



College voor Toetsen en Examens

TERUGBLIK

CENTRAAL EXAMEN ENGELS GT
EERSTE TIJDVAK 2018

Inleiding

Quickscan

Via WOLF (Windows Optisch Leesbaar Formulier) geven examinatoren per vraag de scores van hun kandidaten voor het centraal examen aan Cito door. Ten behoeve van de normering stelt Cito op basis van de WOLF-gegevens een toets- en itemanalyse van het centrale examen op. Via WOLF wordt de examinatoren gevraagd naar hun mening over de moeilijkheidsgraad, de lengte en de aansluiting op het gegeven onderwijs. Ook wordt de examinatoren gevraagd welk rapportcijfer zij aan dit centraal examen zouden geven. De resultaten van deze bevraging via WOLF worden weergegeven via de quickscan. De quickscan geeft een beeld van de mening van de docenten over het centraal examen nadat zij de eerste correctie hebben verricht.

De vragenlijst

Bij enkele vakken zijn de examinatoren ook ná de normering middels een uitgebreidere vragenlijst over het centrale examen eerste tijdvak bevestigd. Met de gegevens die door Cito zijn verzameld d.m.v. de quickscan en vragenlijst, beoogt CvTE om een representatief¹ beeld te verkrijgen van de meningen van de docenten over het centrale examen. CvTE kan dan bezien of in toekomstige centrale examens, met behoud van handhaving van het niveau, ingespeeld kan worden op de wensen en opvattingen uit het veld. Met uitzondering van het rapportcijfer maken de vragen uit de quickscan eveneens deel uit van de vragenlijst. Het is mogelijk dat dezelfde vraag verschillend beantwoord is doordat de quickscan vóór en de vragenlijst ná de normering is afgenomen.

Afnamegegevens

Cito stelt jaarlijks een digitaal examenverslag op de centrale examens: www.cito.nl. Per centraal examen worden daarin afnamegegevens weergegeven die zijn ontleend aan de Toets en Itemanalyse.

Terugblik

De resultaten van de quickscan en de vragenlijst vormen samen met een overzicht van de afnamegegevens de terugblik op het centraal examen. Hierna treft u de terugblik aan op het centraal examen Engels GT eerste tijdvak 2018 en de conclusies van CvTE daaruit.

¹ Ten aanzien van de representativiteit van de steekproef moet worden opgemerkt dat er sprake is van een selectie uit de populatie aangezien het uitsluitend vakdocenten betreft die een groepsrapportage hebben aangevraagd.

1 Resultaten van de quickscan

Bron: www.cito.nl

Kies: voortgezet onderwijs → centrale examens → Examenverslagen → Waardering per examen 2018

Wat is uw oordeel over de moeilijkheidsgraad van het examen?				
			Abs.	Perc.
Ik vind het examen				
te moeilijk			213	13%
moeilijk			929	55%
niet te moeilijk/niet te makkelijk			537	32%
makkelijk			19	1%
te makkelijk			0	0%
		N=	1698	100%

gemiddelde 2,21
standaardafw. 0,66

Wat is uw oordeel over de lengte van het examen in verhouding tot de tijd die de kandidaat ervoor beschikbaar heeft?				
			Abs.	Perc.
Het examen is				
te lang			356	21%
precies goed			1302	77%
te kort			40	2%
		N=	1698	100%

gemiddelde 1,81
standaardafw 0,45

Wat is uw oordeel over de inhoudelijke aansluiting van het examen bij het gegeven onderwijs?				
			Abs.	Perc.
zeer goed			9	1%
goed			411	24%
voldoende			947	56%
onvoldoende			295	17%
slecht			36	2%
		N=	1698	100%

gemiddelde 3,04
standaardafw 0,72

Welk cijfer zou u dit examen geven?				
Ik geef dit examen een:			Abs.	Perc.
	1		3	0%
	2		6	0%
	3		17	1%
	4		100	6%
	5		241	14%
	6		550	32%
	7		615	36%
	8		152	9%
	9		14	1%
	10		00	0%
		N=	1698	100%

gemiddelde 6,25
standaardafw 1,14

2 Resultaten van de vragenlijst

Vorbereiding op het examen en aansluiting bij de syllabus

Welke documenten gebruikt u bij de voorbereiding van uw leerlingen op het examen?

