

# Zullen we het over rekenen hebben? Die ‘situaties in het dagelijks leven’ hebben er niets mee van doen.

Aan de Vaste Kamercommissie van Onderwijs, AO 28-1-2015

Ben Wilbrink

“De #rekenoetsen is intussen de #fyra van het Nederlandse onderwijs aan het worden. Oplappen helpt echt niet meer.”

geciteerd uit een tweet van Ronald Buitelaar

U voert op 28 januari 2015 [Algemeen Overleg](#) met de bewindslieden over hun [beleidsvoornemens](#) die het [advies](#) van de commissie-Steur goeddeels volgen. Hierbij mijn analyse, op basis van annotaties bij brief en rapport.

De Commissie-Steur en de bewindslieden gaan voorbij aan de vraag wat rekenen ook al weer is, en wat goed rekenonderwijs is, maar zij gebruiken deze woorden wel volop. Kom op zeg, dat de ‘rekenoetsen’ geen rekenen toetsen, dat hoeft alleen uitleg tegenover de rekenideologen uit de school van Hans Freudenthal. Uit het antwoord op de vraag naar wat rekenen is, volgt het antwoord op wat goed rekenonderwijs is. Zoiets als:  $1 + 1 = 2$ . Maar ja, ook internationaal ligt hier een enorm probleem, dat constructivisme ([Wilbrink, 2014](#)) heet, en dat leidt tot onderwijs en toetsen waar psychologen op eenvoudige wijze van aan kunnen geven dat deze het leren rekenen en het toetsen van rekenen ernstig belemmeren ([Wilbrink & Kirschner \(2014\)](#)). Wie dat niet voetstoots wil aannemen: geef dan opdracht om dat experimenteel-psychologisch te onderzoeken; in Wilbrink & Kirschner is daar een concrete suggestie voor gedaan.

De ellende met de ‘rekenoetsen’ is vooral te danken aan twee dingen:

- (1) een valse start met de referentieniveaus ‘rekenen in situaties van het dagelijks leven’ (Van Streun, Meijerink), en
- (2) de evidente politieke blokkade om ooit nog over die referentieniveaus — specifiek: de F-niveaus — een discussie aan te gaan (met als vooralsnog weinig betekende uitzondering de inspanningen om een rekenoetsen-3S voor het vwo te ontwikkelen).

Een moeras is alleen maar leefbaar te krijgen door het droog te leggen, iets waar wij ooit enorme expertise in bezaten. Herstel het rekenonderwijs, zet de valse romantiek waardoor het rekenonderwijs verloren is gegaan maar bij het groot vuil. Een goed economisch principe is ook om geleden verlies te nemen — [Sander Dekker](#): een half miljard voor overeind hijsen van referentieniveaus taal en rekenen — en niet te proberen door extra investeren alsnog een succes van een uitzichtloze zaak te maken.

Rekenen is ongelooflijk belangrijk. Na de dramatische val in rekenvaardigheden zoals geregistreerd in de [PPON 2004](#) is het onvermijdelijk dat de politiek er bovenop zit. Heel goed. Maar rekenen is niet van belang om je boodschappen bij de grootgrutter te kunnen doen (‘rekenen in alledaagse situaties’, de psychologische misvatting die Nederland als staatsdidactiek is opgedrongen met de referentieniveaus rekenen). Twee argumenten waarom vlot kunnen rekenen dan wèl van belang is:

(1) De kwaliteit van het onderwijs over breuken blijkt voorspellend te zijn voor latere keuzen die leerlingen maken voor beta-profielen, en bovendien voor de resultaten die leerlingen in de beta-vakken boeken ([Siegler cs, 2012](#))

(2) *Numeracy*, ik vertaal dat maar met getalbegrip, blijkt belangrijk voor de gezondheid in het latere leven. Niet duidelijk is of rekening is gehouden met verschillen in intellectuele capaciteiten, ik neem aan van wel. Dan gaat het niet om het kunnen maken van een staartdeling, maar om de vertrouwde met getallen die het resultaat is goed rekenonderwijs. Wie de kansen en risico’s die de huisarts geeft, niet wil horen — want ja, dat zijn lastige getallen — die schaad uiteindelijk de eigen gezondheid ([Reyna, Nelson, Han & Dieckmann, 2009](#)). Ga er maar vanuit dat wat voor lichamelijke gezondheid geldt, ook geldt voor de gezondheid van de eigen financiën.

Het onderliggende probleem is dat deze ‘rekenoetsen’ vooral toetsen op verschillen in intellectuele capaciteiten. In de mate waarin dat het geval is, kunnen leraren en leerlingen zich niet doeltreffend voorbereiden op dit onderdeel van het eindexamen. De adviezen van de Commissie-Treur gaan grotendeels aan dit probleem voorbij, behalve waar het de taligheid van de ‘rekenopgaven’ betreft. Maar er moet méér gebeuren dan alleen de belasting door taligheid verminderen. Anders blijft het doormodderen in het moeras van het ‘rekenen in situaties van het dagelijks leven’.

Bij vasthouden aan de ‘rekenoetsen’ als nieuw onderdeel van de eindexamens vo en mbo is de schade te beperken door te borgen dat deze rekenoetsen doeltreffend zijn voor te bereiden: zij toetsen rekenvaardigheid en niet daarbovenop ook nog verschillen in intellectuele capaciteiten. De borging is te bereiken door deze buiten invloed van Cito en CvTE te beleggen bij bijvoorbeeld een vakgroep psychologische methodenleer in combinatie met een vakgroep cognitieve psychologie die de mentale belasting bij de rekenopgaven empirisch toetst. De schade is te minimaliseren door als richtlijn te nemen waar ook het Cito in aanvankelijke modelberekeningen op uit kwam: geen enkele leerling extra gezakt op basis van de rekenoetsen. In plaats van de willekeurige en niet uit te leggen ‘maximaal vijf procent extra gezakten’ van de Commissie-Steur, dus nul procent.

Beter is:

Wijzig de Wet op de referentieniveaus Nederlandse taal en rekenen (haal de pseudowetenschap eruit); maak met het onderwijsveld afspraken op basis van het voorstel van [Jan Karel Lenstra](#): rekenen in de onderbouw, afsluiten met diagnostische toetsing.

## Kamerbrief 17-12-2014 voortgang referentieniveaus. We hebben geen eigen beleid meer, maar volgen onze commissies

[Brief](#) van de staatssecretaris en de minister van onderwijs, cultuur en wetenschap. Kamerstuk 31 332 Nr. 36. Doorlopende Leerlijnen Taal en Rekenen. Stand van zaken invoering referentieniveaus taal en rekenen (17 december 2014).

### Aantekeningen bij de tekst

Een groot deel van deze uitvoerige brief betreft detailregelingen waar ik geen aantekeningen bij zal maken.

“Er wordt door scholen en docenten hard gewerkt aan de verbetering van de taal- en rekenvaardigheden van leerlingen in het vo en in het mbo. We constateren dat de betrokkenheid van scholen en docenten groeit. Dat we er nog niet zijn, is bekend. Voordat alle leerlingen de referentieniveaus voor rekenen daadwerkelijk beheersen, zijn er nog stappen te zetten.”

kamerbrief 17 december 2014

Uit deze tekst spreekt de niet op feiten gebaseerde overtuiging dat ‘hard werken’ en ‘nog meer stappen zetten’ uiteindelijk ertoe zullen leiden dat jongeren met adequate taal- en rekenvaardigheden de school verlaten. Zolang de ‘rekentoetsen’ in feite leerlingen sorteren op intellectuele capaciteiten in plaats van op rekenvaardigheid, kan er geen sprake van zijn dat welk beleid van de bewindslieden en van scholen en leraren dan ook, resulteert in mooie resultaten op deze ‘rekentoetsen’. Voor literatuur, zie: [Wilbrink & Kirschner \(2014\)](#); [Wilbrink, Hulshof & Pfaltzgraff \(2012\)](#). Dit alles betreft natuurlijk de kwestie of deze ‘rekentoetsen’ wel valide zijn — nee, dus — waarover nog veel meer valt te zeggen.

“(..) een structurele situatie waarin de taal- en rekenvaardigheden van leerlingen op niveau zijn, scholen hun taal- en rekenonderwijs op orde hebben en waarin alle leerlingen een toets krijgen die uitdagend is.”

kamerbrief 17 december 2014

Wat is dat precies, taal- en rekenonderwijs dat ‘op orde’ is? Want dat staat niet in de ‘Wet op de referentieniveaus Nederlandse taal en rekenen’ beschreven. Terwijl het hier toch gaat om de kern van het probleem: kennelijk is het rekenonderwijs in po en vo de laatste decennia ernstig tekortgeschoten. Maar waarin is het tekortgeschoten? In didactische praktijken misschien? Zo ja, moet daar dan niet eens kritisch naar worden gekeken? Hier is natuurlijk aan de orde of het constructivisme een onderwijsideologie is die van staatswege kan worden opgelegd. Voor dat constructivisme (ook wel situationisme genoemd, in het VK *progressivism* als ik het goed heb, en elders ook vaak reform-didactiek) zie [Wilbrink \(2014\)](#) voor een korte introductie.

Een ‘rekentoets’ moet ‘uitdagend’ zijn? Dat heeft Nederland tot nu toe wel gemerkt! Is dit een nieuw criterium voor de kwaliteit van onderdelen van eindexamens? Dergelijke onnodige kwalificaties belasten het debat over het rekenonderwijs.

“(..) worden docenten en scholen ondersteund bij het vormgeven van goed taal- en rekenonderwijs.”

kamerbrief 17 december 2014

Nogmaals, want daar is het belangrijk genoeg voor. Het probleem dat hier is toegedekt: de bedroevende staat van de rekenvaardigheden van jongeren is het gevolg van precies hetzelfde ideologisch gebaseerde rekenonderwijs dat nu wederom over de band van de referentieniveaus en talrijke ondersteunende diensten en wat niet al over leerlingen in vo en mbo wordt uitgestort. Ik formuleer het maar eens wat minder parlementair. De referentieniveaus voor rekenen, in korte tijd in elkaar geflanst door een werkgroep-Van Streun van vooral belanghebbenden bij reformistische rekendidactiek, is niet boven iedere twijfel van deugdelijkheid verheven ook al proberen velen bij voortdurend dat beeld van deugdelijkheid te vestigen. Bewindslieden voorop, trouwens. Die referentieniveaus zouden wel eens weinig meer dan strohalmen kunnen zijn. Laat dat behoorlijk onderzoeken, niet door een commissie dus.

“Ook het inzicht dat de toets oplevert aan leerling en docent over waar de leerling staat en waar eventueel extra inzet nodig is, is van belang. (..) Daarom worden, op verschillende niveaus, rekentoetsen en -examens afgenomen, die inzicht bieden in de mate waarin leerlingen de benodigde rekenvaardigheden beheersen.”

kamerbrief 17 december 2014

Dit is toch wel een heel bijzondere invulling van wat een eindexamen bedoelt te zijn. Laat ik er even geen verdere woorden aan

vuil maken.

