

Verslag ijkingsonderzoek centrale examens Nederlands VO 2014 aan de referentiesets en referentieniveaus taal

Marieke van Onna, Hendrik Straat, René Alberts

Cito | Arnhem

Vooraf

In deze notitie wordt gerapporteerd over de resultaten uit door Cito uitgevoerd equivaleringsonderzoek. Het gaat om equivalering van de centrale examens Nederlands aan de referentieniveaus Nederlands 2F t/m 4F. In 2013 heeft een voorbereidend onderzoek plaats gevonden en in 2014 de hoofdfase. De bevindingen hebben geleid tot een gedeeltelijke herziening van de te hanteren normen voor de examens vanaf 2015. De gerapporteerde resultaten liggen direct ten grondslag aan het door het CvTE genomen besluit dat in paragraaf 4 als onderdeel van de conclusies vermeld wordt.

1. Inleiding

In opdracht van het College voor Toetsen en Examens heeft Cito de gevolgen van de implementatie van de referentieniveaus Nederlandse taal voor de moeilijkheidsgraad van de centrale examens Nederlands voortgezet onderwijs (VO) en de eventuele gevolgen voor de normering onderzocht.

Voor de centrale examens Nederlands in het VO geldt dat deze in 2015 afgestemd moeten zijn op de referentieniveaus zoals vastgelegd in het 'Besluit referentieniveaus Nederlandse taal en rekenen'. Om te beoordelen in welke mate de tot 2015 gebruikte cesuren afwijken van de beoogde referentieniveaus, is onderzoek uitgevoerd. Hierin zijn de centrale examens 2014 vergeleken met de referentiesets taal 2F en 3F ¹.

In de zomer van 2013 zijn de zogenoemde openbare en geheime referentiesets voor taal en rekenen vastgesteld. Dit zijn sets van opgaven die representatief zijn voor de referentieniveaus. Per set is bepaald wat de cesuur is om te kunnen spreken van beheersing van het referentieniveau. Dat is gebeurd door een standaardbepalingsprocedure, waarin de opgaven uit de referentiesets op moeilijkheidsgraad beoordeeld zijn door panels van experts.

De centrale examens vmbo BB, KB en GL/TL zijn vergeleken met de referentieset 2F. Daarbij is gebruik gemaakt van opname van geheimgehouden referentiesetopgaven in het digitale centrale examen vmbo BB en posttests met zowel referentiesetopgaven als opgaven uit het digitale centrale examen vmbo KB respectievelijk het (papieren) centrale examen vmbo GL/TL.

Het centraal examen Nederlands havo-2014 is via een posttest vergeleken met de referentieset 3F. Voor de vergelijking van het vwo-examen met het 4F-niveau was er geen referentieset beschikbaar. Bij dit examen heeft daarom een 4F-standaardbepaling met experts plaatsgevonden. Daarnaast heeft een posttest plaatsgevonden, waarin opgaven uit de referentieset 3F en examenopgaven van 2014 met elkaar zijn vergeleken.

Deze onderzoeksmethodes worden in paragraaf 2 voor de verschillende niveaus nader toegelicht.

¹ Zie www.referentiesets.nl

2. Methoden

2.1 Digitaal centraal examen BB: anker-in-package

Het digitale centraal examen Nederlands BB bestond in 2014 uit twaalf varianten: negen varianten voor de eerste tijdvakafname en drie herkansingsvarianten. Deze twaalf varianten vormen samen 'de package': het pakket van de digitale centrale examens Nederlands BB 2014.

In drie varianten zijn referentiesetopgaven 2F opgenomen (anker in package). Per variant waren dit respectievelijk vier, negen, en negen items. In totaal betrof dit dertien verschillende opgaven, met totale maximale score dertien.

Iedere variant van het digitale centraal examen werd door tenminste 1300 kandidaten gemaakt.