	Altijd	Meestal	Zelden	Nooit
Een (combinatie van) methode(n)	377(71,0%)	110(20,7%)	27(5,1%)	17(3,2%)
September- en/of Maartmededelingen van CvTE	183(34,5%)	121(22,8%)	122(23,0%)	105(19,8%)
Oude examens	485(91,3%)	45(8,5%)	0(0,0%)	1(0,2%)
Syllabus (gepubliceerd door CvTE)	112(21,1%)	96(18,1%)	141(26,6%)	182(34,3%)
Regeling toegestane hulpmiddelen voor de centrale examens VO	288(54,2%)	113(21,3%)	74(13,9%)	56(10,5%)

Beantwoordt het examen aan de eisen die in de door CvTE gepubliceerde syllabus worden gesteld?
Het examen beantwoordt

volledig aan de eisen	70	13,2%
in voldoende mate aan de eisen	315	59,3%
in onvoldoende mate aan de eisen	38	7,2%
in het geheel niet aan de eisen	2	0,4%
Dit kan ik niet beoordelen	106	20,0%
Totaal	531	100,0%

Wat is uw oordeel over de inhoudelijke aansluiting van het examen bij het gegeven onderwijs?

zeer goed	16	3,0%
goed	224	42,2%
voldoende	203	38,2%
onvoldoende	78	14,7%
slecht	10	1,9%
Totaal	531	100,0%

Uitgangsmateriaal, bronnen en teksten

Hoe beoordeelt u het aantal teksten dat gebruikt wordt als uitgangsmateriaal voor de vragen?

Het aantal teksten is

te groot	136	25,6%
precies genoeg	387	72,9%
te klein	8	1,5%
Totaal	531	100,0%

Hoe beoordeelt u de totale hoeveelheid tekst die gebruikt wordt als uitgangsmateriaal voor de vragen?

De totale hoeveelheid tekst is

te groot	187	35,2%
precies genoeg	337	63,5%
te klein	7	1,3%
Totaal	531	100,0%

Hoe beoordeelt u de geschiktheid van de onderwerpen van de teksten voor de kandidaten?

Ik vind de onderwerpen

allemaal geschikt	32	6,0%
grotendeels geschikt	435	81,9%
grotendeels ongeschikt	64	12,1%
allemaal ongeschikt	0	0,0%
Totaal	531	100,0%

Het is ongewenst dat teksten naar algemeen gevoelens aanstootgevend zijn voor bepaalde bevolkingsgroepen. Wat is uw mening over de aanvaardbaarheid van de teksten?

Ik vind de teksten

allemaal aanvaardbaar	384	72,3%
grotendeels aanvaardbaar	143	26,9%
grotendeels onaanvaardbaar	3	0,6%
allemaal onaanvaardbaar	1	0,2%
Totaal	531	100,0%

Hoe beoordeelt u het examen wat betreft de variatie in onderwerpen in de gebruikte teksten?
Ik vind de variatie in onderwerpen

goed	102	19,2%
ruim voldoende	242	45,6%
voldoende	171	32,2%
onvoldoende	16	3,0%
slecht	0	0,0%
Totaal	531	100,0%

Hoe beoordeelt u de moeilijkheidsgraad van de teksten in het examen?
Ik vind de teksten

zeer moeilijk	43	8,1%
moeilijk	293	55,2%
precies goed	193	36,3%
te gemakkelijk	2	0,4%
Totaal	531	100,0%