“Daarom worden de nodige maatregelen ingezet om zorgvuldig, beheerst en waar nodig stapsgewijs in de periode 2015-2020 toe te groeien naar 2020. In dat jaar moeten de (taal- en) rekenvaardigheden van alle leerlingen op orde zijn.”

kamerbrief 17 december 2014

Hoe is dit mogelijk? Dezelfde bewindslieden die koste wat kost vast wilden houden aan 2015-2016 als jaar van definitief meetellen bij eindexamens, zetten de stip nu op de horizon van 2020? Wat is er gebeurd? Ah, er is wel het nodige beleidsproza nodig om deze draai te verkopen, want las ik niet iets eerder in dezelfde alinea:

“(.) meetellen voor het behalen van hun diploma. Dit wordt onverminderd doorgezet, vanwege het belang voor leerlingen van de beheersing van deze vaardigheden op het niveau van de referentieniveaus en voor het vervolgonderwijs.”

kamerbrief 17 december 2014

En dit is een passage waar ik vooral plaatsvervangende schaamte bij voel.

“Naar aanleiding hiervan zijn onder leiding van het College voor Toetsen en Examens (hierna: het CvTE) een aantal acties en onderzoeken in gang gezet.”

kamerbrief 17 december 2014

Toch blijft dit opmerkelijk: het CvTE is immers ten volle verantwoordelijk voor de kwaliteitsgebreken die zij nu zou moeten herstellen. Waar baseren de bewindslieden hun vertrouwen in dit CvTE op? Afijn, het CvTE onderneemt verbeteracties, zoals beschreven in [bijlage 1](#) bij deze kamerbrief, eigenlijk zou ik het CvTE-stuk daarom eveneens van aantekeningen voorzien, tenslotte ligt het CvTE-stuk eveneens voor in het komende algemeen overleg.

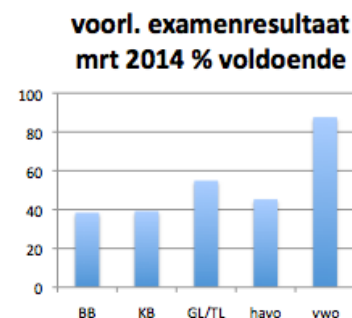
En wat gebeurt er dan met de aanbeveling-Bosker voor meer transparantie en openbaarheid? Het gaat natuurlijk om openbaarmaking van de ‘rekentoets’-vragen, onmiddellijk na afloop van de examens, zoals overigens voor onze examens gebruikelijk is. Maar zo brengen de bewindslieden het niet onder woorden. Een niet zo fijne manier van communiceren met ons Nederlanders. Want wat gebeurt er met digitale examens zoals deze ‘rekentoetsen’: de examenvragen blijven geheim, zodat leerlingen niet in staat zijn om hun recht te halen bij ondeugdelijke vraagstellingen. Ah, zijn er dan ondeugdelijke vragen door de zeef van het CvTE gekomen? Ja, en niet zo’n klein beetje, goed gedocumenteerd op tal van publieke plaatsen zoals de website van Jan van de Craats, en de de mijne. Over geheimhouding van examenvragen, zie [hier](#).

“De domeinrapportages maken voor de docent inzichtelijk hoe er wordt gescoord door zowel de klas als door de individuele leerling op de vier domeinen (getallen, verhoudingen, meten en meetkunde, verbanden). De rekenrapportage van Vensters vo maakt de rekentoetsresultaten op schoolniveau inzichtelijk, evenals de resultaten per klas, afgezet tegen de resultaten op andere vakken.”

kamerbrief 17 december 2014

Dit is een regelrechte aansporing *to teach to the test*. Ik heb daar geen moeite mee, op voorwaarde dat het gaat om een verdraaid goede toets die inderdaad dekt waar het onderwijs voor staat. Zoals Popham zou zeggen. Is dat het geval dan? Hebben we een ‘rekentoets’ die verdraaid goed is? Kom op, zeg. In Den Haag kun je zomaar een kamerlid tegen het lijf lopen die je kan vertellen dat de hamvraag in dit dossier is: ‘Wat is rekenen?’ En hoe kan het dat ‘experts’ en ‘expertisecentra’ Nederland ‘rekentoetsen’ verkopen (zo mag het wel worden genoemd) die geen rekenen toetsen (want ‘rekenen’ mag op de rekenmachine— wist u dat, hebt u wel eens zo’n voorbeeldrekentoets van de [website](#) van het Cito proberen te maken?) maar wel leerlingen rangordenen naar troebele verschillen in intellectuele capaciteiten (taalvaardigheid, Nederlands als moedertaal of niet; omvang werkgeheugen; wereldkennis; toetsbehendigheid)?

Laat ik vast een illustratie geven van dat testen zodanig dat de scores verschillen in intellectuele capaciteiten aanduiden. De gemiddelden van havisten vergeleken met vwo-ers, bijvoorbeeld. Of verfijn dat naar profielen. Kijk, voor zoiets overzichtelijks als rekenvaardigheid zijn leraren en leerlingen uitstekend in staat zich zo op een adequate rekentoets voor te bereiden dat vrijwel iedereen daar probleemloos voor slaagt. Havisten zouden dan gemiddeld even hoog moeten scoren als vwo-ers (op dezelfde rekentoets-3F). Is dat het geval? Kijk zelf eens. Bron: cijfers van het CvTE (uit bijlage bij kamerbrief 13 juni 2014), zie [Wilbrink & Kirschner \(2014b\)](#): “De feiten in de grafiek zijn hard: meer dan 40 procentpunten verschil op dezelfde ‘rekentoets’-3F tussen havo en vwo, met nog één herkansingsgelegenheid te gaan. Binnen het vmbo zijn de verschillen al verkleind door aangepaste cesuur: BB lager, GL/TL hoger, op dezelfde ‘rekentoets’-2F.” Het is klinkklaar dat havisten er niet in slagen zich op deze ‘rekentoets’ behoorlijk voor te bereiden — en dat kan niet anders betekenen dan dat deze ‘rekentoets’ vooral test op persoonlijke verschillen van intellectuele aard, waartoe ook taalvaardigheid behoort.



“Docenten krijgen in komend schooljaar na afloop van elke afnameperiode, op voorwaarde van geheimhouding, de mogelijkheid tot inzage in alle opgaven van de door hun leerlingen gemaakte toetsen, de antwoordmodellen en de door hun leerlingen gegeven antwoorden. Binnen de school of instelling krijgt de docent digitaal inzage in de rekentoets via de examensecretaris of rekencoördinator.”

kamerbrief 17 december 2014

Dit is belachelijk. Hier staat dat de docent zelfs met de eigen leerlingen die inzage niet mag delen. Noch met kamerleden, trouwens. Schande. Maar ja, bewindslieden vinden dat docenten zo ‘een schat aan informatie’ krijgen. Docenten doen er verstandig aan dit gegeven paard eens goed in de bek te kijken.

“Wij hebben begrip voor de behoefte aan transparantie ten aanzien van de toets. Gelijktijdig moet er wel rekening worden gehouden met de uitvoerbaarheid hiervan.”

kamerbrief 17 december 2014

Zeker, wij weten dat discriminatie niet geoorloofd is, maar we moeten gelijktijdig wel rekening houden met de uitvoerbaarheid ervan. Hoe erg is dit? Heel erg. Lees er alles over in het proefschrift van [Job Cohen, 1981](#).

Het komt er nu dus op neer dat eindexamens voortaan aan de wetten van Kafka zijn onderworpen, niet aan die van de rechtsstaat (zie ook Willem Witteveen (2014, posthuum). **De wet als kunstwerk. Een andere filosofie van het recht.** [Boom](#), op 9 januari besproken in NRC Boeken C2-3). Hier zal de rechter uiteindelijk aan te pas moeten komen. De uitspraak van de rechter is 100% voorspelbaar: bij geheimhouding van mogelijk kwalitatief ondermaatse examenvragen worden examenkandidaten direct in hun gerechtvaardigde belangen geschaad. Overigens is er casuïstiek van de UvA rond 1980, waarbij ook de Commissie voor de Bestuurshervorming (Commissie-Slagter) zich in niet mis te verstane bewoordingen over de onwenselijkheid van geheimhouding heeft uitgesproken.

<http://www.benwilbrink.nl/projecten/geheimhouding.htm#18>

Er is een WOB-verzoek om openbaarmaking gedaan; de bewindslieden trokken er hun handen vanaf, gaven het verzoek door aan het CvTE, en dat college doet het voorkomen als zou vrijgeven van examenvragen (na afname) een bedreiging voor de Staat der Nederlanden vormen (want dat is wettelijk eigenlijk de enige grond voor weigering van een WOB-verzoek). Lees er alles over:

<http://www.benwilbrink.nl/projecten/geheimhouding.htm#CvTE-WOB> en

<http://www.benwilbrink.nl/projecten/geheimhouding.htm#CvTE-brochure>

Hoe zijn de bewindslieden in dit moeras terechtgekomen? Dat is de vraag hoe het mogelijk is dat de ‘rekentoets’ wordt afgenomen als digitale toets, in plaats van als probleemloze schriftelijke toets op rekenvaardigheid. Afijn, hier zit een wereld van belangen achter, dat moeten we eindelijk eens willen weten. In de VS zijn daar al enkele onderste stenen boven gekomen, en dat mag een indicatie zijn van hoe de hazen ook in Nederland lopen.

[http://blogs.edweek.org/edweek/on\\_california/2014/09/ipad\\_scandal\\_grows\\_in\\_la\\_raising\\_leadership\\_questions.html](http://blogs.edweek.org/edweek/on_california/2014/09/ipad_scandal_grows_in_la_raising_leadership_questions.html)

“Daarom stelt de commissie Steur voor, net zoals de commissie Bosker, om de resultaten op de rekentoets mee te laten tellen voor het behalen van het diploma vanaf de al eerder vastgestelde school- en studiejaren (...).”

kamerbrief 17 december 2014

Bewindslieden misbruiken hier hun commissies, door ze verantwoordelijk te maken voor het voortzetten van hun eigen beleid. Mogelijk zijn beide commissies alleen maar naïef geweest door op de stoel van de politiek te gaan zitten, maar dan hoeven bewindslieden daar nog niet gretig gebruik van te maken. Het wordt op deze manier problematisch voor integere burgers om nog deel te nemen aan ministeriële commissies. Het zijn mede deze politieke trucs die het ‘rekentoets’-dossier tot een zo taaie materie maken. Saillant ander voorbeeld hiervan: bij de behandeling van de Wet op de referentieniveaus Nederlandse taal en rekenen is ondubbelzinnig (ook in de [Handelingen](#) van 31 maart 2011 dus) vastgelegd dat er bij de rekentoets natuurlijk *niet* met de rekenmachine gerekend gaat worden, maar ‘met het [koppie](#)’ (staatssecretaris Van Bijsterveldt). Enzovoort.

