



> 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

1. MBVOM
2. DGPV
3. VO

Ministerie van OCW	
Zaaknr.	
Dat. in reg.	24 DEC. 2019
nr.	
Directie	VO

Bureau van het CvTE
Muntstraat 7
3512 ET Utrecht
Postbus 315
3500 AH Utrecht
Nederland
www.cvte.nl

Contactpersoon
J. Kastelein
Sectormanager vmbo
030 28 40 700
j.kastelein@cvte.nl

Onze referentie
CvTE 19.01207

Bijlagen
3

Datum 20 december 2019
Betreft Advies over een eigenstandige normering van het tweede tijdvak

Geachte heer Slob,

Deze brief bevat het advies van het College voor Toetsen en Examen over een eigenstandige normering van het tweede tijdvak. De aanleiding voor dit advies is gelegen in het Onderzoek van de Vrije Universiteit Amsterdam (VU) naar de invloed van de correctie op de slaagkans van leerlingen die het centraal examen herkansen in het tweede tijdvak. Het VU-onderzoek was een bijlage bij uw brief 'Toetsing en examinering in het voortgezet onderwijs 2018' van 14 januari 2019 aan de Tweede Kamer.

In de genoemde kamerbrief benoemt u dat de informatiepositie van de correctoren in het tweede tijdvak groter is dan in het eerste tijdvak en heeft u ons opgeroepen om te komen met een advies over een eigenstandige normering van het tweede tijdvak. Cito is betrokken bij de totstandbrenging van dit advies.

In het examensysteem van het voortgezet onderwijs heeft een leerling het recht om één centraal examen te herkansen¹. Daarvoor is het tweede tijdvak bedoeld² en daarom beslaat dat slechts vier dagen. In deze context is een geheel eigenstandige normering van het tweede tijdvak niet mogelijk vanwege

- de kleine aantallen kandidaten (per centraal examen minder dan 5% van het eerste tijdvak);
- het gegeven dat de vaardigheid van de groep, die deelneemt aan het tweede tijdvak, niet representatief is en bovendien van jaar tot jaar sterk kan wisselen³.

In deze zin is het tweede tijdvak niet volwaardig.

Een geheel eigenstandige, op zichzelf staande, normering van het eerste en tweede tijdvak is alleen mogelijk als beide tijdvakken wel volwaardig in omvang en groepsamenstelling zouden zijn. Dat is in de huidige opzet van het eindexamen niet mogelijk. Signalen uit het scholenveld vormen voor ons een aansporing om tot een nadere afweging te komen.

¹ In vmbo-bb en -kb mag een leerling het centraal examen in één algemeen vak herkansen en daarnaast ook het cspe (centraal schriftelijk en praktisch examen) in het beroepsgerichte profielvak.

² Behalve voor herkansen is het tweede tijdvak ook bedoeld voor inhalen. Leerlingen die vanwege een geldige reden (bijvoorbeeld ziekte) een of meer centrale examens niet in het eerste tijdvak hebben kunnen afleggen, maken ook gebruik van het tweede tijdvak. Vanwege de beperkte duur van het tweede tijdvak (vier dagen) zijn sommigen voor inhalen en/of herkansen ook aangewezen op het derde tijdvak in augustus.

³ Dat komt doordat de leerlingen die een onvoldoende uit het eerste tijdvak herkansen relatief oververtegenwoordigd zijn en gemiddeld genomen minder vaardig zijn.

Hierna beschrijven wij eerst het hoe en waarom van de huidige normering van het tweede tijdvak. Vervolgens geven wij een verklaring van de uitkomsten van het VU-onderzoek. Daarna komen wij tot ons advies.