Moeilijkheidsgraad

Wat is uw oordeel over de moeilijkheidsgraad van het examen?
Ik vind het examen

te moeilijk	58	10,9%
moeilijk	264	49,7%
niet te moeilijk / niet te makkelijk	203	38,2%
makkelijk	5	0,9%
te makkelijk	1	0,2%
Totaal	531	100,0%

Wat is uw oordeel over de mate waarin het examen de vaardige en minder vaardige kandidaten onderscheidt?
Het onderscheidend vermogen is

voldoende	334	62,9%
onvoldoende	141	26,6%
Kan ik niet beoordelen	56	10,5%
Totaal	531	100,0%

Lengte

Wat is uw oordeel over de lengte van het examen in verhouding tot de tijd die de kandidaat ervoor beschikbaar heeft?
Het examen is

te lang	91	17,1%
precies goed	426	80,2%
te kort	14	2,6%
Totaal	531	100,0%

Taalgebruik

Hoe beoordeelt u de moeilijkheidsgraad van het taalgebruik in het examen?
Ik vind het taalgebruik

zeer moeilijk	45	8,5%
moeilijk	257	48,4%
precies goed	227	42,7%
te gemakkelijk	2	0,4%
Totaal	531	100,0%

Hoe beoordeelt u de moeilijkheidsgraad van het taalgebruik bij de vragen?
Ik vind het taalgebruik bij de vragen

zeer moeilijk	28	5,3%
moeilijk	210	39,5%
precies goed	291	54,8%
te gemakkelijk	2	0,4%
Totaal	531	100,0%

Vraagvorm

Een gesloten vraag is een vraagvorm waarbij de kandidaat moet kiezen uit een beperkt aantal antwoordmogelijkheden die vooraf gegeven zijn. Een open vraag is een vraagvorm waarbij een kandidaat het antwoord zelf moet formuleren. In het examen worden zowel open als gesloten vragen gebruikt. Wat vindt u van de verdeling van beide vraagvormen in het examen? Ik vind de verdeling van de vraagvormen

goed	343	64,6%
niet (zo) goed, ik heb liever een groter aantal gesloten vragen	37	7,0%
niet (zo) goed, ik heb liever een groter aantal open vragen	118	22,2%
niet (zo) goed, ik heb liever uitsluitend gesloten vragen	32	6,0%
niet (zo) goed, ik heb liever uitsluitend open vragen	1	0,2%
Totaal	531	100,0%

Correctievoorschrift

Het correctievoorschrift bestaat uit een algemeen gedeelte en een vakspecifiek gedeelte. In het algemene gedeelte staan de algemene regels voor de correctie van examenvragen. Het vakspecifieke gedeelte bevat het beoordelingsmodel met het bijbehorende scoringsvoorschrift. In welke mate geeft het beoordelingsmodel u voldoende houvast om bij de antwoorden van een kandidaat te bepalen hoeveel scorepunten u moet toekennen? Het beoordelingsmodel geeft mij

bij alle vragen voldoende houvast	141	26,6%
bij de meeste vragen voldoende houvast	371	69,9%
bij de meeste vragen onvoldoende houvast	17	3,2%
bij alle vragen onvoldoende houvast	2	0,4%
Totaal	531	100,0%

Hoeveel tijd hebt u gemiddeld per kandidaat besteed aan de correctie van het examen?

minder dan 30 minuten	351	66,1%
tussen de 30 en 45 minuten	166	31,3%
meer dan 45 minuten	14	2,6%
Totaal	531	100,0%

Vakspecifieke vragen

Het examen is op B1- tot B2-niveau geijkt aan het Europees Referentie Kader (ERK)
In welke mate vindt u de vragen in het examen passend bij B1- respectievelijk B2-niveau van het ERK
Het aantal vragen in het examen dat passend is bij

B1-niveau is te groot	24	4,5%
B2- niveau is te groot	166	31,3%
B1- respectievelijk B2 niveau is in evenwicht	341	64,2%
Totaal	531	100,0%

Open vraag: Hebt u nog opmerkingen of een toelichting?