Bij de analyse zijn de varianten met referentiesetopgaven geschaald met de referentieset 2F volgens het OPLM². Daarna is de 2F-cesuur van de referentieset 2F overgebracht naar de twaalf varianten van het examen. Zo is de vaardigheid die nodig is om een voldoende score te behalen op de referentieset 2F vertaald naar een score voor iedere examenvariant waarbij de kandidaat dezelfde vaardigheid toont. Deze score per variant is vervolgens vertaald in een N-term per variant, passend bij het referentieniveau 2F.

2.2 Digitaal centraal examen KB: digitale posttest

Het digitale centraal examen Nederlands KB bestond in 2014 uit negen varianten. De digitale posttest bestond uit zes varianten van ieder vijftien of zestien opgaven. Hiervan kwamen er negen of tien uit de referentieset 2F, de resterende opgaven kwamen uit het digitale centraal examen KB 2014. In totaal zijn er negentien opgaven uit de referentieset 2F opgenomen in de posttest en tweeëndertig uit het digitale centraal examen KB 2014. Het aantal leerlingen per variant in de digitale posttest varieerde van 154 tot 261. De digitale posttest KB is afgenomen bij leerlingen in leerjaar 3 vmbo GL/TL.

Bij de analyse zijn de opgaven uit het digitale centraal examen van de posttest-varianten geschaald met de referentieset 2F volgens het OPLM. Daarna is de 2F-cesuur overgebracht naar de negen examenvarianten van het digitale centraal examen. Analoog aan BB (zie 2.1) kon zo voor iedere variant van het digitale centraal examen KB een N-term bepaald worden, die past bij het referentieniveau 2F.

2.3 Centrale examens vmbo GL/TL en havo: posttest

Bij een posttest worden opgaven uit het centraal examen van het actuele examenjaar gecombineerd met opgaven uit een eerder examenjaar (ankeropgaven) en/of met opgaven uit een referentieset. Daardoor kan de moeilijkheidsgraad van het actuele centraal examen vergeleken worden met die van het centraal examen uit het eerdere jaar en/of met de referentieset.

In de posttest voor Nederlands vmbo GL/TL is het centraal examen van 2014 vergeleken met zowel een eerder centraal examen als met de referentieset 2F. Voor havo ging het bij de posttest om een combinatie van opgaven uit het centraal examen van 2014 en de 3F-referentieset.

² OPLM staat voor het One Parameter Logistic Model:

<http://www.cito.nl/onderzoek%20en%20wetenschap/psychometrie/psychometrie/software/oplm>

De posttests worden meestal niet afgenomen in examenklassen, omdat die groepen na de afname van het examen niet meer bereikbaar zijn. De opgaven worden daarom gemaakt door leerlingen met een zo vergelijkbaar mogelijk vaardigheidsniveau. Meestal wordt gekozen voor leerlingen van een hoger leerniveau en een lager leerjaar. De posttests van vmbo GL/TL en havo zijn respectievelijk afgenomen in havo-3 en vwo-5.

In de regel worden de opgaven uit de verschillende opgavensets gecombineerd in een aantal verschillende boekjes. Deze worden gecombineerd uitgezet in de deelnemende klassen. Elke klas krijgt alle versies aangeboden, waarbij elke leerling uiteraard maar één versie krijgt voorgelegd.

Omdat het aantal boekjes voor Nederlands vmbo GL/TL groot was (8) is er voor gekozen om niet acht maar vier boekjes per klas uit te zetten, waarmee er dus een dubbele posttest georganiseerd is. De steekproef van scholen is daarbij opgesplitst. Tabel 1 geeft een overzicht van de aantallen scholen, klassen en leerlingen van de posttests vmbo GL/TL en havo.

Tabel 1. Aantallen leerlingen posttests Nederlands vmbo GL/TL en havo

posttest	niveau	Inhoud			Aantallen items uit		Overige aantallen			
		new exam	ref exam	ref.set	Examen '14	Ref.set	versies	scholen	klassen	observ /boekje
1	GT	GT 2014-1	2007	2F	21	14	4	12	31	154
2	GT	GT 2014-1	2007	2F	21	14	4	13	33	154
3	HAVO	H 2014-1	nvt	3F	15	20	6	20	47	177

In Tabel 1 staan tevens de aantallen opgaven in de posttest die overlaptten met de centrale examens 2014 en de referentiesets.