Bureau CvTE

Datum

20 december 2019

Onze referentie

CVTE 19.01207

Het hoe en waarom van de huidige normering van het tweede tijdvak

Het tweede tijdvak kent geen relatief eigenstandige normering doordat wij bij de publicatie van de N-term van het eerste tijdvak alvast een voorlopige N-term voor het tweede tijdvak bekend maken. 'Voorlopig' houdt daarbij in dat de N-term van het tweede tijdvak niet lager (maar wel hoger) kan zijn, dan de voorlopige N-term. Ervan uitgaande dat de definitieve N-term voor het tweede tijdvak gelijk is aan de voorlopige (wat dus niet hoeft), kan een corrector bepalen hoeveel scorepunten een leerling die herkanst minimaal moet halen om te kunnen slagen. De voorlopige N-term voor het tweede tijdvak is gebaseerd op de N-term die voor het eerste tijdvak is vastgesteld.⁴ De aanname daarbij is dat het eerste en het tweede tijdvak ongeveer even moeilijk zijn, omdat de beide examens gelijktijdig en in samenhang tot stand worden gebracht. Als uit de afnamegegevens blijkt dat het tweede tijdvak moeilijker is dan het eerste tijdvak, dan krijgt het tweede tijdvak een hogere N-term dan de voorlopige N-term. Blijkt uit diezelfde afnamegegevens dat het tweede tijdvak makkelijker is dan het eerste tijdvak, dan krijgt het tweede tijdvak geen lagere N-term, maar is de N-term voor het tweede tijdvak gelijk aan de voorlopige N-term.⁵

Aldus is de N-term voor het tweede tijdvak gerelateerd aan de N-term voor het eerste tijdvak en in die zin is in het tweede tijdvak geen sprake van een eigenstandige normering.

In het eerste tijdvak beschikt de corrector niet over een voorlopige N-term. Daardoor is er verschil tussen het eerste en het tweede tijdvak waar het de informatiepositie van de corrector betreft.⁶

Stoppen met het bekendmaken van de voorlopige N-term voor het tweede tijdvak zou echter betekenen dat leerlingen vaker ten onrechte zakken. Wij vinden dat dit zwaar moet wegen en Cito heeft aangegeven dat er geen methodiek voorhanden is die voldoende nauwkeurig is om dit te voorkomen. Daarom willen wij niet stoppen met de voorlopige N-term voor het tweede tijdvak. Cito steunt deze conclusie. Voor een meer gedetailleerde beschrijving verwijzen wij naar bijlage 1.

Onze verklaring voor de uitkomsten van het VU-onderzoek

Uit het VU-onderzoek komt naar voren dat scores op het centraal examen van het tweede tijdvak die leiden tot onafgeronde eindcijfers als 6,5 (eindcijfer 7) en 5,5 (eindcijfer 6), relatief vaker worden toegekend en scores die leiden tot onafgeronde eindcijfers als 4,45 (eindcijfer 4), 5,45 (eindcijfer 5) en 6,45 (eindcijfer 6) juist relatief minder vaak.

Betekent dit dat sprake is van oneigenlijke cijferbeïnvloeding door correctoren bij het centraal examen van het tweede tijdvak? Wij denken van niet. Het komt doordat de correctoren de beoordelingsruimte die het correctievoorschrift biedt, benutten om een leerling, voor wie de uitkomst van het centraal examen in het tweede tijdvak 'erop of eronder' (geslaagd of gezakt) betekent, niet op één scorepunt laten zakken. Wij vinden dat legitiem. In bijlage 2 gaan wij hier dieper op in.

⁴ De voorlopige N-term voor het tweede tijdvak is in principe gelijk aan de N-term voor het eerste tijdvak. Als in het eerste tijdvak via de N-term gecompenseerd voor een onvolkomenheid, kan de voorlopige N-term voor het tweede tijdvak lager zijn dan de N-term voor het eerste tijdvak.

⁵ Deze alinea beschrijft de hoofdlijn van de normeringssystematiek van het tweede tijdvak. Evenals in het eerste tijdvak kan ook in het tweede tijdvak via de N-term worden gecompenseerd voor een fout.

⁶ In het scholenveld is algemeen bekend dat de N-term niet lager is dan 0,0. Dat doet echter niets af aan het verschil in informatiepositie.

Ons advies

Omdat er goede redenen zijn om in het tweede tijdvak te blijven werken met voorlopige N-termen en vanwege de verklaring die wij zien voor de uitkomsten van het VU-onderzoek, adviseren wij om niet over te gaan tot een (relatief) eigenstandige normering van het tweede tijdvak. Bijlage 3 bevat een nadere toelichting op ons advies.

Bureau CvTE**Datum**

20 december 2019

Onze referentie

CvTE 19.01207

Met vriendelijke groet,



drs. P.J.J. Hendrikse

Voorzitter College voor Toetsen en Examens

Bijlage 1 - Het waarom van de niet-eigenstandige normering van het tweede tijdvak

Bureau CvTE

Datum

20 december 2019

Onze referentie

CVTE 19.01207

Reden voor het bekend maken van de voorlopige N-term voor het tweede tijdvak
Aan het bekend maken van de voorlopige N-term voor het tweede tijdvak – en daarmee aan het verschil in informatiepositie van de correctoren tussen het eerste en het tweede tijdvak – ligt een reden ten grondslag. Die reden ligt in het verschil in meetnauwkeurigheid.