“(..) maar om ook een transitieperiode in te stellen waarin scholen, docenten en leerlingen nog tijd krijgen om in stappen toe te kunnen groeien naar het halen van de referentieniveaus. Het betekent meer tijd om het onderwijs op orde te krijgen en het niveau van leerlingen omhoog te brengen.”

kamerbrief 17 december 2014

Het bovenstaande is het vervolg van wat de Commissie-Steur adviseert. Ik ben benieuwd hoe de Commissie aannemelijk maakt dat dit een realistisch scenario is, op deze plaats in ieder geval geen verwijzing naar bronnen daarvoor. Waarom zou de komende jaren gaan lukken, wat afgelopen jaren zo jammerlijk is mislukt? Ik moet het duidelijker zeggen: de Commissie — en de bewindslieden — gaan sleutelen aan de symptomen. Maar voor een vertrouwenwekkende voorspelling van geleidelijke verbetering in de symptomen, zal het toch nodig zijn een treffende diagnose van de onderliggende kwaal te hebben. Waar is die? Wat is de *oorzaak* — wat zijn de *oorzaken* — van de teleurstellende resultaten op deze ‘rekentoetsen’? Deugen die toetsen in de grond van de zaak wel? Zijn ze *valide*?

“*Werken met vaardigheidsscores*”

kamerbrief 17 december 2014

Een onbegrijpelijk voorstel. Ja, zeker, het Cito heeft een voor leken onnavolgbare methodologie ontwikkeld om gestandaardiseerde vaardigheidsscores te kunnen vaststellen, een techniek die terecht is toegepast in de PPOON-onderzoeken (PPOON: Periodieke Peiling Onderwijs Nederland). Maar waarom zou dit een adequate techniek zijn voor de ‘rekentoets’? Ah, hij is nodig om een over de jaren heen vergelijkbare schaal voor cijfergeving te hebben. Maar dat beantwoordt de vraag niet: waarom is het in het specifieke geval van een aan de eindexamens toegevoegde toets op basisvaardigheden rekenen niet voldoende om alleen een zak-slaaggrens (cesuur) vast te stellen? Op zijn Engels: een *pass-fail score*? Omwille van het mes van Ockham,

eenvoudigheidshalve dus, zou zeker in een overgangperiode kunnen worden volstaan met die *pass-fail score*.

“Aanpassen van de positie van de rekentoets in de uitslagregel”

kamerbrief 17 december 2014

Lees deze sectie in de kamerbrief eens in zijn geheel door (de sectie is te lang om hier in zijn geheel te citeren).

Dit is toch niet te geloven: een regeling die ieder jaar verandert, en per jaar mogelijk pas achteraf bekend is. Wat las ik daar net over, bij de bespreking van het posthume boek van Willem Witteveen over wetgeving?

“Wetten horen geen terugwerkende kracht te hebben, zij moeten begrijpelijk zijn en regels horen niet onderling tegenstrijdig te zijn. Men moet ook geen regels maken die redelijkerwijs niet kunnen worden nageleefd, men moet regels niet zo vaak veranderen dat burgers de draad kwijt raken en men moet de regels toepassen zoals ze bekend zijn gemaakt — en dus niet via rechtspraak en bestuur er iets heel anders van maken.”

Marc Chavannes (9 januari 2015). De rechtsstaat is in Nederland een utopie. NRC Boeken C2-3. Bespreking van Willem Witteveen (2014, postuum). De wet als kunstwerk. Een andere filosofie van het recht. [Boom](#)

“Ook wordt op voorstel van de commissie het aantal herkansingen voor de rekentoets in het vo vergroot naar drie (dat betekent vier kansen in totaal) (..)”

kamerbrief 17 december 2014

Dit voorstel is natuurlijk heel erg noodzakelijk, althans in het *worst case scenario* dat de ‘rekentoets’ *niet* definitief door de Tweede Kamer of door bewindslieden zelf wordt afgeblazen. De vraag is dus: hoe is het eigenlijk mogelijk dat jarenlang slechts *één* herkansing is geboden — met daarna alleen nog de mogelijkheid om staatsexamen af te leggen (mededeling OCW)? Daar ligt een reactie publiceerde op de waarschuwing van [Wilbrink & Hulshof \(2011\)](#) dat de regeling voor de ‘rekentoets’ zou leiden tot slachtoffers onder examenkandidaten: [Zwitser & Béguin \(2011\)](#). Ik wil hier een lange discussie — die destijds ook via email is gevoerd met Anton Béguin — sterk inkorten: commissies Bosker en Steur hebben ondertussen wel duidelijk gemaakt dat het Cito modelberekeningen heeft gepresenteerd die geen relatie hebben met de werkelijkheid van de pilotstudies.

Er is voor de ‘rekentoets’ een interessant alternatief scenario mogelijk, zonder te tornen aan de wens van de minister om de ‘rekentoets’ in de kernvakkenregel mee te nemen. Dat scenario is heel eenvoudig, en verdient dus aandacht voordat allerlei ingewikkelde scenario’s doorgerekend gaan worden.

#### **Geef voor de rekentoets een onbeperkte herkansingsmogelijkheid.**

De toets gaat per computer afgenomen worden, zodat er technisch geen extra problemen zijn. Het is natuurlijk mogelijk om voor herkansingen een strengere eis te stellen, om gokken tegen te gaan, maar dat maakt de regeling weer ingewikkelder. Een andere mogelijkheid om gokgedrag tegen te gaan is om herkansingstoetsen veel langer te maken dan de toets bij eerste gelegenheid. Blijven te kunnen rekenen, daar gaat het uiteindelijk om bij eindexamenkandidaten.

Ben Wilbrink & Joost Hulshof (2011). Opmerking bij het artikel ‘Gaat meefout bij de rekentoets slachtoffers maken? *Examens, Tijdschrift voor de Toetspraktijk*, 8 #4, 23-26. [http://benwilbrink.nl/publicaties/11dupliek\\_rekentoets.htm](http://benwilbrink.nl/publicaties/11dupliek_rekentoets.htm) (concept voor reactie op Zwitser & Béguin (2012))

“De inspectie wordt gevraagd om de rekenresultaten te betrekken in het toezicht. Wij vragen het CvTE en het Cito om de ontwikkeling van de resultaten goed te monitoren en indien dit nodig blijkt zullen wij op basis daarvan bijsturen”

kamerbrief 17 december 2014

Wat stellen de bewindslieden zich hier concreet bij voor? Gaat de Inspectie eindelijk eens onderzoeken of de ‘rekentoets’ wel rekenen toetst? Gaat het Cito zich inspannen om het onbedoelde testen op verschillen in intellectuele capaciteiten uit de ‘rekentoetsen’ te halen? Wordt het CvTE gefaciliteerd met een behoorlijke afdeling Research & Development? Of denken de bewindslieden dat slaagpercentages ‘voor zich’ spreken?

“Maar zoals geconstateerd in de voortgangsrapportages taal en rekenen 2013 en 2014 zijn er groepen leerlingen voor wie rekenen moeilijk is en zal blijven en waarvoor zonder de mogelijkheid van een alternatieve toets het behalen van een startkwalificatie in gevaar komt.”

kamerbrief 17 december 2014

Voor rekenzwakke leerlingen geldt des te sterker de noodzaak om volstrekt helder te zijn over wat rekenen precies is, en wat niet. Toetsen op intellectuele capaciteiten is voor alle leerlingen vervelend, voor zwakke leerlingen een hel. De leerlingen die minder begenadigd zijn wat intellectuele capaciteiten betreft, zijn veelal prima in staat zich basale rekenvaardigheid te verwerven. Zonder de contextflauwekul.

“Tot slot

Met de in deze brief beschreven maatregelen zorgen we er samen voor dat het taal- en rekenonderwijs zo snel mogelijk verder versterkt wordt. Dat doen we op een zo zorgvuldig mogelijke wijze.”

kamerbrief 17 december 2014

Tot die zorgvuldigheid behoort ook, mogen we hopen, een integere behandeling van beschikbare kennis en inzichten, als onderscheiden van de houding dat er nu eenmaal verschillen van opvatting zijn, en wij dus vasthouden aan de onze. Handen in de lucht voor de opvatting dat de aarde plat is, ja? Mooi zo. Er zijn grenzen aan wat politiek maakbaar is ([Witteveen, 2014](#)).

## ‘Steur’: de ‘rekentoetsen’ 2F en 3F deugen niet, maar dat zeggen we niet hardop

Commissie verantwoorde invoering rekentoets (verder Commissie-Steur te noemen) (5 december gedateerd, 17 december 2014 openbaar gemaakt). [Advies over de uitwerking van de referentieniveaus 2F en 3F voor rekenen in toetsen en examens.](#); met [kamerbrief](#) van de bewindslieden.

Ik begrijp het rapport van deze Commissie-Steur niet. Ik krijg sterk de indruk dat drie leden van de Commissie er geen *say* in hebben gehad.

De Commissie heeft kennelijk geen behoorlijke kennis van het rekentoets-dossier, en evenmin van de controverse over het rekenonderwijs in Nederland. Maar door miskenning van de waarschijnlijke oorzaken achter dit leed dat rekentoets heet, komt de Commissie niet verder dan het voorstellen van lapmiddelen, symptoombestrijding, jarenlange reparaties uit te voeren door instanties die om te beginnen zelf verantwoordelijk zijn voor de ontstane reken(toets)-problemen. Mogelijk was dat ook de opdracht aan de Commissie, maar door het ontbreken van de opdrachtbrief blijft dat onduidelijk. Lapmiddelen of niet, daar valt in testpsychologische zin nog best het nodige over te zeggen. Zou ik denken. Maar ik zie daar geen helder spoor van in het rapport van de Commissie. Een voorbeeld: de Commissie doet geen poging om een verklaring te vinden voor het enorme verschil in resultaten tussen havisten en vwo-ers, terwijl het toch evident is dat een goede verklaring iets zal zeggen over de validiteit of juist het ontbreken van validiteit van de ‘rekentoetsen’.

### Aantekeningen bij het advies

De aanbiedingsbrief is een samenvatting. Hierin geen enkel signaal van twijfel aan het rekengehalte van de ‘rekentoets’, geen uiting van ook maar een begin van bezorgdheid over ideologisch geïnspireerde opvattingen over rekenonderwijs zoals deze blijken uit zowel het werk van de werkgroep-Van Streun (Commissie Meijerink, referentieniveaus Nederlandse taal en rekenen) als dat van de rekentoetswijzercommissie-Schmidt ([2F en 3F](#)), geen zorgen over de validiteit van de ‘rekentoetsen’ anders dan een verwijzing naar de aanbevelingen van de Commissie-Bosker over de kwaliteit van de toetsen en de opgaven. Er zijn legio aanwijzingen en rapporten dat er achter elkaar stukken van deze Fyra af donderen, maar deze commissie laat hem gewoon doorstomen.

“Het onderwijs voor rekenen in vo en mbo moet van ver komen.”

Commissie Steur, aanbiedingsbrief

De commissie ziet althans dit deel van het probleem. Belangrijker is nog het rekenen in het bo: is dat rekenonderwijs adequaat, en zo nee, waar schort het aan, en is het vervolgonderwijs wel in staat de tekorten op te vangen en te repareren?

“De Tweede Kamer heeft besloten dat aan het eind van een opleiding moet worden getoetst of het vastgestelde referentieniveau is bereikt.”

