

Van Duitse helmen en verliefd

De vmbo-examens geschiedenis van 2019

Docenten op het examenforum vonden het examen GL/TL 2019-1 een examen met een hoog reproductief karakter, weinig inzicht en (te) veel chronologie. Docenten aanwezig bij de examenbespreking waren veel positiever. En voor het eerst sinds lange tijd werd er nauwelijks geklaagd over een 'talig examen'. Tijd voor een analyse.

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

Hieronder staan schematisch de gegevens die bij Cito via het Wolfsysteem 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 2019 in cijfers (alleen het papieren examen GL/TL):

	GL/TL
gemiddelde score / schaallengte	37,94 / 65
gemiddelde p-waarde*	0,58
vastgestelde N-term*	1,1
% onvoldoende (na vaststelling door CvTE)	21,8
gemiddelde cijfer	6,4
betrouwbaarheid*	0,80
standaarddeviatie*	8,17
aantal kandidaten	24.014
steekproefgrootte	20.525

* 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 gemakkelijke 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 scoreschaal wordt benut.

Zie voor meer informatie de toetstechnische begrippenlijst:

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

Is het mogelijk te slagen voor een examen geschiedenis zonder dat je over parate kennis beschikt? Is het mogelijk om op historische vaardigheden en historisch inzicht getoetst te worden zonder historische kennis? Dit lijken retorische vragen maar jaarlijks duiken ze op bij de bespreking van de examens geschiedenis. Het ene jaar zitten er misschien iets meer reproductieve vragen in het examen dan het andere jaar, maar volgens de deelnemers aan het forum was het aandeel reproductieve vragen deze keer wel erg groot. Het mooie van het forum is dat docenten elkaar kunnen aanvullen of corrigeren.² Dat gebeurde ook deze keer. Een docente schreef het volgende: 'Graag wil ik eerst reageren op het gevoel dat het vooral reproductievragen betreft. Omdat wij bezig zijn met het aanleggen van een toetsenbank waarin we alle examenvragen verwerken, noteer ik bij het maken van de vraag steeds of het een reproductie-, toepassings- of inzichtvraag betreft. Bij sommige vragen is er sprake van zowel R als T, bijvoorbeeld bij vraag 2 die uit twee antwoorden bestaat. Dit heb ik even snel in een Excel-bestand gezet: deze toets telt 7 reproductievragen, 31 toepassingsvragen (daarbij heb ik alle volgordevragen gecategoriseerd als toepassingsvragen, behalve vraag 43 over de geschiedenis van Philips, dat is meer inzicht dan toepassen) en 14 inzichtvragen. Verhouding van ongeveer 1:4:2. Keurig dus.'

Uit de TIA (toets- en itemanalyse, zie kader links) blijkt dat er dit jaar sprake is van een hogere betrouwbaarheid en een grotere standaarddeviatie dan voorgaande jaren. Het examen doet dus wat het moet doen. Leerlingen die geleerd hebben, worden beloond. Het examen is inderdaad een mix van reproductieve vragen, toepassingsvragen en inzichtvragen. Iemand op het forum merkte op: 'Omdat het veel reproductie vraagt, is het niet per definitie gemakkelijker.' Een ander schrijft: 'Ik schrok me dood toen ik het examen voor de eerste keer inzag: wat makkelijk. Ik verwachtte een monsterscore, maar nu ik klaar ben, denk ik dat het wel meevalt.' De inleidende tekst van het verslag van de landelijke examenbespreking begon als volgt: 'Voor het tweede jaar het vernieuwde examen. Aantal

heden

mooie en leuke vragen, ook vragen die talig zijn of inhoudelijk lastig. Meerkeuze- en gesloten vragen zijn goed gemaakt door de leerlingen. De rode draad was goed te vinden in het examen. Het kopje “door de tijd heen” gaf mooi de wisselmomenten van de onderwerpen aan. Staatsinrichting is goed verweven met de rest van de stof.’ Er is dus sprake van een opvallende tegenstelling tussen de teneur op het forum en in het verslag van de examenbespreking. Overigens is het forum niet maatgevend bij de bespreking van het examen. Dat is de landelijke examenbespreking daarentegen wel, maar ook hier geldt dat ‘de status van dit verslag niet dwingend is, slechts adviserend. Het correctievoorschrift is dwingend.’³