In combinatie met het schoolexamen wordt in het eerste tijdvak gemeten of je gezakt of geslaagd bent. Die meting bestaat uit een groot aantal afzonderlijke metingen: de schoolexamens en de centraal examens voor al je vakken. De meting 'gezakt of geslaagd', die na het eerste tijdvak plaatsvindt, bestaat dus uit een groot aantal deelresultaten. Hoe groter het aantal deelmetingen, des te betrouwbaarder is de uitslag en dat maakt de meting 'gezakt of geslaagd' tot een nauwkeurige.

Dat ligt anders in het tweede tijdvak. Ben je als leerling na het eerste tijdvak gezakt, maar kun je nog slagen door één centraal examen in het tweede tijdvak te herkansen dan doe je dat natuurlijk. Dat ene centraal examen in het tweede tijdvak is een afzonderlijke meting, maar die bepaalt wel of je geslaagd of gezakt bent. Daarmee is de nauwkeurigheid van de meting 'geslaagd of gezakt' in het tweede tijdvak aanmerkelijk kleiner dan in het eerste tijdvak. Leidt je score voor het tweede tijdvak tot een onafgerond eindcijfer van 5,4 of 5,45 of 6,4 of 6,45, dan zou je gezakt kunnen zijn op basis van een meetonnauwkeurigheid die voor jou negatief uitpakt. Dat is wel heel zuur. De wijze waarop correctoren in het tweede tijdvak omgaan met de toekenning van scores, dat wil zeggen met kennis van de voorlopige N-term (zie bijlage 2), draagt bij aan het vermijden van meetonnauwkeurigheden die voor de leerling negatief uitpakken, waardoor deze onterecht zakt.

Ons oordeel over de voorlopige N-term voor het tweede tijdvak

De voorlopige N-term voor het tweede tijdvak helpt voorkomen dat leerlingen onterecht zakken. Wij vinden dat dat zwaar moet wegen. Om die reden willen wij niet stoppen met het bekendmaken van de voorlopige N-term voor het tweede tijdvak. Er zijn derhalve goede redenen om onderscheid te maken tussen het eerste en tweede tijdvak waar het de informatiepositie van de corrector betreft.

Geen geheel en ook geen relatief eigenstandige normering van het tweede tijdvak

Eerder is toegelicht waarom een geheel eigenstandige, op zichzelf staande, normering van het tweede tijdvak in de huidige opzet van het eindexamen niet mogelijk is.⁷

Door te stoppen met het bekendmaken van de voorlopige N-term zou het tweede tijdvak wel een relatief eigenstandige normering krijgen. Hierboven is uiteengezet waarom wij niet willen stoppen met de voorlopige N-term.

Een relatief eigenstandige normering van het tweede tijdvak is dus in de huidige opzet van het eindexamen wel mogelijk, maar wij vinden dat vanuit het oogpunt van een zorgvuldige behandeling van de leerling niet wenselijk.

⁷ Zie de tweede alinea van de brief.

Bijlage 2 – Onze verklaring voor de uitkomsten van het VU-onderzoek

Het VU-onderzoek was gericht op het centraal examen van het tweede tijdvak. Het tweede tijdvak wordt vooral afgelegd door kandidaten die een van hun eindcijfers willen verbeteren. Dat kan van 4 naar 5 zijn, maar ook van 8 naar 9. Er is echter ook een groep leerlingen die wel mee moeten doen, omdat zij op basis van de uitslag na het eerste tijdvak zijn gezakt. Voor deze groep, ongeveer de helft van het totaal aantal deelnemers aan het tweede tijdvak, betekent dat ene cijfer dat zij voor het centraal examen in het tweede tijdvak behalen, dus het verschil tussen geslaagd of gezakt. Voor hen is het erop of eronder. Uit het VU-onderzoek komt naar voren dat scores op het centraal examen van het tweede tijdvak die leiden tot onafgeronde eindcijfers als 6,5 (eindcijfer 7) en 5,5 (eindcijfer 6), relatief vaker worden toegekend en scores die leiden tot onafgeronde eindcijfers als 4,45 (eindcijfer 4), 5,45 (eindcijfer 5) en 6,45 (eindcijfer 6) juist relatief minder vaak. Betekent dit dat sprake is van oneigenlijke cijferbeïnvloeding door correctoren bij het centraal examen van het tweede tijdvak? Wij denken van niet.

