|  |
| --- |
| Bureau van het CvTE |
| Muntstraat 7  3512 ET Utrecht  Postbus 315  3500 AH Utrecht  Nederland |
| www.cvte.nl |
|  |
|  |
|  |
|  |
|  |
| Onze referentie |
| CvTE 18.01138 |
| Uw referentie |
|  |

|  |
| --- |
| > Retouradres Postbus 315 3500 AH Utrecht |
| Ministerie van OCW  t.a.v. de Minister voor Basis- en Voortgezet Onderwijs en Media  de heer Drs. A. Slob  Postbus 16375  2500 BJ DEN HAAG |
|  |
| |  |  | | --- | --- | |  |  | | Datum | 29 juni 2018 | | Betreft | Beantwoording vragen uit de blog van Kneyber | |

Geachte heer Slob,

Recent heeft u het College voor Toetsen en Examens (CvTE) verzocht bij te dragen aan de beantwoording van vragen van het lid Kwint (SP) aan uw adres, over de klachten over het centraal eindexamen Frans 2018-I (2018Z10990).

Aanleiding voor de vragen van het lid Kwint was de blog van René Kneyber over hoe het College voor Toetsen en Examens (CvTE) compenseert bij de normering bij examens.

Hieronder geeft het CvTE puntsgewijs een reactie op het artikel van de heer Kneyber.

*De heer Kneyber:*

*Vorige week schreef ik een column in Trouw over de examenperikelen rond het eindexamen Frans, en over de rechtszaak die is aangespannen door een scholiere. Kort gezegd: in het correctiemodel zat een fout, waardoor docenten een goed antwoord fout moesten rekenen. Het College voor Toetsen en Examens (CvTE) wilde de fout in het voorschrift echter niet corrigeren, maar deed dat uiteindelijk na veel maatschappelijke ophef pas op de dag dat de normering werd vrijgegeven. Echter, de gedupeerde leerlingen kregen van het CvTE niet 0,2 punten erbij (wat de vraag waard was) maar 0,1 punt. Precies de 0,1 punt-verschil die deze scholiere nodig had om te slagen. Zij spande een kortgeding aan, en kreeg geen gelijk, omdat het CvTE haar eigen procedure goed had gevolgd. Dit vonnis werd begin dit jaar vernietigd door de Hoge Raad.*

Reactie CvTE:

De compensatie voor de vraag in het examen Frans vwo vorig jaar vond plaats ná blijvende onenigheid tussen docenten en de Examenlijn van het CvTE. Omdat de tijd voor aanvullingen op dat moment verstreken was, volgde toen een besluit te compenseren via de N-term. Deze gang van zaken is door de voorzieningenrechter in Utrecht in orde bevonden. Dit vonnis is later door het Hof Arnhem-Leeuwarden [niet door de Hoge Raad] vernietigd, maar niet om inhoudelijke redenen.

*RK: Het bizarre aan dit verhaal is dat wanneer het CvTE gewoon de fout had toegegeven en het correctiemodel had aangepast deze vrouw 0,2 punt erbij had gekregen en geslaagd zou zijn. Nu kreeg ze slechts 0,1 punt en niet 0,2 punten omdat, zo stelden ze in de rechtszaak, er anders teveel kinderen zouden slagen.*

CvTE: Het CvTE streeft zo lang mogelijk naar overeenstemming tussen docenten en Examenlijn als het gaat om discussies over uiteenlopende visies. Er komt echter een moment waarop de tijd van aanvullingen voorbij is en er een knoop moet worden doorgehakt. Voor het uitbrengen van een aanvulling is het dan te laat. Daarom is gekozen voor compensatie bij de normering, een gebruikelijke oplossing in dat soort gevallen.

De essentie van de normeringssystematiek is dat leerlingen jaar in, jaar uit aan dezelfde eisen moeten voldoen, niet dat een bepaald aantal leerlingen (teveel of te weinig) slaagt. Door de beproefde normeringssytematiek ligt de ‘lat’ ieder jaar even hoog. Dat biedt mede de garantie dat een diploma van een bepaalde schoolsoort in Nederland over de jaren dezelfde waarde heeft.