Commissie Steur, Inleiding - historie

Dit is toch lastig: er is geen bron gegeven voor deze bewering, en hij lijkt mij onjuist. De formulering is ook dubbelzinnig: ‘toetsen’ kan zowel de betekenis van examineren hebben, als van evalueren. Mij staat wel scherp op het netvlies dat bij de behandeling van de Wet nadrukkelijk de mogelijkheid van evalueren open werd gehouden (Handelingen, 31 maart 2010).

“De niveaus 2S en 3S zijn niet wettelijk aan een sector toegewezen.”

Commissie Steur, Inleiding - historie, voetnoot 5

Dit gegeven had bij de Commissie de nodige alarmbellen moeten doen rinkelen. Het is immers niet te verenigen met deze PR-taal:

“Met de invoering van de referentieniveaus wordt een kwaliteitsverhoging beoogd, van zowel de rekenresultaten van het Nederlands onderwijs als geheel alsook van een vastgelegd beheersingsniveau voor rekenen door iedere leerling (en dat op verschillende momenten in de schoolloopbaan).”

Commissie Steur, Inleiding - historie

Hoogdravende taal. Maar dan zou toch allereerst de aandacht op de niveaus 2S en 3S gericht moeten zijn! Nope.

“Met de uitslag op de toets/het examen kan de vervolgopleiding worden geïnformeerd over de mate waarin een leerling het gevraagde referentieniveau heeft bereikt.”

Commissie Steur, Inleiding - historie

De commissie gaat er hier stilzwijgend van uit dat de ‘rekentoets’ veel meer is dan alleen een toets of een gewenst niveau is bereikt. Toch is dit volstrekt niet vanzelfsprekend, en misschien ligt hier zelfs een van de oorzaken waarom de ontwikkeling van deze ‘rekentoetsen’ door het Cito in een drama is geëindigd. Namelijk door een niet behoorlijk doordacht streven om leerlingen op deze toets zo betrouwbaar mogelijk van elkaar te onderscheiden op *whatever it is that the test measures*. En dat laatste is vooral een onheldere mix van intellectuele verschillen, waaronder verschillen in taalvaardigheid, maar zeker ook verschillen in mentale belastbaarheid.

“De behaalde leerlingresultaten op de (pilot)rekentoetsen en de pilot centrale examens rekenen vallen tot op heden tegen.”

Commissie Steur, Inleiding - tegenvallende rekenresultaten

Dat is maar net hoe je het bekijkt: leraren in het veld hadden deze ‘tegenvallende’ resultaten wel degelijk verwacht, voor hen vielen ze dus niet tegen. Ik vermoed dat de betekenis hiervan groter is dan de leden van deze commissie, misschien met uitzondering van Greetje van de Werf, hebben vermoed.

De commissie rept hier met geen woord over de vraag of de resultaten op de ‘rekentoetsen’ een stijgende lijn laten zien. En dat doen ze niet, afgezien dan van de versoepelingen die gaandeweg door het CvTE zijn aangebracht (minder vragen, verhoudingsgewijs minder contextvragen, aangepaste cesuur voor verschillende opleidingswegen in het vmbo). De betekenis van het antwoord op deze vraag is niet gering: immers, wanneer bij steeds adequater inspanningen van de scholen de resultaten desondanks gelijk blijven, is dat een signaal dat deze ‘rekentoetsen’ maar tot op beperkte hoogte echt goed zijn voor te bereiden door leraren en hun leerlingen.

“(.) voorkomen moet worden dat leerlingen de dupe worden van onvolkomenheden in de toetsen en van rekenonderwijs dat nog niet op orde is.”

Commissie Steur, Opdracht

Dit is een prevelementje van politici. Hoe gaat de Commissie ermee om? Wel, naar onvolkomenheden in het rekenonderwijs wordt kennelijk verder niet gekeken. Dat is toch bijzonder riskant, want de Commissie veronderstelt dan in feite en stilzwijgend dat dat rekenonderwijs zelf niet problematisch is, dat het met een beetje goede wil van de scholen prima voor elkaar zal komen. Quod non.

“Ons advies gaat niet over de kwaliteit van de toetsen/examens, we gaan ervan uit dat daaraan wordt gewerkt volgens de adviezen van de commissie-Bosker.”

Commissie Steur, Opdracht

Wel, dat is tenminste duidelijk. Het doet denken aan die [tweet](#) van bestuursvoorzitter Marten Roorda van het Cito (9 oktober 2013) “Voor alle duidelijkheid: #Cito is geen partij bij vaststelling van inhoud Rekentoets en typen opgaven. Cito voert de richtlijnen van CvE uit”. Ik weet het niet, hoor, maar toch heb ik zo’n vermoeden dat dit (de opstelling van de Commissie Steur zowel als die van de bestuursvoorzitter van het Cito) in flagrante strijd is met de [Standards](#) (ook in Nederland richtinggevend voor in beginsel iedereen die verantwoordelijkheid draagt voor tests, toetsen en examens), of gewoon met de zorgplicht die OCW (en zijn Commissie) en Cito hebben naar leerlingen.

“Daarbij mag het ook niet zo zijn dat scholen die hun rekenonderwijs nog niet op orde hebben, worden ‘beloond’.”

Commissie Steur, Opdracht

Dit is een zorgwekkende opdracht, hoe is de Commissie hiermee omgegaan? Er spreekt ook wantrouwen naar het onderwijsveld uit, om niet te zeggen: een vooronderstelling van kwade trouw. En dat blijkt bepaald niet alleen uit dit ene zinnetje in de Opdracht-sectie. Werpt de Commissie dit verre van zich, of huppelt ze vrolijk mee?

“Wat betreft de literatuur willen we in het bijzonder noemen: (..) de startrapportages van de intensiveringstrajecten rekenen van de Steunpunten taal en rekenen vo en mbo. Een ander belangrijk bestudeerd document is de Regeling omzetting scores in cijfers centrale examens en rekentoets VO 2014.”

Commissie Steur, Werkwijze

- Regeling omzetting scores in cijfers centrale examens en rekentoets VO 2014  
[http://wetten.overheid.nl/BWBR0035058/geldigheidsdatum\\_03-10-2014](http://wetten.overheid.nl/BWBR0035058/geldigheidsdatum_03-10-2014)

Over rekenen gesproken. Startrapportage Intensiveringstraject rekenen vo. Steunpunt Taal en Rekenen mbo.

<http://www.steunpunttaalenrekenenvo.nl/sites/default/files/20140120%20Startrapportage%20mbo%20Over%20rekenen%20gesproken.p>

Rekenen op het vo. Startrapportage Intensiveringstraject rekenen vo. Steunpunt Taal en Rekenen vo.

[http://www.steunpuntaalenrekenen.nl/sites/default/files/Rapportage%20Steunpunt%20vo.%20intensiveringstraject\\_1.pdf](http://www.steunpuntaalenrekenen.nl/sites/default/files/Rapportage%20Steunpunt%20vo.%20intensiveringstraject_1.pdf)

“Om meer zicht te krijgen op de techniek van de rekentoets/het ce-rekenen, de overwegingen en de achtergronden die daarbij hebben gespeeld heeft de commissie gesproken met experts van CvTE en Cito.”

Commissie Steur

Deze commissie was natuurlijk zelf prima in staat om de verhalen van Cito en CvTE door te prikken, waar dat nodig mocht zijn. Toch blijft hier de schijn hangen van een commissie die een gewillig oor heeft voor Cito en CvTE, de instanties die het meest direct zijn betrokken bij de ‘rekentoets’-ellende, en geen tegengeluiden heeft willen horen, zoals bijvoorbeeld in de [hoorzitting](#) die de onderwijscommissie van de Tweede Kamer op 4 december 2013 hield. In [Wiskunde-brief](#) #654 zijn veel van de ingediende stukken op te halen, zoals ook die van [Jan Karel Lenstra](#) met een uitstekend en sober alternatief voor de ‘rekentoetsen’

“Het conceptadvies is besproken met de opdrachtgever. Over de (technische) implicaties van het advies is ook gesproken met experts van CvTE en Cito.”

Commissie Steur

Hoe onafhankelijk is dit advies van de Commissie eigenlijk?

“4.1 Rekenonderwijs”

Commissie Steur, Rekenonderwijs

In deze sectie: geen moeilijke vragen over die referentieniveaus. Nee, voor die referentieniveaus is grote steun in het onderwijsveld, volgens de Commissie. Ik had toch graag enige aandacht gezien voor een heel andere stelling: dat de huidige ellende met de ‘rekentoetsen’ mogelijk volgt uit een valse start die juist met die referentieniveaus voor rekenen is gemaakt. Het is ingewijden immers volstrekt helder dat althans de niveaus 1F, 2F en 3F zijn ingevuld met wat ‘realistisch rekenen’ heet (in andere landen zijn daar andere benamingen voor — reform math, progressivism — of gewoon de vertaling RME: realistic mathematics education (from Holland); voor wie dat bekender in de oren klinkt: dit zijn vormen van wat ook wel ‘het nieuwe leren’ heet. Met andere woorden: wat de wetgever in het beton van de ‘Wet op de referentieniveaus Nederlandse taal en rekenen’ heeft gegoten, komt neer op staatsdidactiek. Staatsdidactiek. Ik herhaal het maar even, want het is ongehoord!

“Rekenen is in het voortgezet onderwijs formeel wettelijk geen vak. Er bestaat dan ook geen bevoegdheid voor het geven van rekenonderwijs. Er is daardoor geen sprake van bevoegde rekendocenten: iedere bevoegde docent mag rekenen geven.”

Commissie Steur, 4.1 Rekenonderwijs, *bevoegdheid, bekwaamheid en ondersteuning*

Wie heeft dit idee de Commissie ingefluisterd? De NVvW, de Nederlandse Vereniging van Wiskundeleraars? Waarom zou een leraar wiskunde geen rekenonderwijs kunnen geven? Is rekenen iets anders dan wiskunde? Kijk nog eens naar de inbreng van [Jan Karel Lenstra](#) in de hoorzitting van 4 december 2013.

Een commissie die gelooft dat rekenonderwijs iets anders is dan wiskundeonderwijs, is vatbaar voor tal van verdere drogredenen uit het omvangrijke circuit van rekenondersteuners. Er is ook wel humor, zoals in deze passage: “Ook in onze panelbijekomsten kwam naar voren dat het bewustzijn pas later doordrong dat rekenonderwijs vraagt om docenten met specifieke kennis en vaardigheden op het gebied van de rekendidactiek.”. Grootmoeder heeft een grote mond, en scherpe tanden, maar roodkapje krijgt maar geen argwaan.

De actualiteit heeft de Commissie op dit punt ingehaald: zie het standpunt van de NVvW zoals verwoord in de [brief](#) van 20 januari aan TK en OCW: het rekenonderwijs kan het best door wiskundeleraars worden verzorgd.

“Digitale afname”

Commissie Steur 4.1 Rekenonderwijs, *digitale afname*

Er zijn probleempjes, maar verder ziet de Commissie er alleen maar heil in. Het verbaast mij. Zeker Don Mellenbergh weet heel goed dat de gebruikte afnametechniek van invloed kan zijn op de resultaten; op zijn minst zou op tafel moeten komen dat digitale afname geen nadelen voor leerlingen oplevert. Dat bewijs is er niet, bewijs van het tegendeel is in ruime mate voorhanden. Tot de benadeling hoort ook de aan deze digitale afname gekoppelde geheimhouding van toetsvragen en de ondoorzichtige ingewikkeldheid van bepalen van de cesuur voor de specifieke gebruikte digitale toets.

