardigheid en realistisch rekenen: specifiek taalkundige moeilijkheden en de invloed op het begrijpen van realistische rekenopgaven*. Doctoraalscriptie, Utrecht (supervisors P.P.M. Leseman en H. van Luit).
- Wilbrink, B. (1977). Het verborgen vooroordeel tegen andere dan meerkeuze vraagvormen. *Congresboek ORD*, 219-222. <http://goo.gl/bMSe7>
- Wilbrink, B. & Hulshof, J. (2011). De wet, het rekenen, en de rekentoets in de eindexamens havo/vwo. *Examens, Tijdschrift voor de toetspraktijk*, 8(3), 18-22.
- De heer B. Wilbrink is psycholoog-onderwijsonderzoeker. www.benwilbrink.nl. De heer prof.dr. J. Hulshof is hoogleraar Wiskundige Analyse aan de Vrije Universiteit en bestuurslid van Beter Onderwijs Nederland. De heer H. Pfaltzgraff is leraar wiskunde havo/vwo sinds 1965, CEVO medewerker aan de examens wiskunde B vwo 1987-1993.

Gesignaleerd

Ervaringscertificaten moeten beter - 25 juni 2012 - <http://www.onderwijsinspectie.nl/actueel/nieuwsberichten/>

De kwaliteit van de ervaringscertificaten, het document waarmee een traject van erkenning van eerder of elders verworven competenties (evc) wordt bekrachtigd, moet beter. Dit blijkt uit onderzoek van de Inspectie van het Onderwijs.

De inspectie heeft daarnaast bekeken hoe examencommissies in het middelbaar en het hoger beroepsonderwijs omgaan met ervaringscertificaten. De staatssecretaris van Onderwijs heeft het rapport, "Examencommissies & ervaringscertificaten - over evc in het mbo en het hbo" vandaag aan de Tweede Kamer aangeboden.

Hoofdconclusie is dat in het mbo en in het hbo de kwaliteit van ervaringscertificaten in de meeste gevallen ernstig tekortschiet. Het certificaat biedt examencommissies daardoor onvoldoende informatie om verantwoorde beslissingen te nemen over vrijstelling, diplomering of maatwerk. Evc-aanbieders moeten de certificaten verbeteren, vooral wat betreft de onderbouwing van de erkende competenties. De overheid zou er goed aan doen om ook de komende jaren de regierol voor de kwaliteit van evc-voorzieningen voort te zetten.