*RK: Om de regeling die ze hiervoor gebruikten juridisch beter dicht te timmeren heeft het CvTE vorige week, met instemming van de minister haar rekenmethode voor dit soort problemen gepubliceerd.*

CvTE: De heer Brussee wees er tijdens de zitting van de zaak in Utrecht rondom de leerling Frans op dat hij de formule voor de compensatie nergens kon vinden. Die zou, was zijn mening, net als de regeling Omzetting scores in cijfers, gepubliceerd moeten zijn. Met de publicatie ervan, de explicatie van de formule, is het verzoek van Brussee ingewilligd. De formule zelf is al jaren in gebruik.

*RK:* *Mijn eerste reactie: OMG.*

*Ik sluit me volledig aan bij de conclusie van advocaat Wilco Brussee dat er van deze formule helemaal niets klopt, en ik zal voor de leken onder u eerst duidelijk proberen te maken wat er niet deugt aan deze formule, en wat de consequenties hiervan zijn.*

CvTE: De blog mist hier hoor en wederhoor. De gebruikte vergelijking gaat mank en daardoor loopt de blog uit het spoor. Het hieronder gebruikte voorbeeld is een te sterke simplificatie; de werkelijkheid is veel complexer. Technisch klopt het evenmin, dat begint al met de keuze voor een scoreschaal die uitgaat van 100 punten, in tegensteling tot de scoreschaal van 90 punten die bij de examenpraktijk gebruikt wordt. Dan zou de rekenexercitie ook passen bij de gehanteerde formule.

*RK:* *De formule*

*In het geval er een onvolkomenheid in het correctiemodel wordt geconstateerd, maar deze niet tijdig gepubliceerd wordt, kan het CvTE het cijfer besluiten op te hogen via de N-term, om leerlingen zo te compenseren. De daarvoor gehanteerde formule is: 9\*P\*M/L, waarbij M het maximale te halen punten is voor de vraag, L het maximale aantal punten voor het examen en P het gemiddelde percentage punten dat door alle leerlingen is gehaald op die vraag. “Een P-waarde van 0,63 betekent dat de kandidaten gemiddeld 63% van M behaald hebben.”*

*Om duidelijk te maken waarom dit rekenmiddel bizar is, neem ik u mee in een volstrekt theoretische vraag.*

*‘Jan krijgt 50 cent wisselgeld. Welk antwoord geeft waarschijnlijk aan welke munten hij terugkreeg.’*

*A: een munt van 50 cent.*

*B: 5 muntjes van 10 cent.*

*Beide antwoorden zijn goed, maar in het correctiemodel staat dat het antwoord B moet zijn. Navraag bij het CvTE leert dat antwoord A echt fout gerekend moet worden. (Het klinkt bizar, maar zo gaat dat dus echt!)*

*We gaan eens kijken naar een aantal fictieve situaties. De vraag is in dit geval twee punten waard, op een totale examenlengte van 100 punten.*

CvTE: Zoals eerder aangegeven is de praktijk veel complexer. In werkelijkheid gaat een discussie tussen docenten en de Examenlijn juist over de grensgevallen, waarbij een eenduidig antwoord zoals in de casus die door de heer Kneyber gegeven wordt, niet direct en overduidelijk voorhanden is. Dat komt voor, b.v. indien een vraag-antwoord combinatie in de praktijk toch niet volledig beantwoordt aan zijn doel. De ervaring bij de Examenlijn is dat docenten vaak opmerkingen van hoge kwaliteit indienen. Doordat zowel zij, andere docenten met andere opmerkingen, en de leden van de Examenlijn die met deze opmerkingen aan de slag gaan, allemaal vakexperts zijn staan meningen en overtuigingen soms echt lijnrecht tegenover elkaar. Dan komt het voor, ook na langdurig overleg, dat men het niet altijd eens wordt.

De tijd is hierbij een limiterende factor. In een eerste tijdvak worden zo’n 4000 vragen voorgelegd aan 200.000 leerlingen en 26.000 docenten. Bij de Examenlijn kwamen er dit jaar tot 7 juni 1540 opmerking binnen. In de meeste gevallen lossen problemen zich geruisloos op, meestal omdat de eerste en tweede corrector er samen uitkomen. Soms niet, dan is het nemen van een beslissing onvermijdelijk: dat is als een vraag wordt voorgelegd aan de Examenlijn en de normeringsdag nadert.