Het correctievoorschrift (cv) bevat voor iedere vraag een antwoordmodel. Met inachtneming van de algemene regels van het cv, kennen de correctoren bij iedere vraag conform het antwoordmodel scorepunten toe aan de antwoorden van de leerlingen. Het cv is bindend, maar dat wil niet zeggen dat er geen beoordelaarsruimte is. Het kan zijn dat er goede argumenten zijn om een leerlingantwoord op vraag 17 met twee scorepunten te honoreren, maar evenzeer met drie scorepunten.

Als dat voor vraag 22 evenzeer het geval is, zou een corrector ervoor kunnen kiezen om bij de ene vraag twee en bij de andere vraag drie scorepunten toe te kennen. Dit met als argumentatie dat daardoor de eindscore (de optelling van alle scorepunten die per vraag zijn toegekend) het best past bij de prestatie die de leerling tijdens de afname heeft geleverd. Het cv, hoe gedetailleerd ook, garandeert geen eenduidige uitkomst. De beoordeling van een examen is een proces waarin de corrector vanuit zijn professionaliteit, met inachtneming van het cv, de eindscore bepaalt.

In de huidige situatie kan de corrector, met de kennis van de voorlopige N-term voor het tweede tijdvak, zelf bepalen welk voorlopig CE-cijfer bij welke eindscore hoort. Bij een eindscore van 47 zou het voorlopige CE-cijfer van 5,7 kunnen horen en bij een eindscore van 48 scorepunten het voorlopige CE-cijfer van 5,8. Voor een leerling kan het CE-cijfer van 5,7 dan wel 5,8 (in combinatie met diens SE-cijfer het verschil betekenen tussen een 5 en 6 als eindcijfer (of tussen een 6 en een 7)). Ook de in 2012 ingevoerde CE-eis⁸ kan maken dat een CE-cijfer van 5,7 of 5,8 het verschil betekent tussen zakken of slagen. Als het dan zo is, dat een eindscore van zowel 47 als 48 scorepunten een verdedigbare eindscore is en je realiseert je daarbij dat het verschil tussen 47 en 48 scorepunten ook het verschil tussen zakken of slagen kan betekenen, wat doe je dan als leraar van een leerling die je de afgelopen twee jaar in je klas hebt gehad en van wiens vaardigheden je een goed beeld hebt doordat je hem vele proefwerken en schoolexamens hebt afgenomen? De leraar/corrector zal kiezen voor 48 scorepunten. Je laat je leerling niet op één scorepunt zakken. Over die eindscore van 48 moet hij het dan wel eens worden met zijn tweede corrector, want de beide correctoren stellen in overleg de eindscore vast. Het professionele oordeel van de eerste corrector vergt dus ook nog eens een bevestiging door een tweede corrector, die zelf ook leraar is.

Dit verklaart waarom eindscores die leiden tot onafgeronde eindcijfers van 5,5 vaker voorkomen dan eindscores die leiden tot onafgeronde eindcijfers van 5,45.

⁸ Om te kunnen slagen moet het gemiddelde van de centrale examens van de leerling tenminste 5,5 zijn.

Bureau CvTE

Datum

20 december 2019

Onze referentie

CvTE 19.01207

Als 48 scorepunten echt niet verdedigbaar is, dan kijk je bij het tweede tijdvak als corrector of behalve 47 scorepunten ook 46 of 45 een verdedigbare eindscore kan zijn. Als de tweede corrector hiermee instemt, ziet de leerling aan zijn resultaat dat hij niet één, maar twee of drie scorepunten tekort kwam. Nog steeds zuur, maar gemakkelijker te accepteren dan op één scorepunt zakken. Dit verklaart waarom eindscores die leiden tot onafgeronde eindcijfers van 5,45 of 6,45 relatief minder vaak voorkomen.

Bureau CvTE

Datum

20 december 2019

Onze referentie

CvTE 19.01207

Dit is hoe het in de praktijk van het centraal examen in het tweede tijdvak werkt. Volgens het VU-onderzoek slagen jaarlijks zo'n 1400 leerlingen doordat de voorlopige N-term bekend is. Dat is ongeveer 0,7% van het totaal. Er zijn ongeveer 1200 VO-scholen in Nederland. Dat betekent dat er gemiddeld genomen iets meer dan één leerling – verspreid over de schooltypen vwo, havo en vmbo en over alle vakken – slaagt door benutting van de beoordelaarsruimte.