Ik zou het zeer gewaardeerd hebben wanneer de Commissie zou hebben geprobeerd te achterhalen waarom eigenlijk door wie is besloten tot digitale afname. En daar een helder antwoord op zou zijn gevonden. In dit advies gepubliceerd.

“Domeinoverstijgende vaardigheden

Bij rekenopgaven in een complexere context zoals die in referentieniveau 3F voorkomen, zijn domeinoverstijgende vaardigheden nodig die soms hebben geleid tot gekunstelde denkstappen.”

Commissie Steur, 4.2 Onvolkomenheden in de toetsen/examens, *domeinoverstijgende vaardigheden*

Helaas, ‘domeinoverstijgende vaardigheden’ bestaan niet. Er is wel zoiets als intelligentie, maar dat is geen vaardigheid.



‘Rekentoetsen’ die ‘domeinoverstijgende vaardigheden’ toetsen, zijn gewoon beroerde intelligentietests. Ik kan me niet voorstellen dat Van de Werf en Mellenbergh hier geen punt van hebben gemaakt, binnen de commissie. En dat moeten ze dan hebben verloren. De Commissie-Steur verkoopt hier amateurpsychologie die nergens op slaat, zoals de [Rekentoetswijzer](#)commissie-Schmidt dat al eerder deed (daar komt het idee van die ‘[domeinoverstijgende vaardigheden](#)’ vandaan).

De oplettende lezer herkent hier de thematiek van ‘vaardigheden van de 21e eeuw’: probleemoplossen, creativiteit. Helaas, dat zijn geen vaardigheden, ze bestaan niet, wat internationale (OECD) en nationale (WRR, Onderwijsraad, bewindslieden OCW) organisaties ook mogen beweren. Probleemoplossen en creativiteit, in niet-triviale verschijningsvormen, zijn gebonden aan kennisdomeinen, aan expertise ([K. Anders Ericsson and others \(2006\)](#)).

Koningin van de ‘domeinoverstijgende vaardigheden’: ‘kritisch leren denken’ De graal van de onderwijsvernieuwers. Helaas, het is een fata morgana, zoals nog eens mooi uitgelegd door [Daniel Willingham \(2007\)](#): *Critical thinking, why is it so hard to teach?*

“*Verskillende definities van contexten*”

Commissie Stuur, 4.2 Onvolkomenheden in de toetsen/examens, *Verskillende definities van contexten*

Voor de Commissie is het idee van rekenvaardigheid toetsen in contexten een vanzelfsprekendheid. Maar dat is het natuurlijk niet. Over het toetsen van rekenvaardigheid met redactiesommen werd voor de oorlog al strijd geleverd. Vanuit psychologische optiek — en die mag juist van deze Commissie worden verwacht, toch? — is het toetsen van rekenvaardigheid iets anders dan die rekenvaardigheid toetsen in contexten. De Commissie beseft dat ergens ook wel, gezien de tweede alinea in deze paragraaf: basale rekenvaardigheden toetsen in ‘eenvoudige’ contexten. Het springende punt is dat de bijzondere opvatting van toetsen van rekenvaardigheid in contexten behoort bij de onderwijsideologische stroming van het situationisme, dat de laatste decennia versmolten is geraakt met het eerder genoemde constructivisme. Als dit waar is — maar de Commissie wil dit niet onderzoeken — dan schrijven de ‘rekentoetsen’ staatsideologische didactiek voor. John Anderson, Herbert Simon en Lynne Reder hebben het eind negentiger jaren onder de aandacht van de internationale psychologische gemeenschap gebracht, zoals in [Anderson, Reder & Simon \(1998\)](#).

Een wetenschappelijke commissie kan niet om de publicaties van Anderson, Reder en Simon over constructivisme en situationisme in het onderwijsveld van rekenen en wiskunde heen. Deze Commissie is zich van geen mogelijk kwaad bewust, maar lijkt dan ook vooral een politieke commissie te zijn.

Bij gelegenheid van de hoorzitting van 4 december 2013 is een uitvoerig stuk over constructivisme en situationisme ingediend: [Wilbrink \(december 2013\)](#). Op de hoorzitting zelf heeft [Wilbrink \(4-12-2013\)](#) uitgelegd hoe het toetsen van rekenvaardigheid in situaties van het dagelijks leven zou kunnen passen bij personeelselectie, maar zeker niet in een eindexamen (de validiteitskwesitie, alweer).

“De inhoudelijke eisen die gesteld zijn aan de rekentoets/de ce's zijn beschreven in de rekentoetswijzers (vo) en examensyllabi (mbo) en gebaseerd op de referentieniveaus. Er wordt onder meer vanuit gegaan dat voor de oplossing van problemen in complexere situaties domeinoverstijgende vaardigheden nodig zijn. Het gevolg daarvan is dat opgaven niet alleen een beroep doen op rekenvaardigheden, maar ook op inzicht in contexten, vaktaal en taalvaardigheid, zoals woordkennis en tekstbegrip. Een psychometrische eis is dat de scores van alle leerlingen afgebeeld moeten worden op dezelfde vaardigheidsschaal. Deze psychometrische eis is echter alleen met de inhoudelijke eisen verenigbaar als de toetsvragen voor geen enkele leerling onbekende woorden en te moeilijke zinsconstructies bevatten. De (eerder afgenomen) toetsen bevatten echter opgaven met woorden en zinsconstructies die moeilijk te begrijpen zijn voor taalzwakke leerlingen en leerlingen die Nederlands niet als thuistaal hebben, ondanks het feit dat een woordenboek mag worden gebruikt. Het gevolg is dat deze opgaven bij deze leerlingen een mix van reken- en taalvaardigheden meten, terwijl ze bij de overige leerlingen alleen rekenvaardigheden meten. Onze commissie is van mening dat de scores van leerlingen in dezelfde schoolsoort/leerweg in vo en mbo afgebeeld moeten worden op een schaal die hetzelfde meet, mits het taalgebruik begrijpelijk is voor alle leerlingen, of dat er bij minder eenvoudig taalgebruik, gewerkt wordt met verschillende schalen voor verschillende groepen leerlingen.

De commissie-Bosker heeft aanbevolen in de rekentoetsen/examens opgaven met eenvoudigere contexten op te nemen. Onze commissie vraagt hierbij aandacht voor het alleen opnemen van die vragen die geen woorden en zinsconstructies bevatten die moeilijk zijn voor taalzwakke leerlingen en leerlingen die Nederlands niet als thuistaal hebben.”

Commissie Stuur, 5.1 Validiteit van de rekentoets/het centraal examen rekenen

Kijk, in sterk contrast met de voorgaande hoofdstukken waait hier ineens wel de psychologische wind. Laten we eens analyseren en onderbouwen wat de Commissie hier stelt. Dit wordt een uitgebreidere analyse, met veel achterliggende literatuur. Ook in de tekst van de Commissie wordt impliciet verwezen naar achterliggende stukken: ik zal proberen links naar die stukken geven.

De Commissie kijkt hier naar de ‘rekentoets’ als meetinstrument. Dat is in ieder geval nuttig als een begin, maar het is niet voldoende. De ‘rekentoets’ is een examen, en de kwaliteitseis daarvoor is dat leraren en leerlingen zich daar doeltreffend op kunnen voorbereiden ([De Groot, 1970](#)). Ik ga in mijn annotatie mee met de analyse van de Commissie; later zal er nog gelegenheid zijn het net wijder uit te werpen, en ook de transparantie — dus ook het rekenonderwijs — erbij te betrekken. Terzijde: commissielid Don Mellenbergh was eind zestiger jaren medewerker van A. D. de Groot in de Afdeling Examenteknieken van de psychologische faculteit, en bekleedde later de leerstoel van De Groot: methodologie.

De Commissie overweegt hier in feite het volgende. De ‘rekentoets’ is gebaseerd op referentieniveaus en [rekentoetswijzers](#); oké, laten we dat althans voorlopig als gegeven nemen, en dan analyseren of een dergelijke ‘rekentoets’ een psychologisch monstrem is, en zo ja, of dat repareerbaar is. Dat is een prima werkwijze, het maakt de vraagstelling misschien wat overzichtelijker.

“Een psychometrische eis is dat de scores van alle leerlingen afgebeeld moeten worden op dezelfde vaardigheidsschaal.”

Commissie Steur, 5.1 Validiteit van de rekentoets/het centraal examen rekenen

Bij een examen(onderdeel) waar je één cijfer voor krijgt moet het zo zijn dat verschillen in cijfers het gevolg zijn van verschillen in één domein van kennis en/of vaardigheid; althans niet een hutsput van van alles en nog wat. De Commissie spitst dat toe op de spanning tussen (verschillen in) rekenvaardigheid en taalbeheersing. Een belangrijk deel van wat de Commissie bedoelt met taligheidsproblemen bij deze ‘rekentoetsen’ is aan de orde in het proefschrift van [Van Silfhout \(2014\)](#), waarover de Commissie kennelijk nog niet heeft kunnen beschikken.

Direct na de behandeling op 31 maart 2010 van de Wet op de referentieniveaus was er discussie over een motie Van Dijk & Van Dijk om taalverzorging mee te beoordelen in eindexamenwerk. Dat zou betekenen dat het cijfer voor bijvoorbeeld wiskunde een mistige afbeelding zou worden van beheersing van wiskunde en van spelling en syntax van het Nederlands (zie ook [Wilbrink, Borsboom & Couzijn, 2011](#)). Dat soort mistigheid is wat de Commissie hier bedoelt. Samengevat in de woorden van Denny Borsboom (vervang meebeoordelen van taalverzorging door meebeoordelen van werken in contexten, van werken met een rekenmachine):

#### Testtheoretische problemen samengevat

Psychometrisch gezien valt te verwachten dat meebeoordelen van een tweede domein zoals taalverzorging leidt tot een lagere validiteit van het examenonderdeel, tot meer ruis — onsystematische variantie — in de scores en tot het ontstaan van bias. Denk aan jongens versus meisjes of leerlingen met Nederlands als tweede taal: systematische groepsverschillen in scores die niet zijn terug te voeren op het oorspronkelijke kennisdomein, zoals natuurkunde, zijn uitingen van bias ofwel partijdigheid. Het valt moeilijk te voorspellen in welke mate deze problemen zich in de praktijk voordoen, maar het is evident dat een maatregel die de psychometrische kwaliteit van examens eigenlijk alleen maar kan verlagen, hooguit genomen mag worden wanneer het echt niet anders kan. Aan deze voorwaarde is in het voorliggende geval niet voldaan.