Minder schrijven

Wat ook opvallend is, is het veel minder voorkomen van de opmerking dat het examen te ‘talig’ zou zijn. Waarschijnlijk heeft het te maken met het wat meer gesloten karakter van het examen vergeleken met voorgaande jaren. Veel ruimte om afwijkende of aanvullende antwoorden goed te rekenen bood het correctievoorschrift dit jaar niet, behalve bij enkele inzichtvragen. Leerlingen hoefden minder te schrijven en de bronnen waren (nog) meer in leerlingvriendelijke taal aangepast. Dat was ook nodig want uit berichten in de media blijkt dat er zich met betrekking tot de leesvaardigheid van vmbo-leerlingen zorgelijke ontwikkelingen voordoen.⁴ Nederland behoort tot de PISA-landen waar de prestaties van zwakkere en sterkere lezers ver uit elkaar liggen. Dit duidt erop dat de kloof in leesvaardigheid in Nederland in vergelijking met andere landen groot is. Uit PIRLS-onderzoek onder 15-jarigen blijkt bovendien dat Nederland, samen met Frankrijk, nog wel boven het internationale gemiddelde zitten, maar ook tot de weinige landen behoort waar de leesvaardigheid in vergelijking met 2001 achteruit is gegaan, met name bij vmbo-leerlingen.⁵ Een geschiedenisexamen met veel bronnen en inzichtvragen is voor deze leerlingen dan vaak wel lastig. Daarom moet er sprake zijn van enige redelijkheid wat betreft de eisen die aan de leesvaardigheid van een geschiedenisexamen gesteld worden. Net als voor havo en vwo geldt ook voor vmbo dat ‘het geschiedenisexamen het examenprogramma voor geschiedenis toetst, niet voor Nederlands.’⁶ Inzichtvragen waarbij veel gelezen en geschreven moet worden, horen ook in een vmbo-geschiedenisexamen thuis, maar het moet wel in verhouding zijn en in redelijkheid. En dat was bij dit examen ook het geval, gezien de reacties van de docenten en het ontbreken van tijdnoed.

Vraag 16

bron 6

Een verklaring van een Duitse generaal over de afloop van de Eerste Wereldoorlog (1919):
‘Het Duitse leger heeft geen schuld. De prestatie van het leger is bewonderenswaardig. Waar de schuld van de nederlaag ligt, is dus duidelijk. Het is voldoende bewezen dat anderen schuldig zijn. Zelfs onze vijanden zeggen het. Zo heeft een Engelse generaal tegen mij gezegd dat ons leger is verraden.’
Gebruik bron 6.

2p **16** Stel: je doet onderzoek naar de nederlaag van het Duitse leger aan het eind van de Eerste Wereldoorlog en je vindt deze bron.

■ Geef een argument waaruit blijkt dat de bron **bruikbaar** is voor je onderzoek.

■ Geef ook een argument waaruit blijkt dat de inhoud van de bron **niet zo betrouwbaar** is voor je onderzoek.

Doe het zo:

De bron is bruikbaar, omdat ...

De inhoud van de bron is niet zo betrouwbaar, omdat ...

CV:

16 maximumscore 2

Voorbeeld van een juist antwoord is:

De bron is bruikbaar, omdat de generaal een ooggetuige/tijdgenoot was/zijn mening geeft over de nederlaag. **1**

De inhoud van de bron is niet zo betrouwbaar, omdat de (Duitse) generaal de gebeurtenissen opzettelijk anders kan hebben voorgesteld/ hij de schuld (van het verlies) afschuift op anderen. **1**

Opmerking

Het antwoord bij het tweede deel van de vraag dat het de mening van slechts één persoon is, mag niet goed gerekend worden.

