

Nog nooit zoveel kandidaten

Het centrale eindexamen vmbo-geschiedenis 2021

Na een jaar overgeslagen te zijn vanwege corona, gingen de centrale eindexamens dit jaar wel door. In hoeverre had corona invloed op de leerprestaties van de leerlingen? Hoe reageerden leerlingen op het examen? Hoe is de normering tot stand gekomen? En speelde de duimregeling een rol? Huub Kurstjens bespreekt het vmbo-examen.

Huub Kurstjens is toetsdeskundige voor de vmbo-examens geschiedenis en staatsinrichting bij Cito.

Het was even wennen: leerlingen die weer in gymzalen of andere ruimtes voorovergebogen over tafeltjes eindexamen deden. Na de afwezigheid van de centrale examens in 2020 was het de vraag of de leerlingen wel voldoende voorbereid waren op de centrale examens. Hadden ze genoeg tijd gehad om te leren? Hadden ze klassikaal onderwijs gemist? Was het digitaal volgen van lessen afdoende? En hoe zat het met de motivatie? Deze ongewone omstandigheden leidden onder meer tot een extra tijdvak. Toch hebben de leerlingen massaal deelgenomen aan het eerste tijdvak. Wat het effect is van de extra herkansingsmogelijkheden of de duimregeling, weten we nog niet precies.

Wat we wel hebben kunnen afleiden uit de toetstechnische gegevens die via het Wolfstelsysteem bij Cito zijn binnengekomen (zie kader), is dat de N-term dit jaar iets hoger ligt dan in andere jaren. Het lijkt erop dat de resultaten van dit jaar gemiddeld genomen iets achtergebleven zijn bij de resultaten die onder normale afnameomstandigheden verwacht zouden mogen worden. Om hier meer over te kunnen zeggen is een verdere analyse van de resultaten nodig. De onderzoeksresultaten hiervan worden in het najaar verwacht.¹

Reacties van leerlingen

In dit artikel staat de leerling centraal. Dat gebeurt met behulp van de reacties die het LAKS verzameld heeft na afloop van het examen, zoals het indienen van een klacht.

Er werd door leerlingen veel geklaagd over de

problemen waarmee ze te maken hadden als gevolg van de *lockdown*. Examenleerlingen mochten wel naar school komen, maar de voorbereiding op het examen voelde toch anders dan normaal. Enkele reacties van leerlingen hierover: 'Vanwege het thuisonderwijs heb ik mij niet fatsoenlijk op het examen kunnen voorbereiden.' 'Ik heb dit jaar amper les gehad en kon me dus amper voorbereiden.' 'Heb door corona echt achterstand.' 'Bijna de helft van mijn toets is nooit in het echt uitgelegd.' 'Tijdens de lessen geschiedenis heeft de docent niet alle onderwerpen even duidelijk uitgelegd.' 'Door die stomme online lessen hebben we de helft gewoon gemist.'

Bij de bespreking van het vmbo-examen geschiedenis in *Kleio* wordt al jarenlang opgemerkt dat de beheersing van de Nederlandse taal te wensen overlaat. Het is algemeen bekend dat Nederlandse leerlingen lezen niet leuk vinden. PISA-onderzoeken laten zien dat de beheersing van het Nederlands achteruit gaat. Er zou zelfs sprake zijn van 'ontlezing'.² Wat duidelijk wordt uit de berichten van docenten is dat leerlingen minder goed zinnen formuleren, minder goed in staat zijn langere teksten te lezen en minder goed de vraagstelling begrijpen dan enkele jaren geleden. In de woorden van de leerlingen zelf: 'De vragen waren vaag en ik snapte gewoon niet altijd waar ze het over hadden.' 'Er waren woorden in die we nooit hebben gehad op school.' 'Ik begreep de meeste woorden niet die werden gebruikt waardoor ik de vraag niet goed kon beantwoorden.' 'In de vraag staat het woord ideologie. Ik heb geen idee wat

¹ In het najaar verschijnt een onderzoek door CvTE (waarbij Cito ook betrokken zal worden). Zie staatsblad: Staatsblad 2021, 136 op officielebekendmakingen.nl.