[Wilbrink, Borsboom & Couzijn, 2011](#), blz. 8

Jammer is dat de Commissie er niet nadrukkelijk op wijst dat er meer mogelijke oorzaken kunnen zijn voor schending van de eenduidigheid van de vaardigheidsschaal. Wie de verwickelingen rond de ‘rekentoets’ een beetje heeft gevolgd, kan ze wel opnoemen: nadrukkelijk *niet* toetsen op algoritmische rekenvaardigheid en vlot rekenen ([Rekentoetswijzer](#)commissie-Schmidt), de mogelijkheid geven om te rekenen met de digitale rekenmachine ([Wilbrink 2012 LinkedIn](#)), en vooral ook mentale overbelasting bij contextopgaven ([Wilbrink & Kirschner, 2014](#)).

“Deze psychometrische eis is echter alleen met de inhoudelijke eisen verenigbaar als de toetsvragen voor geen enkele leerling onbekende woorden en te moeilijke zinsconstructies bevatten.”

Commissie Steur, 5.1 Validiteit van de rekentoets/het centraal examen rekenen

Laten we vooral zorgvuldig formuleren. Het woordje ‘alleen’ suggereert dat uitpoetsen van het aangeduide probleem van taligheid, een valide ‘rekentoets’ oplevert. Dat is natuurlijk niet zo. Er zijn tal van bedreigingen van rekentoetsvaliditeit, zie [Wilbrink, Hulshof & Pfaltzgraff \(2012\)](#) en [Wilbrink & Kirschner, 2014](#). Misschien is dit wel een geschikt moment om erop te wijzen dat Wilbrink c.s. (2012) ernstige twijfels aan validiteit van de ‘rekentoetsen’ hebben geuit, dat het Cito daar niet inhoudelijk op heeft gereageerd (ik laat de tweet van CEO Marten Roorda hier even terzijde) en dat het CvTE in zijn reactie van de hand van [Kastelein \(2012\)](#) louter op correcte procedures wijst, wat heeft niets met validiteit heeft te maken. Het probleem met een en ander is dat er een ernstige klacht ligt over ontbrekende validiteit van de ‘rekentoetsen’, waarop geen adequaat antwoord is gegeven. Het Nederlandse publiek weet het dus niet. Ondertussen hebben beide instellingen de staatssecretaris verklaard dat de ‘rekentoetsen’ wel degelijk valide en betrouwbaar zijn. Punt. Dus zonder ondersteunend empirisch bewijs. Het CvTE heeft december 2014 in een briefing van de onderwijscommissie van de Tweede Kamer nogmaals gesteld dat de ‘rekentoetsen’ valide zijn. Wederom zonder daar de nodige onderbouwing voor te geven.

Is die validiteit van belang, dan? Wat is er eigenlijk mee bedoeld? Ik verwijs naar de behandeling van dit onderwerp in de [Standards](#). Zonder validiteit heeft een toets of een test absoluut geen bestaansrecht, en is ieder gebruik van toetsscores misbruik, dat door de rechter nietig zal worden verklaard.

In deze sectie 5.1 over validiteit snijdt de commissie het thema in feite nauwelijks aan. Vergelijk daartoe de onderwerpen in het hoofdstuk over validiteit in de [APA Standards](#). Tot de hamvragen behoren zeker de vraag naar wat rekenen en rekenvaardigheid is (wat hebben die contexten er eigenlijk mee te maken?), en of leerlingen zich doeltreffend op deze ‘rekentoetsen’ kunnen voorbereiden (wat niet het geval is).

Er is nadrukkelijk en expliciet door [Wilbrink c.s. 2012](#) op gewezen dat er op meerdere punten sprake is van tekortschietende of ontbrekende validiteit van de ‘rekentoetsen’. De Commissie was van deze kritiek op de hoogte, evenals van het ontbreken van een adequaat antwoord van het Cito en van het CvTE op deze kritiek. De Commissie laat de validiteit van de ‘rekentoetsen’ met de tekortschietende behandeling in deze sectie 5.1 in feite verder bungelen, het aan politici overlatend om te kiezen tussen de expertise van Wilbrink c.s. en die van de Commissie op het onderdeel validiteit van de ‘rekentoetsen’. Een eveneens glasheldere kritiek op ontbrekende validiteit van de ‘rekentoetsen’ is gegeven door [Wilbrink en Kirschner](#), en ook deze kritiek, hoewel aan de Commissie bekend gemaakt, is door de Commissie genegeerd.

Per saldo weten bewindslieden, kamerleden en het Nederlandse publiek dus niet, met dit commissierapport in de hand, of de geuite kritiek op de ontbrekende validiteit van de ‘rekentoetsen’ hout snijdt of niet, althans in de ogen van deze Commissie. Toch gaat de Commissie aan de slag met cesuurbepaling etcetera, voor een toets van onbestemde validiteit. Wellicht ten overvloede: ook het *gebruik* van de toetsresultaten is een validiteitskwestie. En laten zakken of slagen is ongeveer het belangrijkste *gebruik* dat van de ‘rekentoets’-resultaten valt te maken.

## “5.2 Cesuur”

Commissie Steur, 5.2 Cesuur

De eerste alinea van deze heel erg lange paragraaf: let op allerlei woordjes en uitspraken die zekerheden en kwaliteit suggereren waar ze er niet zijn. Het irritante gedoe over die vaardigheidsschaal hoort daar ook bij: de term suggereert veel en veel meer dan hij waar kan maken. De commissie heeft hier overigens te maken met een erfenis van het Cito, dat bij de PPO die vaardigheidsschalen gebruikt om vergelijkingen van 19989 met 2004 te kunnen maken, etcetera; niets mis mee. Bij examens is dat anders, omdat psychometrische technieken helder uitgelegd moeten kunnen worden naar leraren en leerlingen toe. Hoe doe je dat met deze vaardigheidsschalen? Door het te versimpelen. Is dat niet een rad voor ogen draaien, dan? Mogelijk. Ben ik unfair? Ik dacht het niet, omdat overduidelijk is dat de contextopgaven in de ‘rekenoetsen’ ook of zelfs vooral verschillen in intellectuele capaciteiten registreren en dus juist *geen* vaardigheden. Mentale belastbaarheid (werkgeheugen) hoort daar ook bij: geen vaardigheid, maar een psychologisch gegeven waar leerlingen op verschillen ([Wilbrink & Kirschner, 2014](#)).

Ik wil dus niet meegaan met die vaardigheidsschaal-hocus-pocus. Een echt passend alternatief is er niet, behalve recht-voor-zijn-raap ‘percentage goed’; als *next best* komt ‘beheersing’ in aanmerking: ‘vaardigheid’ geeft aan dat deze nadrukkelijk is verworven, ‘beheersing’ laat iets meer in het midden hoe die beheersing is verkregen/ontstaan. Het is niet alleen maar een kwestie van zuiver taalgebruik: spreken over een vaardigheidsschaal suggereert telkens ook dat de vaardigheid op valide wijze is bepaald. En dat is en open wond, excuus, kwestie.

Ik zal in het vervolg fud spreken over **beheersing** van het soort vragen dat in ‘rekenoetsen’ wordt gesteld (hoeveel procent van dat soort vragen maakt de betreffende leerling ‘goed’). Ook als dat onzinnige vragen zijn, vragen die eigenlijk in een IQ-test thuishoren, of die op capaciteitsverschillen in het werkgeheugen aanspraak doen. Laat ik een eenvoudig voorbeeld van de gekkigheid van de ‘rekenoetsen’ geven, iets waar het Cito en het CvTE nog geen antwoord op hebben gegeven. Leerlingen die bij wijze van spreken een foutloze wiskunde-olympiade maken en in de top eindigen, kunnen zij een foutloze ‘rekenoets’-3F maken? De vraag stellen is haar beantwoorden. Een ander, toch in de grond van de zaak ook heel eenvoudig, voorbeeld: wanneer vwo-leerlingen 70% voldoende scores op de ‘rekenoets’-3F, en havisten maar 30%, wat denkt u: is dat een vaardigheidsverschil, of een verschil in intellectuele capaciteiten? Laten we alleen over vaardigheid spreken wanneer het inderdaad gaat om iets dat geleerd, geoefend, of getraind kan worden. Als de ‘rekenoetsen’ een vaardigheid zouden toetsen, zouden toppers uit de wiskundeolympiade probleemloos 100 procent goed scoren. Quod non.

“De referentiecensuur is bepaald met een standaardsettingsprocedure door een groep beoordelaars (docenten uit de betrokken schoolsoorten en opleidingen, rekenexperts en toetsdeskundigen) onder leiding van CvTE en Cito11 in 2013. Het gaat hierbij om een inhoudelijke inschatting van wat leerlingen ten minste moeten kennen en kunnen.”

noot 11: Deze procedure is beschreven in: [Regeling omzetting scores in cijfers centrale examens en rekenoets VO 2014](#)

Commissie Steur, 5.2 Cesuur, *Standaardsettingsprocedure*

En dat is ongeveer alles wat de Commissie over het wonder van de cesuur voor deze ‘rekenoetsen’ heeft te melden. Alsof alles wat Cito en CvTE hier hebben verricht, goed verricht is. En dat is dus evident niet het geval, dan zouden er immers geen commissies Bosker en Steur nodig zijn geweest. Ongelooflijk, maar waar. Wat valt er dan meer over te zeggen? Wat mij betreft begint dat met de samenvatting van de [CvTE-aanpak](#) zoals gegeven in de [Wiskunde-brief 620](#) met kritische kanttekeningen van [Wilbrink](#): ‘De lat op de rekenoets. Willekeur?’ in de volgende [Wiskunde-brief 621](#). Uitvoerige behandeling: ‘Normering van de rekenoetsen-2F en -3F. Alternatieven voor, en kritische analyse van de CvE-opstelling’ [Wilbrink, website](#).

Voorlopige conclusie na deze alinea’s: vaardigheid is een term die de lading niet dekt, vervang deze door beheersing; de cesurbepaling is abacadabra, in plaats van daarmee verder te modderen, wat de Commissie en de bewindslieden voorstellen, zou hier radicaal ingegrepen moeten worden en bijvoorbeeld een compromismethode ([Hofstee \(1983\)](#)) uitgewerkt en uitgevoerd worden (maar niet dan nadat de vraag is beantwoord naar wat rekenvaardigheid en rekenonderwijs moeten zijn). Nou ja, de koninklijke weg is natuurlijk het voorstel van [Jan Karel Lenstra \(4 december 2013\)](#) te volgen.

“Uit andere situaties is bekend dat leerlingen beter presteren als een examenonderdeel meetelt voor diplomering.”

Commissie Steur, 5.2 Cesuur, *standaardsettingsprocedure*

Dit is een uitspraak waar toch wel een kanttekening bij moet worden geplaatst. Immers, de verschillen tussen bijvoorbeeld havisten en vwo-ers op de ‘rekenoets’-3F zijn zo enorm, dat moeilijk valt vol te houden dat ze iets hebben te maken met ‘nog niet echt meetellen voor het examen’. In de overleggen van de Tweede Kamer met de bewindslieden is dit argument bij herhaling gebruikt als bezweringsformule dat het allemaal wel goed komt. Heel ongelukkig, dit. Vergeet dus deze uitspraak van de Commissie. Hij is niet gedaan. Nooit gezien. Zand erover.