Bij de analyse zijn de opgaven uit het centraal examen 2014 die in de posttest-varianten voorkwamen geschaald met de referentieset volgens het OPLM. Bij vmbo GL/TL zijn de opgaven uit de referentieset 2F meegeschaald, bij havo betrof het de opgaven uit de referentieset 3F.

Vervolgens is voor zowel vmbo GL/TL als havo de cesuurscore op de respectievelijke referentieset overgebracht naar het centraal examen 2014. Analooq aan BB (zie 2.1) en KB (zie 2.2) kon zo voor de centrale examens GL/TL en havo een N-term bepaald worden, die past bij het referentieniveau 2F respectievelijk 3F.

2.4 Centraal examen vwo

2.4.1 combi pre/posttest

Voor de vergelijking van het vwo-examen met het 4F-niveau was er geen referentieset beschikbaar. Voor vwo heeft zowel een pretest als een posttest plaatsgevonden, waarvan de resultaten gecombineerd zijn. De afname van de pre/posttest bestond uit opgaven uit het centraal examen van 2014, 3F-referentiesetopgaven en enkele 'nieuwe stijl' opgaven. De opgaven uit het centraal examen vwo 2014 waren vergelijkbaar met opgaven uit eerdere vwo-examens. De referentiesetopgaven representeerden het 3F-referentieniveau (dat voor havo is aangewezen). Bij de opgaven 'nieuwe stijl' ging het om opgaven die door de CvTE-vaststellingscommissie op theoretische gronden als moeilijk waren beoordeeld en die mogelijk als ankeropgaven kunnen worden ingezet bij toekomstige pretests.

De gecombineerde afname voor vwo is gehouden onder leerlingen van vwo-5. Bij de vwo-afname is per klas slechts één boekje uitgezet. De boekjes waren (op puntenaantal) in gelijke mate samengesteld uit de drie bronnen. De aantallen scholen en klassen staan in Tabel 2.

Tabel 2. Aantal leerlingen bij de pre/posttest Nederlands vwo

posttest	niveau	Inhoud			aantallen			
		new exam	ref.set	anker/pretest	versies	scholen	klassen	observ /boekje
1	VWO	V 2014-1	3F	anker	3	14	27	230

Uit analyses van deze gecombineerde afnames vwo kwam naar voren dat het invoeren van de 3F-norm een N-term van 2,6 voor het examen Nederlands vwo 2014 zou betekenen. Dit ligt in lijn met het verschil in moeilijkheidsgraad tussen de centrale examens havo en vwo. Ook blijkt hieruit dat de gehanteerde vwo-norm een stuk strenger is dan 3F-referentieniveau.

2.4.2. Standaardbepaling

Voor de vergelijking van het vwo-examen met het 4F-niveau was er geen referentieset beschikbaar. Daarom is naast de in paragraaf 2.4.1 beschreven test een panel van beoordelaars gevraagd voor een standaardbepaling op het centraal examen 2014. Dit panel was divers van samenstelling en bestond uit: vwo-docenten die ook aan havo-klassen onderwijs gaven, mbo-docenten, en een representant uit het basisonderwijs. De beoordelaars kwamen op zaterdag 10 mei 2014 bijeen voor instructie en oefening met het centraal examen vwo 2013. Nadat het centraal examen 2014 was afgenomen, hebben zij hun oordelen over het centraal examen 2014 schriftelijk toegestuurd aan Cito. Daarbij kregen zij de volgende instructie:

Goat u bij de beoordeling uit van de niveaubeschrijvingen uit van de referentieniveaus.

Geeft u op het scoreformulier per set van opgaven duidelijk aan welk (heel) aantal punten een grenskandidaat op 4F-, respectievelijk 3F-niveau, volgens u behaalt. Met dit aantal punten toont de grenskandidaat dus volgens u aan dat hij/zij het 4F-, respectievelijk 3F-niveau, beheerst.