² Zie bijvoorbeeld het artikel van Adriaan van Dis in *de Volkskrant* van 2 juni 2021 waar hij onder de titel 'Onverschilligheid over ontlezing heerst alom. Moet er een leesknokploeg worden opgericht?' de alarmbel luidt over dit verschijnsel.



dat woord betekent en heb dit ook niet kunnen afleiden uit de zinnen waarin het staat.' 'Dit examen van geschiedenis had VERSCHRIKKELIJK moeilijke woorden, de zinnen waren opgesteld alsof een minister de vragen moest beantwoorden.' Bij het corrigeren van de antwoorden is het voor docenten soms lastig om te bepalen wat ze goed of fout moeten rekenen. Wat doe je bijvoorbeeld als een leerling in plaats van het recht van amendement het recht van amusement opschrijft of van vrijheid van drukpers persdruk maakt, of de naam van Churchill verbastert tot Sjursjill? Het verhaal van de klok en de klepel. Het maakt eens te meer duidelijk dat geschiedenis een talig vak is waarvoor een bepaalde mate van lees- en schrijfvaardigheid vereist is.

Vraagstelling begrijpen

Het voorgaande raakt ook de vraagstelling. Als je de vraagstelling niet goed begrijpt omdat er een bepaalde leesvaardigheid vereist wordt, dan wordt het beantwoorden van die vraag ook lastig. 'Die vraagstellingen waren veel te moeilijk voor wat we hebben geleerd.' 'De vragen waren echt bullshit en alle vragen hadden niks te maken met geschiedenis.' 'Ik snapte niks van de vraagstelling dus heb maar wat ingevuld dit is echt niet leuk!'

‘Wat doe je als een leerling in plaats van het recht van amendement het recht van amusement opschrijft, of van drukpers persdruk maakt?’

'Superwarrige vragen. Ook heel vreemd specifieke vragen over onderwerpen waar je 0,01 milliseconde les over hebt gehad.' 'Het was moeilijk te begrijpen wat ze bedoelen.' 'Ik vond dat de vragen slecht geformuleerd waren. Ik zou het antwoord wel geweten hebben, maar ze hebben er een paar instinkers in gezet.'

Dit soort reacties geeft aan dat de spreektaal van leerlingen en de examentaal soms ver uit elkaar kan liggen. Oefenen met oude examens en de vraagstellingen die daar gebruikt worden, ligt dan ook als oplossing voor de hand. Daarnaast werd er af en toe inzicht gevraagd, 'Hele lastige inzichtvragen.' 'Domme vragen.' Ook het verwoorden van een antwoord vereist een bepaalde taalvaardigheid. Een veel gehoorde klacht over de jaren heen is het verschil tussen de syllabus en de uitwerking daarvan in lesmethoden. Daarom delen veel docenten aan de leerlingen ter voorbereiding van het examen ook de syllabus uit. Daar staat beknopt de leerstof weergegeven die de basis van de eindexamens vormt. De uitwerking van die leerstof en de accenten die daarbij worden aangegeven zijn de verantwoordelijkheid van de educatieve uitgever-

De gegevens van de toets- en itemanalyse (TIA) op basis van de versnelde correctie

Hieronder staan schematisch de gegevens die bij Cito via het Wolf-systeem zijn binnengekomen op basis van de versnelde correctie. Deze gegevens zijn nodig om een afgewogen oordeel te kunnen geven over de aanvaardbaarheid, validiteit en betrouwbaarheid van het examen. Ook is het mogelijk om op basis van deze gegevens een vergelijking te maken met eerder afgenomen examens. Samen met het commentaar dat na afname op de examens gegeven wordt, leidt dit tot de vaststelling van de normering (de N-term) door het CvTE.

Het vmbo-examen 2021 in cijfers (alleen het papieren examen GL/TL)

	GL/TL
gemiddelde score / schaallengte	37,35 / 65
gemiddelde p-waarde*	0,57
vastgestelde N-term*	1,2
% onvoldoende (na vaststelling door CvTE)	21,1
gemiddelde cijfer	6,4
betrouwbaarheid*	0,81
standaarddeviatie*	8,49
aantal kandidaten	22.930
steekproefgrootte	19.801

* De **gemiddelde p-waarde** is een maat voor de moeilijkheidsgraad van een examen. Van eenzelfde orde is de **p'-waarde van een vraag**. Een p of p'-waarde van 0,10 betekent een erg moeilijk examen of een erg moeilijke vraag. Een waarde van 0,90 betekent een erg gemakkelijk examen of een erg gemakkelijke vraag. De p/p' wordt berekend door de gemiddelde score op een examen/vraag te delen door de maximaal haalbare score op dat examen/die vraag. De waardes worden vaak aangegeven als een percentage.

* De **N-term** of de **normeringsterm** is een ijkvariabele die gebruikt wordt om de score op het examen om te zetten in een cijfer. Hoe hoger de N-term, hoe (relatief) moeilijker het examen was.