In de betreffende alinea buitelen de zinnen, de uitspraken over standaarden die zich nog zouden moeten ‘zetten’, over elkaar heen. Trek u er niets van aan, er staat eigenlijk niets verstandigs in. Vergeet heel deze alinea. Uitgevinkt. Schone lei.

“Die onzekerheid maakt het lastig om de slaagcesuur te leggen op het niveau van de referentiecensuur. We vinden het daarom vooralsnog onverstandig om de slaagcesuur te leggen op het niveau van de referentiecensuur.”

Commissie Steur, 5.2 Cesuur, *ambitie en empirie*

Dit klinkt allemaal heel gewichtig, alsof het volstrekt duidelijk is wat slaagcesuur en wat referentiecensuur is. En dat is het niet. Op zijn Latijns: quod non.

“We vinden het daarom vooralsnog onverstandig om de slaagcesuur te leggen op het niveau van de referentiecesuur. Leerlingen zouden dan de dupe worden van onvolkomenheden in de toets of het onvoldoende voorbereid zijn op het examen. Aan de andere kant willen we de cesuur ook niet veel lager leggen, omdat er dan onvoldoende uitdaging van uit zou gaan. Daarom denken we dat vastgehouden zou moeten worden aan de norm, zoals die door CvTE met de referentiecesuur is gelegd in de standaardsettingsprocedure.”

Commissie Steur, 5.2 Cesuur, *ambitie en empirie*

Merk op hoe ijzersterk hier de bewijsvoering is. Moet ik daar nog meer over zeggen? De Commissie weet het gewoon niet, en wil de bewindslieden niet tot last zijn. Of zo iets.

“Met de opname van rekenen in het eindexamen vo en het examen mbo worden de exameneisen zwaarder. Enige stijging van het percentage gezakten is op grond daarvan te billijken.”

Commissie Steur, 5.2 Cesuur, *ambitie en empirie*

Wat de commissie hier doet is onbegrijpelijk: volkomen uit de lucht gegrepen zeggen dat (1) exameneisen zwaarder worden en (2) het daarom te billijken zou zijn dat er meer kandidaten zakken voor het examen dan anders het geval zou zijn. Advocaten kunnen dit vast noteren. Ik ken geen vindplaats in de wet of in de wetsbehandeling waarin sprake is van een opzet om eindexamens zwaarder te maken. En al zou dat zo zijn, omdat er gewoon een extra onderdeel is met een risico om daar een riskante uitslag op te scoren, dan is het niet te billijken dat examenkandidaten daar de dupe van worden. Examenregelingen zijn mensenwerk, tegenwoordig vooral werk van politici trouwens, en kunnen aan veranderde omstandigheden worden aangepast. Zo is dat.

“Door van te voren een maximum toename van het percentage gezakten als gevolg van de rekentoets vast te stellen, wordt voorkomen dat het onredelijk veel leerlingen betreft. De commissie stelt voor het eerste jaar van het voorstel een maximum afwijking van 5%-punten voor in vergelijking met het laagste percentage gezakten per schoolsoort/leerweg in de drie jaar voorafgaand aan het meetellen van de rekentoets/het ce-rekenen voor diplomering.”

Commissie Steur, 5.2 Cesuur

De Commissie solt hier met mensen. Het is naar mijn bescheiden mening zelfs niet aan politici om iets zoals dit voor te stellen of te regelen. Het is echt over de top. Ik ben geen jurist, maar ik voel dat hier iets wordt voorgesteld dat in strijd is met in het algemeen rechtsgevoel levende beginselen van behoorlijk bestuur. Misbruik van bevoegdheden waartoe deze niet gegeven zijn, lijkt me een goede eerste kandidaat.

Het enig redelijke dat hier gesteld zou kunnen worden: een ‘rekentoets’ als toegevoegd bijzonder onderdeel van de eindexamens, mag er niet toe leiden dat er meer kandidaten voor het eindexamen zakken dan anders het geval zou zijn geweest. Dat lijkt me ook in juridisch opzicht juist: het eindexamen, bestaande uit schoolonderzoek en centraal schriftelijk, blijft het eindexamen. Dat er om bijzondere redenen een bijzonder onderdeel op een bijzondere manier is toegevoegd, kan daar niets aan veranderen. Ik kan mij niet anders voorstellen dan dat de rechter zakken voor een eindexamen louter en alleen vanwege het resultaat op de ‘rekentoets’ nietig verklaart. Dat zullen we dan af moeten wachten, als Kamer en bewindslieden dit boven de hoofden van eindexamenkandidaten willen laten hangen.

De Commissie lijkt te denken dat het bij de ‘rekentoetsen’ om een regulier eindexamenonderdeel gaat, maar dat is evident niet het geval, in geen enkel opzicht. Dat roept de vraag op: heeft de Raad van State zich ooit uitgesproken over deze rekentoetsregelingen? Over de Wet op de referentieniveaus Nederlandse taal en rekenen ongetwijfeld wel, maar ja, is mijn dossier niet al omvangrijk genoeg?

- <http://www.steunpuntaalenrekenen.nl/informatie/kaders-en-regels/wet-referentieniveaus>
- [http://www.steunpuntaalenrekenen.nl/sites/default/files/Advies%20raad%20van%20state%20op%20wetsvoorstel%20t&r%20vo\\_0.pdf](http://www.steunpuntaalenrekenen.nl/sites/default/files/Advies%20raad%20van%20state%20op%20wetsvoorstel%20t&r%20vo_0.pdf)
- <http://www.steunpuntaalenrekenen.nl/sites/default/files/Wet%20referentieniveaus%20Nederlandse%20taal%20en%20rekenen.pdf>
- <http://www.steunpuntaalenrekenen.nl/sites/default/files/Memorie%20van%20toelichting%20op%20wetsvoorstel.pdf>

“Onze commissie vindt het belangrijk dat er druk blijft om de rekenprestaties te verbeteren.”

Commissie Steur, 5.3 Slaag/zak-regeling

Waarom neemt de Commissie deze uitspraak in het rapport op? Wat zijn dit voor fratsen? Om de bewindslieden te gerieven?

“Onze commissie hecht aan transparantie voor vervolgonderwijs over de rekenvaardigheid van leerlingen.”

Commissie Steur, 5.4 Transparantie

De Commissie schijnt onkundig te zijn van het feit dat de ‘rekentoetsen’ nadrukkelijk *niet* toetsen op algoritmische rekenvaardigheden (zie de [rekentoetswijzers-F](#)), wat toch voor het ontvangende ho ongeveer het enige is dat ècht van belang is. Dump die referentieniveaus-F in de Hofvijver, het ho zal er geen traan om laten ([Wilbrink c.s. 2012](#)). Dit is natuurlijk ook een cruciaal punt voor de validiteit van de ‘rekentoets’-3F: is dit het rekenen dat het ho verwacht van zijn aankomende studenten?

De tekst van deze paragraaf 5.4, evenals die van de voorgaande 5.3 over de zak-slaagregeling, is opvallend onprofessioneel geformuleerd en overladen met details uit regelingen. Over transparantie gesproken.

“5.5 Focus op resultaten”

Commissie Steur, 5.5 Focus op resultaten

Is deze slotparagraaf geschreven door de PR-mensen van OCW?

“Onze commissie adviseert om de rekentoets in vo en het ce-rekenen voor mbo-4 vanaf 2015-2016 mee te laten tellen bij de examinering en voor mbo-2 en mbo-3 een jaar later (vanaf 2016-2017), en dus de invoering daarvan niet (nog eens) uit te stellen.”

Commissie Steur, 6. Advies

Dit hebben we eerder gezien, in de aanbiedingsbrief van het rapport van de Commissie-Bosker: “Voorgaande neemt niet weg dat de commissie voorstelt om ernaar te streven, de invoering van de rekentoetsen en coe’s volgens planning door te laten gaan”. Het is een zinsnede die is toegeschreven op gebruik door bewindslieden. Minister Bussemaker, in het vragenuurtje direct na vrijgeven van het rapport-Bosker, citeerde deze passage tot driemaal toe.

“Samenvattend stellen wij het volgende voor:”

Commissie Steur, 6. Advies

We kennen ondertussen de trits van jaar tot jaar oplopende lapmiddelen. ‘Lapmiddelen’, want ze werken op de symptomen van de niet door de Commissie gediagnosticeerde ziekte: rekenonderwijs dat niet meer minimaal adequaat is. Om de geloofwaardigheid te vergroten doet de Commissie er nog een overigens niets toevoegende figuur van de ‘Vaardigheidsschaal rekenen’ bij. Dit is een bekende truc: als de boodschap vaag is, voeg er een kleurendiagram of iets dergelijks aan toe, en de geloofwaardigheid springt omhoog.

“*flankerende maatregelen*”

Commissie Steur,

Geloof het of niet (in het laatste geval: sla het commissie-rapport open deze bladzijden), maar de Commissie heeft bladzijde 28 tot 34 nodig om de voorgestelde *flankerende maatregelen* op te sommen. Over transparantie gesproken. Ik besteed er verder geen annotaties meer aan, deze ‘rekentoetsen’ horen niet in welk eindexamen dan ook een plaats te krijgen.

“Geraadpleegde experts”

Commissie Steur, Geraadpleegde experts

Een opmerkelijke lijst: iedereen heeft direct of indirect belang bij de ‘rekentoetsen’ in hun huidige vorm. Mogelijk met uitzondering van Roel Bosker, de voorzitter van de vorige commissie, maar hij is ook voorzitter van een wetenschappelijke raad van het Cito. Heeft de Commissie-Steur geen belangstelling voor tegenspraak gehad?

## Verwijzingen

AERA, APA & NCME (1999). **The Standards for Educational and Psychological Testing.**

<http://www.apa.org/science/programs/testing/standards.aspx> Nu in de 2014-editie. De voorgaande editie is van 1996. Ik heb daarvan oa Validity en Reliability gescand 32 Mb

[http://www.ben-wilbrink.nl/standards\\_1999\\_1-6\\_validity\\_reliability\\_etc.pdf](http://www.ben-wilbrink.nl/standards_1999_1-6_validity_reliability_etc.pdf) For personal use only! Voor mijn fans: voor het NIP-1988 hfdst over toetsen in het onderwijs is o.a. Ben Wilbrink adviseur geweest ;-). De Standards zijn helaas niet free access. Ook niet in de KB (Royal Library of the Netherlands). De editie-2014 nog niet in UB Leiden. Evers ea (subcie vd Kommissie Testaangelegenheden Nederland KOTAN) (1988) Richtlijnen. Ned. Instituut van Psychologen. [Hoofdstuk] 8. Toetsgebruik in het onderwijs. [http://www.ben-wilbrink.nl/Richtlijnen\\_Toetsgebruik\\_in\\_het\\_onderwijs.pdf](http://www.ben-wilbrink.nl/Richtlijnen_Toetsgebruik_in_het_onderwijs.pdf) uit Richtlijnen voor ontwikkeling en gebruik van psychologische tests en studietoetsen

Anderson, J. R., Reder, L. M. & Simon, H. (1998). Radical constructivism and cognitive psychology. In D. Ravitch (Ed.) *Brookings papers on education policy 1998* (227-278). Washington, DC: Brookings Institute Press. [pdf](#) [Ook beschikbaar via JSTOR, registratie vereist om free online te kunnen lezen] Een betere versie in pdf: <http://goo.gl/6ULfY4>