Bruikbaarheid en betrouwbaarheid

Een voorbeeld van een inzichtvraag is vraag 16 (zie kader hierboven). Leerlingen moesten een argument geven waarom een bron wel bruikbaar was voor een onderzoek, maar ook een argument geven waarom de inhoud ervan niet zo betrouwbaar was. Veel leerlingen lazen over het woord ‘inhoud’ heen en gaven allerlei antwoorden die er niet toe deden of die gingen over de herkomst van de bron. Groot was ook de verontwaardiging bij de docenten over de opmerking in het correctievoorschrift: ‘Het antwoord bij het tweede deel van de vraag (over betrouwbaarheid, H.K.) dat het de mening van slechts één persoon is, mag niet goed gerekend worden.’ In een eerder examen werd een dergelijk antwoord namelijk wel goed gerekend. Dat laatste klopt, maar is achteraf gezien nogal ongelukkig, want het is historisch-theoretisch niet juist. Het gaat in dit ►

Peter Jager en

Huub Kurstjens zijn toetsdeskundigen geschiedenis bij Cito voor respectievelijk de digitale en papieren examens geschiedenis vmbo (Peter.Jager@cito.nl en Huub.Kurstjens@cito.nl).

¹ Dit thema duikt ook regelmatig op naar aanleiding van de discussies bij #Onderwijs2032 en curriculum.nu. Aleid Truijens, onderwijscolumniste van *de Volkskrant*, merkt hierover op: ‘Vakkennis lekt langzaam weg, alles gaat om ‘betekenisvol’ leren, om communicatie en presentatie. (...) Zonder kennis kun je niet kritisch denken.’ (*de Volkskrant* 8 juni 2019).

² Overigens betreft het een relatief kleine groep docenten, vaak dezelfde mensen, die de afgelopen jaren gebruikmaakt van dit medium. De vraag is hoe representatief deze groep is voor de docenten geschiedenis in het vmbo.

³ Uit: het verslag landelijk VGN-overleg centraal examen geschiedenis en staatsinrichting vmbo t/g 2018-2019, eerste tijdvak.

⁴ www.lezeninhetvmbo.nl en bericht op ad.nl via tiny.cc/K5ADV5.

⁵ www.leesmonitor.nl/nl/leesprestaties-kinderen en www.pisa2018.nl en de column van Aleid Truijens op *volkskrant.nl* via tiny.cc/VKAT.

⁶ Zie ook het artikel van Stefan Boom en Saskia de Boer, *Historisch denken in de centrale eindexamens, Kleio* 3, mei 2019, p. 70-72.

Vraag 11

bron 2

Duitse soldaten op weg naar het front.



Gebruik bron 2.

1p **11** Hoort de foto bij het **begin** of bij het **eind** van de Eerste Wereldoorlog? Verklaar je antwoord met behulp van de foto.

Doe het zo:

De foto hoort bij het ... (kies uit: begin / eind) van de Eerste Wereldoorlog, want ...

CV:

11 **maximumscore 1**

Voorbeeld van een juist antwoord is:

De foto hoort bij het begin van de Eerste Wereldoorlog, want dat blijkt uit het enthousiasme/de opgewekte blik van de soldaten/de vrouwen/aan het eind van de oorlog was niet iedereen meer zo enthousiast/uitgelaten / waren de mensen droeviger/verdrietiger.

Opmerking

Alleen als na een juiste keuze een juist onderdeel van de foto volgt, wordt 1 scorepunt toegekend.

geval immers om representativiteit. Als het de mening van één persoon betreft, moet je wel eerst vaststellen in hoeverre deze representatief is, maar deze kan best heel betrouwbaar zijn. Lastig voor vmbo-leerlingen, maar een antwoord moet historisch-theoretisch wel correct zijn. De vraag scoorde overigens met een p'-waarde van .55 niet eens zo heel slecht.