* De **betrouwbaarheid** is de mate waarin de scores consistent, nauwkeurig en reproduceerbaar zijn, kortom vrij van meetfouten.

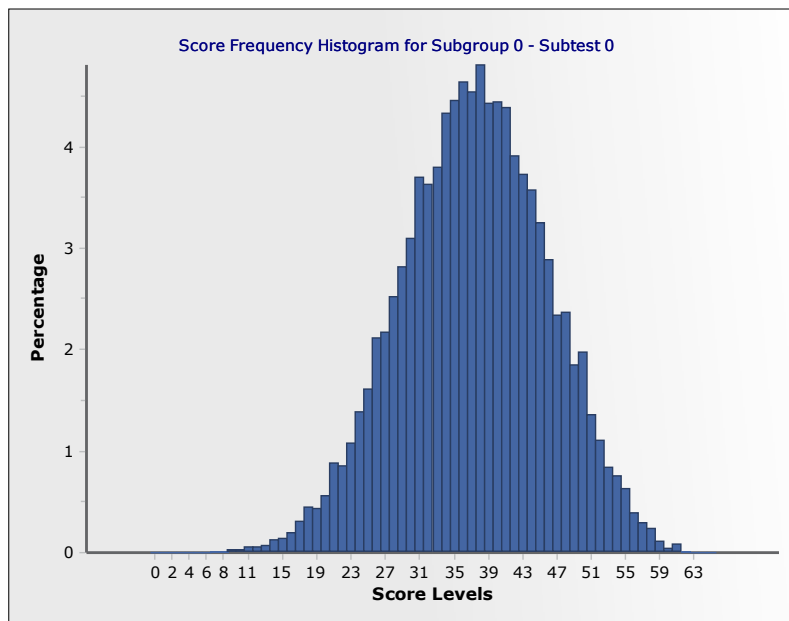
* De **standaarddeviatie** is een spreidingsmaat. Hoe groter de spreiding, hoe beter de hele score-schaal wordt benut.

Zie voor meer informatie de toetstechnische begrippenlijst:

www.cito.nl/static/oenw/ttb/beglist1.htm.

rijen en de docent. Dat er enige ruimte zit tussen de syllabus en de uitwerking ervan in lesstof, maken de volgende citaten duidelijk: 'Het examen kwam niet overeen met de lesstof.' 'Er kwamen vragen in die we niet behandeld kregen in het boek en in de les.' 'Er stonden vragen in die niet tot de examenstof behoren.' 'De vragen waren niet representatief voor de totale leerstof.' 'Er klopte niks van de vergelijking tussen de syllabus en het examen zelf. Van alles wat we hebben moeten ►

leren kwamen er maar een paar dingen overeen.’ Deze reacties geven aan dat er tijdens de voorbereiding van het examen meer aandacht nodig is voor de interpretatie van de eindtermen in de syllabus. Die kan moeilijk aan een leerling zelf overgelaten worden, wat helaas wel vaak het geval is bij onderwijs op afstand. Daarom speelt de instructie van de docent een belangrijke rol bij de ‘vertaling’ van de eindtermen.



Histogram met onderaan het aantal te behalen scorepunten (maximaal 65). De staafjes geven de scorepercentages aan van de leerlingen.

Meer vragen dan anders

Een veel gehoorde klacht betrof de lengte van het examen. ‘Allereerst was het examen veel te lang, de opgaven waren vooral open vragen waardoor je nog meer in tijdnood kwam. Ook zaten er te veel bronnen in.’ ‘De vragen waren prima te doen, dus dat viel mee alleen vond ik 51 vragen wel erg veel.’ ‘Ik heb veel oefenexamens gemaakt en meestal waren er ongeveer 43 tot 45 vragen, maar het examen van dit jaar bestond uit 51 vragen. Dit vond onze docent ook echt megalang.’