*RK:* *Scenario 1:*

*Alle leerlingen die het examen maakten hebben het ‘verkeerde’ antwoord gekozen, namelijk A. De P-waarde is nu 0,00. In de rekenmethode krijgen al deze leerlingen daarom (9\*0\*2/100=)0 punten compensatie, oftewel helemaal geen compensatie.*

*Consequentie: leerlingen die het antwoord A hadden, dat inhoudelijk wel goed was, krijgen 0 punten erbij voor de vraag. Leerlingen met antwoord B zouden 0,2 punten krijgen ( maar die waren er niet).*

*Scenario 2:*

*De helft van de leerlingen heeft het goede antwoord gekozen. De P-waarde is nu 0,5. Oftewel, alle leerlingen krijgen (9\*0,5\*2/100=)0,09 punt, dus 0,1 punt compensatie.*

*Consequentie: Degenen die A hadden krijgen 0,1 punt erbij. Degenen die B hadden, en dus van hun docent punten hebben gekregen voor de vraag, omdat dat mocht, krijgen nu dus 0,3 punt voor de vraag.*

*Scenario 3:*

*Alle leerlingen hadden antwoord B. De P-waarde is nu 100% dus 1. Alle leerlingen krijgen nu (9\*1\*2/100)=0,18 dus 0,2 punt compensatie, oftewel het volle pond.*

*Consequentie: Ze hebben al punten gehad voor de vraag, want iedereen had B, maar ze krijgen toch compensatie. Uiteindelijk levert deze vraag ze daarom 0,4 punten op.*

CvTE: Technisch kloppen deze scenario’s niet. Binnen het eigen voorbeeld en de eigen rekenwijze van Kneyber wel, maar hierbij wordt geen rekening gehouden met het feit dan leerlingen al compensatie hebben gekregen hebben via de N-term voor het hele examen.

Zie de presentatie over normering: <https://www.youtube.com/watch?v=SYYFy1QIJso>

Zie voor een uitleg over compensatie voor onvolkomenheden via de N-term: <https://www.examenblad.nl/document/cvte-de-hoogte-van-de-ophoging-bij>

Bij scenario 2 ontbreekt dat de N-term in deze situatie al 0,1 punt hoger is/was dan in de situatie waarbij iedereen de vraag goed had. Bij scenario 3 ontbreekt dat de N-term 0,2 punt lager is dan in de situatie dat iedereen de vraag fout had.

Kneyber focust in zijn redenering alleen op de ophoging en neemt het zogenaamde N-term-effect, het effect dat het antwoordgedrag (de antwoorden) van alle andere kandidaten heeft op de N-term van het examen hierin niet mee. De N-term reageert op het aantal goede antwoorden en voorziet dientengevolge in veel gevallen al in een deel van de compensatie.

*RK: Intermezzo:*

*Wat er hier gebeurt is dus heel vreemd. Leerlingen die benadeeld worden door het onjuiste correctievoorschrift worden minder gecompenseerd dan degenen die per ongeluk wel goed gokten wat er in het correctievoorschrift zou staan. Maar het wordt nog erger!*

CvTE: Wat hier gebeurt is methodologisch niet juist. Kneyber vergelijkt drie verschillende situaties met elkaar, terwijl vergeleken moet worden hoe leerling A behandeld wordt ten opzichte van leerling B. Dit verschil kan met de normering niet meer recht worden gezet. De scores zijn al gegeven en vastgesteld. Wat met compensatie via de normering wel kan is iedereen er iets bij geven zodat alle leerlingen ten minste het cijfer krijgen dat ze hadden gekregen als de vraag niet in het examen had gezeten. Op die manier krijgt leerling A het cijfer dat hij verdient en leerling B krijgt iets teveel, maar is geen enkele leerling benadeeld.