Tijdens de instructiebijeenkomst hebben de beoordelaars uitleg gekregen over 'de grenskandidaat' (een kandidaat met een vaardigheidsniveau dat net wel / niet genoeg is voor niveau 4F) en over de werkwijze. Deelnemers hadden vooraf grondig kennis genomen van het centraal examen vwo 2013 en de referentieniveaus taal 3F en 4F zoals vastgelegd in het 'Besluit referentieniveaus Nederlandse taal en rekenen'. De examenopgaven uit dit examen waren geclusterd in vier relatief homogene setjes van opgaven: kort-antwoord-vragen, lange antwoordvragen, tekstfragmenten en samenvatting. Elke deelnemer gaf bij elke vraag een score voor de grenskandidaat. De somscore per setje werd vervolgens geïnventariseerd en deelnemers met lage c.q. een hoge somscore werd gevraagd een toelichting te geven. Na de toelichting/discussie kon elke deelnemer zijn/haar oordeel nog bijstellen. Tot slot van de bijeenkomst is de deelnemers gevraagd aan te geven hoeveel scorepunten een 3F-kandidaat zou behalen op deze setjes van examenopgaven.

De beoordelaars moesten op vijf setjes van opgaven uit het centraal examen 2014 aangeven hoeveel punten een grenskandidaat 4F zou scoren. De scores op de setjes zijn opgeteld tot de 4F-cesuur in de verwerking van de gegevens, het mediane oordeel op deze totaalscore is gebruikt voor de bepaling van de 4F-cesuur. Ook hebben de beoordelaars een inschatting gemaakt van de scores van een 3F-grenskandidaat, waarmee de 3F-cesuur op het examen vwo is geschat. Deze cesuren 4F en 3F zijn vervolgens vertaald in N-termen voor het centraal examen Nederlands vwo 2014

3. Resultaten

Alle puntschattingen voor de N-term behorend bij het aangewezen referentieniveau staan in Tabel 3. In een afzonderlijke kolom zijn de N-termen van de examens zoals vastgesteld bij de normering van 2014 erbij gezet.

Bij vmbo-BB zou een N-term uitgaande van de referentiecesuur 2F 1,1 punt strenger zijn. Gewogen naar het aantal afnames per variant zou 42,5% van de BB-examenkandidaten van 2014 het 2F-niveau behalen.

Bij vmbo KB zou een N-term uitgaande van de referentiecesuur 2F gemiddeld 0,4 punt lager/soepeler zijn, met 91,1% van de KB-examenkandidaten van 2014 die het 2F-niveau behalen.

De cesuur van het centraal examen vmbo-GL/TL ligt met zijn huidige N-term 1,5 punt boven de referentiecesuur 2F. Als de 2F-norm aangehouden zou worden, zou er maar liefst 1,5 punt soepeler genormeerd worden en zou 99% van de GL/TL-examenkandidaten van 2014 het 2F-niveau behalen.

De cesuur van het centraal examen havo ligt met zijn huidige N-term 0,7 cijferpunt boven de 3F-norm. Als de 3F-norm aangehouden zou worden, zou er soepeler genormeerd worden en zou 93,1% van de havo-examenkandidaten van 2014 het 3F-niveau behalen.

De cesuur van het centraal examen vwo ligt volgens de standaardbepaling met zijn huidige N-term 0,4 cijferpunt onder de 4F-norm. Als de 4F-norm aangehouden zou worden, zou er strenger genormeerd worden en zou 76,5% van de vwo-examenkandidaten van 2014 het 4F-niveau behalen. Verder kwam uit de standaardbepaling dat de afstand tussen de cesuur bij de N-term vwo en de referentiecesuur 3F 1,6 cijferpunten bedraagt en de afstand tussen 3F en 4F 2,0 cijferpunten.