Dat correctoren de beoordelingsruimte die het correctievoorschrift biedt, benutten om een leerling niet op één scorepunt laten zakken, vinden wij legitiem.

De VU-onderzoekers hebben ook geconstateerd dat er verschillen zijn tussen vakken en tussen scholen. De verschillen tussen vakken zijn ontstaan door verschillen in proporties open vragen. De centrale examens wiskunde bestaan voor 100% uit open vragen. Bij de moderne vreemde talen ligt de proportie open vragen onder de 25%, omdat daar de meest gebruikte vraagvorm die van de meerkeuzevraag is. De andere vakken zitten daar qua proportie open vragen tussenin. Hoe groter de proportie open vragen hoe groter de beoordelingsruimte voor de corrector.

De verschillen tussen scholen ontstaan (mede) doordat de eerste corrector tevens de docent van de leerling is. Dat is zowel bij het eerste als het tweede tijdvak het geval. Dat het gemaakte werk ook door een tweede corrector, in casu een docent van een andere school, wordt nagekeken, neemt verschillen tussen scholen niet weg. Tweede correctie is een goede zaak, want het draagt bij aan de zorgvuldigheid waarmee het gemaakte werk wordt beoordeeld, maar het is geen garantie voor uniformiteit.

Bijlage 3 - Nadere toelichting op ons advies

Bureau CvTE

Voor leerlingen van wie de uitslag na het eerste tijdvak 'gezakt' luidt, biedt het centraal examen van het tweede tijdvak een laatste kans. Zij zijn gebaat bij het houvast dat de voorlopige N-term biedt. In de context van het tweede tijdvak vinden wij dat naar leerlingen toe redelijk en billijk. Het bepalen van de eindscore is een professioneel oordeel over het kennis- en vaardigheidsniveau zoals de leerling dat tijdens het afleggen van zijn centraal examen heeft getoond. Naar de docenten toe zien wij dat als een erkenning van hun professionaliteit. Ons inziens is daarmee de zorgvuldigheid van de zak-slaag-beslissing in het voortgezet onderwijs goed geborgd. Zowel na het eerste tijdvak, alsook na het tweede tijdvak voor hen die daarop aangewezen zijn.

Datum

20 december 2019

Onze referentie

CvTE 19.01207

Dat er verschillen tussen vakken zijn doet daar niets aan af. De gebruikte vraagvormen (open of meerkeuze) sluiten aan bij de traditie van het vak en het draagvlak daarvoor onder de vakdocenten, al dan niet verenigd in een vakvereniging. Ook de leerling weet globaal hoe het centraal examen voor zijn vakken eruit ziet: bij wiskunde geen, bij de talen veel meerkeuzevragen. Deze kennis laat hij – meestal impliciet – meewegen als hij bepaalt op welk centraal examen hij zijn recht op herkansing wil uitoefenen. Dankzij de voorlopige N-term weet hij op welke definitieve N-term hij minimaal mag rekenen. Dat de corrector ook bekend is met de voorlopige N-term minimaliseert de kans dat de ene meting in het tweede tijdvak false negative uitpakt, waardoor de leerling ten onrechte zou zakken.

De verschillen tussen scholen neem je niet zo maar weg. Die zijn er niet alleen bij het tweede tijdvak, maar ook bij het eerste tijdvak en bij de schoolexamens. Ze zijn er sinds de invoering van de huidige centrale examens in 1974. Wij wijzen er daarbij op dat de uniformiteit van de beoordeling in het VO dankzij de centrale examens aanmerkelijk hoger is dan in het MBO, HBO en WO. Deze sectoren kennen geen centrale examens⁹, geen bijbehorende uniforme en bindende correctievoorschriften, geen verplichte tweede correctie en evenmin de regel dat het schoolexamen moet zijn afgesloten voordat het centraal examen mag worden afgelegd. In het examensysteem van het voortgezet onderwijs zijn voor de kwaliteitsborging van de diploma's meer checks ingebouwd dan in de andere sectoren.

Er zijn goede redenen zijn om in het tweede tijdvak te blijven werken met voorlopige N-termen. In de uitkomsten van VU-onderzoek zien wij geen reden om daar anders over te denken.

Daarom adviseren wij om niet over te gaan tot een relatief eigenstandige normering van het tweede tijdvak.

⁹ De centrale examens Nederlands en Engels in het MBO daargelaten.