*RK: Realistisch scenario 4 (variant op 2)*

*De helft van de leerlingen heeft het ‘goede’ antwoord gekozen. Echter, leraren nemen geen genoegen met het onzinnige voorschrift van het College, zij gaan toch geen goede antwoorden fout rekenen! Dat is de helft van de docenten, met theoretisch de helft van de leerlingen. Zij besluiten antwoord A toch gewoon goed te rekenen.*

*De P-waarde is nu 0,75. Compensatie is nu (9\*0,75\*2/100)=0,14 punt dus 0,1 punt.*

*Nu wordt het ingewikkelder. Er zijn nu leerlingen die A hadden waarbij de docent het wel had goed gerekend, en leerlingen waarvan A niet goed is gerekend.*

*Score voor de vraag A B*

*Docent rekende A wel goed 0,3 0,3*

*Docent rekende A niet goed 0,1 0,3*

*Consequentie: Leerlingen met hetzelfde antwoord (A), krijgt vanwege een andere opvatting van de corrector 0,2 punten lager!! (Terwijl het antwoord dus gewoon goed is).*

CvTE: Het correctievoorschrift is een algemeen verbindend voorschrift. Juist doordat docenten zich daar niet aan houden, ontstaat ongelijkheid en die kan de normering daarna niet meer, of onvoldoende, rechttrekken. Juist doordat afgesproken is om allemaal het correctievoorschrift te volgen, wordt elke leerling zoveel mogelijk op dezelfde wijze beoordeeld.

RK: *Conclusies:*

*Dat correctiemodellen niet worden aangepast, maar dat er ‘gecompenseerd’ wordt in de normering is kennelijk al een lang gebezigde praktijk. Ongetwijfeld zal het College hameren op de correcte wijze waarop ze de procedures hebben gevolgd. Maar de rest van de wereld zal toch echt van mening zijn dat deze methodiek moreel volstrekt ongeloofwaardig is.*

*En deze verwerpelijke methodiek wordt dus vaak toegepast. Sterker nog: telkens als een fout in het correctievoorschrift niet is aangepast wordt het op deze wijze gecompenseerd.*

CvTE: Het CvTE past het correctievoorschrift aan, als de tijd dat toelaat. Het CvTE blijft zo lang mogelijk in gesprek met vraagstellers. Daarom is, zodra de tijd een aanvulling van het correctievoorschrift niet meer toelaat, het CvTE gehouden leerlingen op de andere vastgestelde manier recht te doen: de compensatie via de n term.

De gehele normeringssystematiek, inclusief de compensatie via de N-term en de pre/posttest- en ankervraagmethodiek, is meermalen door externe deskundigen tegen het licht gehouden en als juist beoordeeld in 2010 (Hoijtink en Sijtsma); 2014 (Glas en Meijer) en onlangs nog in 2017 (RCEC rapport dat op 6 september naar de Tweede Kamer is gestuurd). Het bureau van het CvTE heeft in samenwerking met Cito op 2 februari 2017 en 8 maart 2018 bijeenkomsten georganiseerd waarin diepgaander op de normeringssystematiek is ingegaan. De heer Kneyber is daarvoor uitgenodigd, maar heeft, ondanks rappel, geen gehoor gegeven.

*Een aantal vragen doemen direct in mij op:*

*1) Waarom niet gewoon het correctievoorschrift aanpassen? Wat kan in vredesnaam het probleem zijn?*

*2) Waarom keurt de minister een volstrekt ondeugdelijke compensatieregeling goed?*

*3) Hoe juridisch houdbaar is deze willekeur eigenlijk?*

*4) Als er al dit soort problemen in de compensatieregeling zitten, hoe betrouwbaar is dan de pretest/posttest- of ankervraag-methodiek (die ook niet openbaar is) dan eigenlijk?*

*Ik weet het, het CvTE is niet dol op inhoudelijk reageren op columns en blogartikelen, en ze doen het daarom zelden. Toch stel ik het op prijs, als u dit stuk deelt op sociale media zal het ongetwijfeld helpen..*

CvTE: We denken met bovenstaande beantwoording voldoende te zijn ingegaan op de blog van de heer Kneyber.

Hoogachtend,

Drs. P.J.J. Hendrikse

Voorzitter College voor Toetsen en Examens