Tabel 3. Overzicht van de ijkingsresultaten bij de examens Nederlands 2014

Schooltype	Variant	N-term Examen	N-term 2F	N-term 3F	N-term 4F	Percentage IIn dat referentieniveau haalt	Vershil N-term
BB	1a	1,1	-0,2			44,8	-1,3
	1b	1,0	-0,2			50,1	-1,2
	1c	1,0	-0,1			41,3	-1,1
	2a	0,9	-0,2			42,7	-1,1
	2b	0,7	-0,4			35,9	-1,1
	2c	1,1	0,0			47,3	-1,1
	3a	1,1	0,0			43,4	-1,1
	3b	0,9	-0,2			29,8	-1,1
	3c	0,8	-0,4			33,5	-1,2
KB	1a	0,5	0,9			92,0	0,4
	1b	0,4	0,9			92,7	0,5
	2a	0,5	0,8			89,4	0,3
	2b	0,9	1,4			89,0	0,5
	3a	0,7	1,1			91,3	0,4
	3b	0,4	0,7			92,3	0,3
GL/TL	CE	0,4	1,9			99,0	1,5
havo	CE	1,2		1,9		93,1	0,7
vwo	CE	1,3		2,9	0,9	76,5	-0,4

4. Conclusie

In Tabel 3 worden de puntschattingen van de referentie-N-termen gegeven. Bij vmbo-BB en vwo zou op basis van dit onderzoek de N-term naar beneden bijgesteld moeten worden om op het beoogde referentieniveau (resp. 2F en 4F) cijfers te geven. Bij vmbo KB, vmbo GL/TL en havo zou met een hogere/soepelere N-term gewerkt kunnen om nog steeds volgens het aangewezen referentieniveau (2F dan wel 3F) cijfers te geven.

Eerdere bevindingen

In 2013 is er gerapporteerd over een onderzoek naar ijkings van centraal examens Nederlands 2011 en 2012 aan de referentieniveaus. Deze ijkings verliep grotendeels via twee methoden van standaardbepaling. De onzekerheid bij dat onderzoek was vrij groot. Vergelijken we de resultaten van 2014 met dit eerdere onderzoek, dan zijn de resultaten van BB en vwo in grote lijnen in overeenkomst met de resultaten van 2013: de norm zou verzwakt moeten worden. Bij havo en vmbo GL/TL zijn de beoordelaars in 2013 strenger geweest dan nu uit de empirische ijkings aan de respectievelijke referentiesets komt: in 2013 leek de verschuiving in N-term negatief of 0 te zijn. Nu blijkt dat de prestatie-eis voor een voldoende bij deze examens boven de beoogde referentieniveaus liggen.

Gevolgen voor de normering van de centraal examens Nederlands

De uitkomsten van dit onderzoek hebben het CvTE tot het volgende besluit gebracht:

BB: normering wordt 0,1 cijferpunt strenger

Het CvTE zal voor de centrale examens Nederlands vmbo BB van 2015 N-termen vaststellen die 0,1 cijferpunt strenger zijn dan de N-termen tot en met 2014. Gevolg: De lat bij de CE's Nederlands vmbo BB wordt 0,1 cijferpunt hoger gelegd.

vwo: normering wordt 0,4 cijferpunt strenger

Het CvTE zal voor de centrale examens Nederlands vwo van 2015 N-termen vaststellen die 0,4 cijferpunt strenger zijn dan de N-termen tot en met 2014. Gevolg: De lat bij de CE's Nederlands vwo wordt 0,4 cijferpunt hoger gelegd.

KB, GL/TL en havo: geen verandering in prestatie-eisen

De prestatie-eisen voor de centrale examens Nederlands vmbo KB, vmbo GL/TL en havo hoeven niet te worden aangepast ten opzichte van de examenjaren tot en met 2014. Bij deze schooltypen verandert de hoogte van de lat dus niet.

Dit besluit en de toelichting daarop zijn aan het scholenveld meegedeeld met een brief 'Gevolgen van de referentieniveaus taal voor de normering van de centrale examens Nederlands 2015' (kenmerk CvE-14.02279; datum: 24 september 2014). De brief is via Examenblad.nl verspreid. Docenten Nederlands zijn tevens op de brief geattendeerd via de vakspecifieke informatie Nederlands in zowel de September- als de Maartmededeling van het CvTE over het eindexamen van 2015.