Het klopt dat er in het examen van dit jaar een paar vragen meer zaten dan in vorige examens. Dat komt omdat er meer vragen dan gebruikelijk in zaten met slechts één scorepunt. Wat betreft het totaal aantal scorepunten: dat ligt al jaren rond de 65 punten, zo ook dit jaar. Daarnaast wordt voor elk examen het aantal woorden geteld in het vragenboekje én het bronnenboekje. Ook dat week niet significant af van voorgaande jaren: rond de 4000 woorden. Een indicatie over tijdnood zou het overslaan van de laatste paar vragen van het examen kunnen zijn. Die vragen gaven daar geen aanleiding toe. Opmerkelijk was wel in dit examen dat er veel meer vragen werden overgeslagen dan gebruikelijk, met als uitschieter vraag 3. Uit de omschrijving moesten de leerlingen in een open vraag de naam van (Abraham) Kuyper opschrijven. Ruim 10 procent sloeg deze vraag over. We weten

niet of dit kwam door gebrekkig leerwerk of dat er sprake is geweest van de invloed van corona of een combinatie daarvan. Daarom was er voor alle vakken een compensatieregeling bedacht in de vorm van een extra herkansingsmogelijkheid, meer afnamemomenten en de zogenaamde duimregeling. Op de meeste vragen werd geen of nauwelijks commentaar gegeven. Dat is het beste compliment dat je kunt krijgen als toetsenmaker. Een paar vragen sprongen eruit bij dit examen. Twee daarvan worden hieronder kort besproken.

Vraag 25 ging over enkele gebeurtenissen over het leven van de Oostenrijkse kapitein Von Trapp uit de film *The Sound of Music*. Die gebeurtenissen moesten in een juiste chronologische volgorde worden gezet. Belangrijk bij de omschrijving van die gebeurtenissen zijn de signaalwoorden, zoals Oostenrijk-Hongarije, *Anschluss*, Verdrag van Versailles en de economische crisis. Je hoefde de film niet gezien te hebben om de vraag toch goed te kunnen beantwoorden. Het leverde verontwaardigde reacties op bij leerlingen die dit niet gewend waren. ‘Bij de vraag over de *Sound of Music* vond ik dat mensen die die film hadden gekeken echt een voordeel hadden. Ik heb die film zelf persoonlijk niet gekeken en had de vraag dan ook fout. Mijn moeder heb ik de vraag thuis ook laten maken en die had hem goed. Zij had verder niet geleerd. Alleen simpelweg de film wel gezien.’ ‘De vraag ging over de *Sound of Music* en daarin waren wel soort van dingen verwerkt die we geleerd hadden maar als je niet wist wat de *Sound of Music* is, had je veel minder kans op het goed beantwoorden van de vraag dan iemand die wel wist wat de *Sound of Music* is. Nu had je geen mogelijkheid om je daarop voor te bereiden.’ Het was overigens niet de eerste keer dat een chronologievraag op deze manier, met behulp van signaalwoorden, werd gesteld.

Economie of geschiedenis

De vraag die de meeste stof deed opwaaien was vraag 50. Vragen over de Europese Unie roepen traditiegetrouw veel reacties op, want Europa is een ver-van-mijn-bedshow. Daar kwam bij dat er, naast een politieke reden, ook een economische reden werd gevraagd waarom landen in de Europese Unie dezelfde tijd zouden moeten aanhouden. Nog afgezien dat er veel plausibele antwoorden mogelijk waren, was de algemene reactie dat een dergelijke vraag eerder in een examen economie dan in een geschiedenisexamen thuishoorde. ‘Hele onduidelijke vage vraag. Gaat meer over economie dan over geschiedenis.’ ‘Het waren vragen over de politiek van nu en dat heeft niks te maken met geschiedenis.’ ‘Het ging over een tijd en dat heeft niks te maken met geschiedenis.’ ‘Ik snapte de vraag niet en ik vond het heel lastig om een economische en politieke reden in te vullen daarbij, omdat het iets heel anders is dan wat we hebben

3 H. Kurstjens, ‘Koester het centraal eindexamen’, in: *Kleio 2*, maart 2021, pp. 64-69.

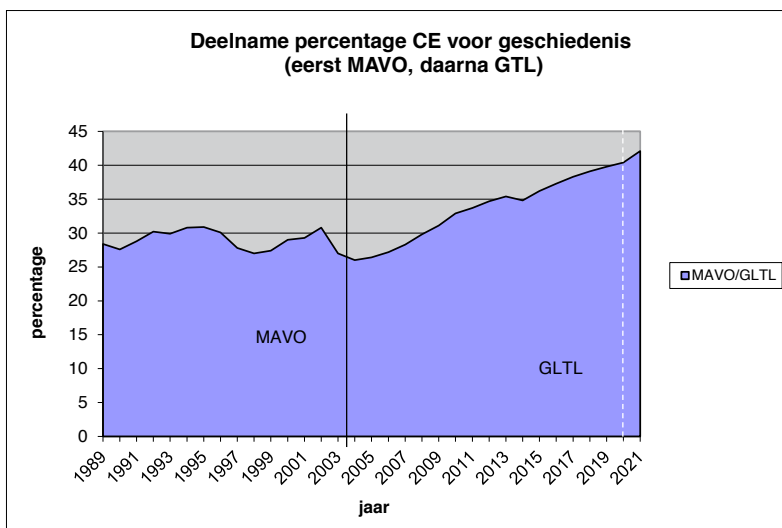
4 B. Haeck en D. Bleus, ‘In de jaren ‘90 begon het al fout te gaan in ons onderwijs’, in: *De Tijd*, zaterdag 5 juni 2021, p. 17.