Job M. Cohen (1981). **Studierechten in het wetenschappelijk onderwijs.** Tjeenk Willink. Proefschrift Rijksuniversiteit Leiden. Nog steeds een unieke bron als het gaat om rechten van leerlingen en studenten bij toetsen en examens, ook internationaal gezien. [http://www.ben-wilbrink.nl/Cohen\\_1981\\_Studierechten\\_inh\\_3.1\\_samenv.pdf](http://www.ben-wilbrink.nl/Cohen_1981_Studierechten_inh_3.1_samenv.pdf)

Commissie-Steur (december 2014). Doordacht doorzetten naar een hoger niveau. Advies van de commissie verantwoorde invoering rekentoets in vo en mbo. [pdf](#)

CvTE (december 2012). Normering met een vaardigheidsschaal, zoals bij Rekentoets VO en COE [http://www.cve.nl/document/normering\\_met\\_een](http://www.cve.nl/document/normering_met_een)

College voor Toetsen en Examens (11 december 2014). Notitie **Acties naar aanleiding van de aanbevelingen van de commissie**

**Bosker** (bijlage bij de kamerbrief van 17 december) [pdf](#)

- A.D. de Groot (1970). Some badly needed non-statistical concepts in applied psychometrics. *Nederlands Tijdschrift voor de Psychologie en haar Grensgebieden*, 26, 360-376. Didakometrisch en Psychometrisch Onderzoek, juni 1970.  
<http://www.benwilbrink.nl/publicaties/70degroot.htm>
- Willem K. B. Hofstee (1983). The Case for Compromise in Educational Selection and Grading. In Scarvia B. Anderson and John S. Helmick (Eds) (1983). On educational testing. San Francisco: Jossey-Bass Publishers. p. 109-127.  
[http://benwilbrink.nl/publicaties/83hofstee\\_compromise.htm](http://benwilbrink.nl/publicaties/83hofstee_compromise.htm)
- Jan Karel Lenstra (4 december 2014). Position paper hoorzitting onderwijscommissie Tweede Kamer over de rekentoetsen. Ophalen via <http://www.wiskundebrief.nl/654.htm#1> of direct <http://wisforum.nl/hoorzitting/lenstra.pdf>
- Nederlandse Vereniging van Wiskundeleraren (20 januari 2015). Notitie rekenen. Brief aan de Tweede Kamer en aan OCW.  
<https://www.nvvw.nl/19628/nvvw-rekenen>
- 31332-36 Brief regering d.d. 17 december 2014 - S. Dekker, staatssecretaris van Onderwijs, Cultuur en Wetenschap **Stand van zaken invoering referentieniveaus taal en rekenen in vo en mbo** [download](#); daar ook de bijlagen op te halen, namelijk
- Acties naar aanleiding van de aanbevelingen van de commissie Bosker [pdf](#)
  - Advies van de commissie verantwoorde invoering rekentoets in vo en mbo [pdf](#)
- 13 juni 2014. Kamerbrief over voortgangsrapportage invoering referentieniveaus taal en rekenen 2014 [ophalen](#) met bijlagen:
- Prestaties basisonderwijs taal en rekenen 2013. Rapport van Cito over de prestaties van basisschoolleerlingen van groep 8 uit 2013 in termen van het referentiekader taal en rekenen. [ophalen](#)
  - Voortgangsrapportage invoering referentieniveaus taal en rekenen 2014. Invoering centrale toetsing en examinering referentieniveaus nederlandse taal en rekenen. [ophalen](#)
- E.C.E. de Kler (18 december 2014). Convocatie algemeen overleg rekentoets in vo en mbo. [pdf](#)
- Valerie F. Reyna, Wendy L. Nelson, Paul K. Han & Nathan F. Dieckmann (2009). How Numeracy Influences Risk Comprehension and Medical Decision Making. *Psychological Bulletin*, 135, 943-973.  
Open access <http://www.ncbi.nlm.nih.gov/pubmed/19883143>
- Frank van der Schoot (2008). Onderwijs op peil? Een samenvattend overzicht van 20 jaar PPO. Stichting Cito. [pdf ophalen](#)
- R. S. Siegler, G. J. Duncan, P. E. Davis-Kean, K. Duckworth, A. Claessens, M. Engel, M. I. Susperreguy & M. Chen (2012). Early predictors of high school mathematics achievement. *Psychological Science*, 23, 691-697.  
concept: <http://www.psy.cmu.edu/~siegler/Siegler-et-al-inpressPsySci.pdf>
- Gerdineke van Silfhout (2014). **Fun to read or easy to understand? Establishing effective text features for educational texts on the basis of processing and comprehension research.** Proefschrift Utrecht. Landelijke Onderzoekschool Taalwetenschap.  
<http://www.lotpublications.nl/fun-to-read-or-easy-to-understand-fun-to-read-or-easy-to-understand-establishing-effective-text-features-for-educational-texts-on-the-basis-of-processing-and-comprehension-research>
- Ben Wilbrink (december 2014). Situationisme, cognitieve psychologie, en rekentoets. Ingediend voor Hoorzitting onderwijscommissie Tweede Kamer 4 december 2013.  
[http://benwilbrink.nl/benwilbrink\\_rekentoets\\_hoorzitting\\_4-12-2013.pdf](http://benwilbrink.nl/benwilbrink_rekentoets_hoorzitting_4-12-2013.pdf)
- Ben Wilbrink (4 december 2014, mondelinge presentatie) Over het ontwerp van vragen die rekenvaardigheid in situaties van het dagelijks leven pretenderen te toetsen. Hoorzitting onderwijscommissie Tweede Kamer.  
[http://www.benwilbrink.nl/projecten/hoorzitting\\_4-12-2013\\_situationisme.htm#rondetafelgesprek](http://www.benwilbrink.nl/projecten/hoorzitting_4-12-2013_situationisme.htm#rondetafelgesprek)
- Ben Wilbrink (2014). Al die teksten in onze eindexamens, waar komt dat fenomeen toch vandaan? Kan Sander Dekker precies 75 L zand in een kruwagen scheppen? *Van Twaalf Tot Achttien, Vakblad voor Voortgezet Onderwijs*, 24, 28-30.  
[http://benwilbrink.nl/projecten/hoorzitting\\_4-12-2013\\_situationisme.htm#concept](http://benwilbrink.nl/projecten/hoorzitting_4-12-2013_situationisme.htm#concept)
- Ben Wilbrink (). De lat op de rekentoets. Willekeur? Wiskunde-brief 621  
<http://www.wiskundebrief.nl/621.htm#1>
- Wiskunde-brief 620 Normering rekentoets  
<http://www.wiskundebrief.nl/620.htm#4>
- Ben Wilbrink (website). Normering van de rekentoetsen-2F en -3F. Alternatieven voor, en kritische analyse van de CvE-opstelling  
[http://benwilbrink.nl/projecten/rekentoets\\_normering.htm](http://benwilbrink.nl/projecten/rekentoets_normering.htm)
- Ben Wilbrink, Denny Borsboom en Michel Couzijn (2010). Spelling, grammatica en interpunctie meebeoordelen op eindexamens? Tweede Kamer motie van de leden Jan Jacob van Dijk en Jasper van Dijk. *Examens, Tijdschrift voor de Toetspraktijk*, 7 #3, 5-9.  
[http://www.ben-wilbrink.nl/EXAMENSCOMPLEET\\_sept\\_2010.pdf](http://www.ben-wilbrink.nl/EXAMENSCOMPLEET_sept_2010.pdf)
- Taalverzorging meebeoordelen bij eindexamens is eenzelfde probleem als er een rekentoets in opnemen. Daarom is dit artikel over spellen ook relevant voor de rekentoets.
- Ben Wilbrink & Joost Hulshof (2011). De wet, het rekenen, en de rekentoets in de eindexamens havo/vwo. Invoering van de rekentoets havo/vwo in 2014. *Examens, Tijdschrift voor de Toetspraktijk*, 8 #3, 18-22.

<http://www.ben-wilbrink.nl/ex-11-3.pdf>

Ben Wilbrink, Joost Hulshof & Henk Pfaltzgraff (2012). Examens, Tijdschrift voor de Toetspraktijk, 9 #3, 26-31.

[http://www.ben-wilbrink.nl/wilbrink\\_hulshof\\_pfaltzgraff\\_2012\\_De\\_rekentoetsen\\_3F\\_zijn\\_niet\\_valide.pdf](http://www.ben-wilbrink.nl/wilbrink_hulshof_pfaltzgraff_2012_De_rekentoetsen_3F_zijn_niet_valide.pdf)

Ben Wilbrink & Paul Kirschner (2014). Met die contexten toetsen wij geen rekenen maar vermogen tot (zelf)ontdekkend leren. blog

[http://benwilbrink.nl/publicaties/14wilbrink\\_kirschner\\_rekentoetsen.htm](http://benwilbrink.nl/publicaties/14wilbrink_kirschner_rekentoetsen.htm)

Ben Wilbrink & Paul Kirschner (2014). Rekentoetsen-met-contexten toetsen geen rekenen. Poster [KNAW-seminar](#) 'Rekenen: het onderwijs, het vak en de toets' 30 juni 2014

[http://www.ben-wilbrink.nl/14Wilbrink\\_Kirschner\\_poster.pdf](http://www.ben-wilbrink.nl/14Wilbrink_Kirschner_poster.pdf)

Daniel Willingham (2007). Critical thinking, why is it so hard to teach? *American Educator*, summer 2007, 8-19.

[http://www.aft.org/sites/default/files/periodicals/Crit\\_Thinking.pdf](http://www.aft.org/sites/default/files/periodicals/Crit_Thinking.pdf)

Willem Witteveen (2014, posthuum). **De wet als kunstwerk. Een andere filosofie van het recht.** [Boom](#)

Robert Zwitser & Anton Béguin (2011). Gaat meefout bij de rekentoets slachtoffers maken? Het effect van de rekentoets op het percentage misclassificaties. *Examens, Tijdschrift voor de Toetspraktijk*, 8 #4, 23-26.

Met repliek van Wilbrink & Hulshof op blz. 26.

[http://www.ben-wilbrink.nl/zwitser\\_beguin\\_2011\\_wilbrink\\_hulshof.pdf](http://www.ben-wilbrink.nl/zwitser_beguin_2011_wilbrink_hulshof.pdf)

Voor Commissie-Schmidt rekentoetswijzers 2F en 3F (ook Commissie-Van de Craats rekentoetswijzer 3S trouwens) zie [website SLO](#)

Voor geheimhouding, zie literatuur en casuïstiek

<http://www.benwilbrink.nl/projecten/geheimhouding.htm>.

Regeling omzetting scores in cijfers centrale examens en rekentoets VO 2014. Verkregen op 3 oktober 2014, van

[http://wetten.overheid.nl/BWBR0035058/geldigheidsdatum\\_03-10-2014](http://wetten.overheid.nl/BWBR0035058/geldigheidsdatum_03-10-2014)

[Kijk in dit document naar de paragraaf **Procedure rekenen**, B.W.]