In totaal hebben 531 docenten deze vragenlijst ingevuld en van die 531 hebben vervolgens nog 162 gebruik gemaakt van de mogelijkheid een opmerking te maken. Hieronder een samenvatting van de opmerkingen die gemaakt zijn. Grofweg waren er wel een aantal (overlappende) thema's te herkennen bij de reacties die kritisch van aard waren.

1 Moeilijkheidsgraad

Opmerkingen in deze categorie zoomden vaak in op het in het examen gebruikte vocabulaire (in vragen en/of teksten), te groot aantal technische vragen, te groot aandeel van vragen op B2 niveau. Daarnaast werd ook een aantal keren aandacht gevraagd voor dyslectische leerlingen, zwakke leerlingen, leerlingen met taalachterstand. In het verlengde hiervan zijn ook de opmerkingen gerubriceerd over de lengte van het examen, dan wel het te grote aantal lange teksten waardoor leerlingen moeite hadden geconcentreerd te blijven. Ook hebben een aantal docenten kritiek op de gebruikte teksten, die volgens hen slecht of niet aansluiten bij de belevingswereld van de kandidaten, met als gevolg lagere scores.

2 Aanvullingen op het CV

Het examen behoeft aanvullingen en dat betekent volgens sommigen dus dat er slecht werk is geleverd. Deze aanvullingen zorgden voor extra werk. De aanvullingen zelf waren onlogisch, behalve de aanvullingen waren er volgens sommigen nog wel meer vragen die van discutabele aard waren. In het verlengde hiervan werd er ook getwijfeld aan de werking van de 2de correctie.

3 De N-term

Uit een aantal opmerkingen blijkt dat er nog veel onduidelijk is over hoe de N-term tot stand komt en wat deze precies doet.

4 Het eindexamen toetst maar één vaardigheid

Docenten vinden het bezwaarlijk dat alleen lezen centraal getoetst worden en dat dit vervolgens voor 50% (of meer) het cijfer voor het vak Engels bepaalt. Lezen is volgens sommigen van hen nu juist de vaardigheid die voor GI/TI leerlingen het minst van belang is. Ook het feit dat dit een verschil tussen CE en SE cijfers kan veroorzaken wordt als bezwaarlijk ervaren. Er worden adviezen verstrekt om dit op te lossen: schrijven terug in het examen, een digitale toets, aangevuld met kijk/luistervaardigheid, een leertoets met grammatica en/of kennis over de Engelse cultuur.

Bij de conclusies zullen we nog nader op bovenstaande kritiek in gaan.

3 Analysegegevens

Bron: www.cito.nl

Kies: voortgezet onderwijs → centrale examens → Examenverslagen → Resultaten per examen 2018
Open de Excel-applicatie.

Door de cel 'P'-waarde totale steekproef' aan te klikken vindt u een omschrijving van het begrip 'P'-waarde'.

Ten aanzien van de representativiteit van de steekproef moet worden opgemerkt dat er sprake is van een selectie uit de populatie aangezien het uitsluitend vakdocenten betreft die een groepsrapportage hebben aangevraagd.

Schooltype/Leerweg	VMBO GL/TL
Soort examen	CSE
Vaknaam	Engels
Totaal aantal kandidaten	60.970
Steekproefgrootte	50.912
Aantal items	41
P'-waarde totale steekproef	0,62
P'-waarde, alleen jongens	0,65
P'-waarde, alleen meisjes	0,60
Maximumscore	45
Gemiddelde score	28,1
Normeringsterm	0,9
Gemiddeld cijfer	6,5
Percentage onvoldoendes (< 5.5)	24,2
Standaardafwijking	7,4
Betrouwbaarheid (asympt. GLB)	0,86
Standaardmeetfout	2,8