moeten leren.’ Het antwoord op deze vraag heb ik nooit kunnen leren.’ Deze vraag verdiende niet de schoonheidsprijs. Bij de correctie van deze vraag zijn de docenten coulant geweest.

De eerste indruk van docenten was dat het veel was, maar goed te doen. Bij het corrigeren blijken de antwoorden van de leerlingen en de resultaten toch een beetje tegen te vallen. Veel vragen werden overgeslagen, het verschil tussen goede en minder goede leerlingen was groot (zie histogram), af en toe was het een beetje veel van hetzelfde, een echt leerexamen, soms lastige vraagstellingen. Wat precies de gevolgen van het ‘duimeffect’ zijn geweest, is nog niet duidelijk. Uit reacties van docenten blijkt dat nogal wat leerlingen er bewust voor gekozen hebben om het vak te laten vallen en dus niet of nauwelijks geleerd hebben voor dit examen. Ook werd er opgemerkt dat er meer ‘clustervragen’ rondom een bepaald onderwerp werden verwacht, omdat die waren toegezegd. Daar is inderdaad tijdens een vmbo-studiebijeenkomst over gesproken, maar het maken van een examen kost veel tijd. In de komende examens wordt daar wel een poging toe ondernomen. Vermeden moet worden dat er een stapeleffect ontstaat, dat wil zeggen dat als je de ene vraag fout hebt, ook de andere vraag automatisch fout hebt.

Belevingswereld

Procentueel hebben nog nooit zoveel leerlingen een vmbo(GTL)-eindexamen geschiedenis gemaakt (zie grafiek). Dat is een goed teken! Maar er is ook kritiek mogelijk op de examens. Natuurlijk blijft het maken van examens mensenwerk. De betrokkenheid vanuit het onderwijsveld is erg belangrijk om de examens goed te laten aansluiten op de belevingswereld van de leerlingen. Daarom nodigen we docenten van harte uit om hieraan bij te dragen, bijvoorbeeld als lid van een constructiegroep. Zie tinyurl.com/KCDCE21 voor meer informatie. Bij deze dan ook een oproep om gehoor te geven aan het verzoek om een constructiegroep te versterken. Dat kan door een mail te sturen naar klantenservice@cito.nl, maar ook door onze website met vacatures in de gaten te houden. Met elkaar zorgen we voor een diploma



met een ‘civiele waarde’, een diploma dat algemeen geaccepteerd wordt en dat toegang geeft tot het vervolgonderwijs.

Onder de titel ‘Koester het centraal eindexamen’ werd in maart van dit jaar in *Kleio* teruggeblikt op ruim 40 jaar centrale examens geschiedenis en staatsinrichting in mavo/vmbo.³ Daar werd onder andere gezegd dat er in het buitenland jaloers naar ons examensysteem gekeken wordt. De Belgische onderwijsexpert Dirk Van Damme herhaalde het nog eens in juni van dit jaar: ‘Ik ben geleidelijk aan een sterk voorstander van centrale toetsing geworden. We hebben in Vlaanderen geen goede cultuur van evaluatie. Scholen pakken het heel verschillend aan. Ik denk dat centrale examens nodig zijn om gelijke kansen te bevorderen.’⁴ Een herhaald pleidooi tot koestering van een centraal examen, met alle manco’s die het ook heeft. Wij zijn er samen verantwoordelijk voor. ■

De zwarte verticale streep is het jaar dat de mavo overging in GTL. De witte verticale streep geeft aan dat er in 2020 vanwege corona geen examens zijn afgenomen.

De auteur dankt voor commentaar en inbreng: Pieter van der Bent, Peter Jager, Raymond de Kreek, Lody Smeets en Corine Theuns. Dit was de laatste examenbespreking van Huub Kurstjens die na 28 jaar stopt als toetsdeskundige bij Cito. Reacties kunt u mailen naar Huub.Kurstjens@cito.nl.