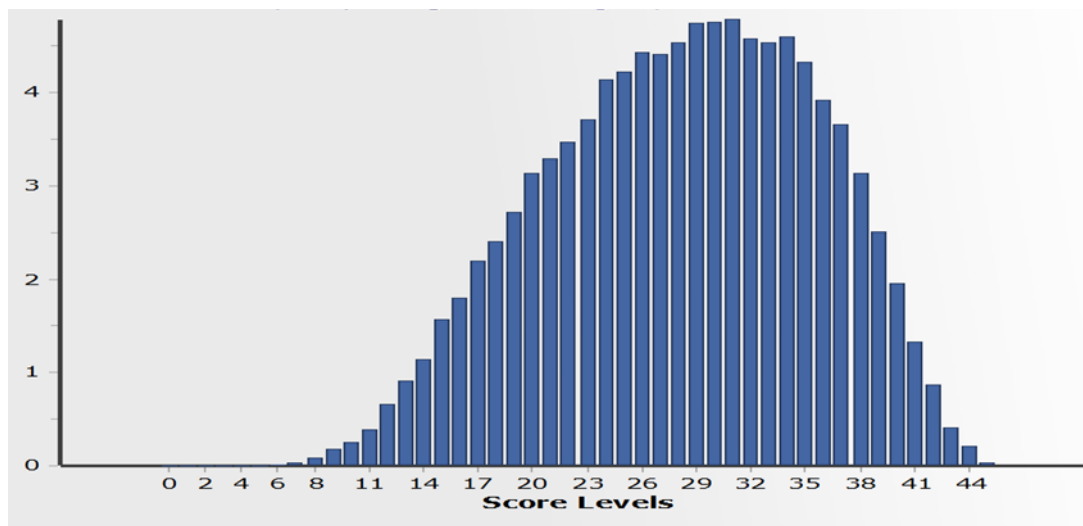
4 Conclusies en aandachtspunten voor de toekomst

Over het algemeen waren de docenten redelijk positief over het centraal examen Engels GL/TL zoals uit de cijfers blijkt. Het examen werd beoordeeld met een 6,25 en het streven is om dat cijfer in de nabije toekomst wat hoger te krijgen. In de laatste jaren heeft het centraal examen Engels GL/TL een ontwikkeling doorgemaakt. Allereerst wordt de zakelijk brief niet meer als verrijksdeel in het centraal examen Engels getoetst. Daarnaast is de laatste jaren het vaardigheidsniveau van de kandidaten voor het vak Engels gestegen, wat vaak resulteerde in een vrij lage N-term. Om een betere spreiding te krijgen tussen de kandidaten is het examen daarom op sommige punten iets moeilijker gemaakt, wat ook een gunstig effect heeft gehad op de spreiding en de uiteindelijk vastgestelde N-term.

De opmerkingen die docenten bij het centraal examen Engels GL/TL maakten kunnen in een paar categorieën ingedeeld worden (zie pagina 10):

- 1 Moeilijkheidsgraad
- 2 Aanvullingen op het CV
- 3 De N-term
- 4 Het eindexamen toetst maar één vaardigheid

Ad 1: Hoewel het streven is om elk jaar vergelijkbare centrale examens te maken kan het niveau per jaar verschillen. Daarnaast kan ook het gemiddelde vaardigheidsniveau van de leerlingen niet ieder jaar gelijk zijn. Sterker nog, we weten uit onderzoeken en metingen dat het vaardigheidsniveau Engels lezen van GI/TL leerlingen de laatste jaren gemiddeld behoorlijk hoger is dan pakweg 10 jaar geleden. De prestatie-eis (wat moet de leerling tenminste aantonen te beheersen voordat we vinden dat deze prestatie een voldoende waard is) is echter ongewijzigd, en de N-term wordt dan ook zo berekend dat bij gelijke prestaties in verschillende jaren de waardering hetzelfde is. Om de verschillen tussen de voldoende en de goede lezers in beeld te brengen zullen we wel een examen moeten voorleggen met vragen daarin die deze verschillen blootleggen, maar het merendeel van de vragen moet door ruim de helft van de groep goed gemaakt kunnen worden. De teksten worden aangeleverd door ervaren docenten. Het constructieteam van docenten maakt ook de opgaven. De vaststellingscommissie van het CvTE, die vervolgens deze teksten en opgaven keurt, bestaat eveneens uit docenten. Hun inschatting was dat zowel de teksten als de opgaven van een bij de GI/TL passend niveau waren. De toets is bedoeld om een onderscheid te maken tussen leerlingen die het beoogde niveau wel hebben en hen die dat, om wat voor reden dan ook, (nog) niet hebben. Gemiddeld haalden de leerlingen 62,47 procent van de scorepunten. De scoreverdeling van GI/TL 2018 was evenwichtig en gebalanceerd en niet anders dan voorgaande jaren.



Docenten kunnen exact zien hoe moeilijk het examen in zijn geheel (en elke vraag apart) is geweest. De statistische gegevens en analyses (TIA's) van dit examen zijn openbaar en hier te vinden: *LINK INVOEGEN*.

ad 2: Het maken van een examen blijft mensenwerk. Alle opgaven zijn in het constructie- en vaststellingsproces op verschillende momenten bekeken en op juistheid gecontroleerd: meerdere docenten, toets-deskundigen, een neerlandicus, een native speaker. Maar dan nog kan het zo zijn dat er iets gemist wordt of dat een bepaalde interpretatie door de screeners over het hoofd wordt gezien of dat leerlingen onvoorziene antwoorden geven. De examens worden gescoord door docenten; we vertrouwen er op dat die de richtlijnen volgen. Als blijkt dat een vraag multi-interpretabel is of dat het model-antwoord niet duidelijk genoeg is, zullen er aanvullingen tijdens de correctieperiode gepubliceerd worden. En soms wordt er dan ook voor gekozen om deze aanvulling wat ruimer te maken om meer recht te doen aan

kandidaten die de essentie van de tekst begrepen hebben, maar net een andere formulering hebben gekozen dan in het CV staat. . Het examen bestaat uit tenminste 40 items om onvolkomenheden te kunnen verdisconteren, zonder dat de betrouwbaarheid van de uitslag aangetast wordt.

ad 3: Scores krijgen pas betekenis als ze omgezet worden in waarderingen, oftewel cijfers. Bij de eindexamens wordt daarvoor een N-term toegepast. Kort en bondig uitleggen hoe de N-term tot stand komt is onmogelijk. Het feit dat deze kan variëren van 0 tot 2 is juist om over de jaren te zorgen dat leerlingen bij een gelijke prestatie eenzelfde waardering krijgen. Er bestaat geen goede of slechte N-term, een N-term moet passend zijn bij het specifieke examen waar het aan gekoppeld is. Hierover meer lezen? Ga dan naar:

<https://www.examenblad.nl/onderwerp/normering-centrale-examens/2018> of
<https://www.examenblad.nl/onderwerp/ik-heb-een-examenklas/2018>

ad 4: Het eindexamen bestaat uit twee delen, het SE en het CE, die beiden even zwaar wegen. Dat het CE voor de moderne vreemde talen alleen leesvaardigheid toetst is wettelijk bepaald in het eindexamenprogramma, waarvoor de minister van OCW verantwoordelijk is. Scholen hebben zelf de vrijheid om het SE in te richten, maar daarvoor zijn wel richtlijnen gegeven. Voor de talen worden expliciet de andere vaardigheden (luisteren, schrijven en spreken) genoemd als SE inhouden, maar op welke manier scholen dit doen staat hen verder vrij. De overheid is bezig met een onderzoek naar een herziening, dat wellicht gaat leiden tot een nieuw curriculum en uiteindelijk een nieuw eindexamenprogramma. Informatie hierover en inspraakmogelijkheden vind je op <http://curriculum.nu>.