Jongens en meisjes

Opmerkelijk waren de verschillen tussen de antwoorden van jongens en meisjes. Zo waren er jongens die de vragenmakers te slim af waren bij vraag 11 (zie kader hierboven). Bij de vraag of een foto van Duitse soldaten die op weg zijn naar het front gedateerd moest worden aan het begin of het einde van de Eerste Wereldoorlog, waren er enkele jongens die in hun antwoord ingingen op de verschillende helmen van de Duitse soldaten. De foto móest wel van het begin van de Eerste

Wereldoorlog zijn, want toen droegen de soldaten een *Pickelhaube* en niet de *Stahlhelm* die pas later in de Eerste Wereldoorlog gedragen werd. Daar zou je bijna een bonuspunt voor willen geven. Meisjes scoorden iets beter bij vraag 32, over een Nederlandse vrouw die tijdens de Tweede Wereldoorlog verliefd werd op een Duitse soldaat (zie kader). Sommige mensen zien haar als collaborateur, anderen niet. Meisjes reageerden verontwaardigd op de veronderstelling dat verliefd worden, op wie en onder welke omstandigheden dan ook, als collaboratie gezien kan worden. Dat overkomt je nu eenmaal. Het empathisch vermogen van de meisjes overtrof dat van de jongens. Overigens scoren jongens bij de vmbo-examens geschiedenis al jaren achter elkaar iets beter dan meisjes. Iets om nader te onderzoeken?

Vernieuwing

Dit examen was het tweede examen op basis van een vernieuwde syllabus. Docenten waarden de vernieuwing in de syllabus en in het examen. Een examen dat opgebouwd is op basis van chronologie en dat voorzien is van tussenkopjes en 'Door

Vraag 32

bron 4

Een vrouw vertelt over haar ervaringen tijdens de Tweede Wereldoorlog in Nederland:

Ik was goed op school, maar als er een schoolfeest was zat ik altijd alleen. En toen kwamen er overal jonge Duitse soldaten die 'mooi meisje' tegen me zeiden. Als ik ging dansen of in de schouwburg was, kreeg ik ineens aandacht. Eén van die Duitsers nodigde me uit voor een feestje bij hem thuis. Hij was kapitein en werkte in concentratiekamp Vught. We kregen verkering. Zijn chauffeur bracht me na elke afspraak 's avonds weer naar het station om naar huis te gaan.

Gebruik bron 4.

2p **32** Tegenwoordig vinden sommige mensen dat de vrouw een collaborateur was, anderen vinden van niet.

■ Geef voor beide meningen één argument.

Doe het zo:

De vrouw was een collaborateur, omdat ... (geef een argument).

De vrouw was géén collaborateur, omdat ... (geef een argument).

CV:

32 **maximumscore 2**

Voorbeeld van een juist antwoord is:

De vrouw was een collaborateur, omdat ze een relatie had met een Duitser/de vijand. **1**
De vrouw was géén collaborateur, omdat verliefd worden op/aandacht krijgen van iemand geen collaboratie is/ze zelf niet in het kamp/met Duitsers had (samen)gewerkt. **1**

de tijd heen'-blokjes wordt positief beoordeeld. Datzelfde geldt voor de integratie van staatsinrichting in het Historisch Overzicht. Vraag 22 (p' .43) is daarvan een voorbeeld. In het antwoord moest aangegeven worden welk Nederlands klassiek grondrecht zou botsen met een situatie zoals die bestaat in totalitaire staten. De integratie van staatsinrichting en Historisch Overzicht was ook terug te vinden in vraag 27 (p' .94) en 28 (p' .33) waarbij een vergelijking gemaakt moest worden tussen een rechtsstaat en een totalitaire staat. Dit zijn bestaansverhelderende vragen die niet makkelijk zijn, maar wel betekenisvol zijn voor onze huidige samenleving.

Veel klachten van docenten en leerlingen gingen over de chronologievragen (33, 34, 42, 43, 48). Het betrof vooral de grote hoeveelheid vragen, de aard van de vraagstelling en de plaats van de vragen in het examen. Met name vraag 43 (p' .31), over de geschiedenis van Philips maakte veel los. Via het LAKS kwamen klachten binnen als 'Het ging over het bedrijf Philips en dat heeft niks te maken met geschiedenis', 'De stof is nooit behandeld, niks hierover staat in mijn examenbundel of geschiedenisboek', 'Focking Phillips', 'Vijf hele zinnen op chronologische volgorde. Hoe zie ik in welke tijd het speelde?' Maar ook dit soort vragen is onderdeel van de vernieuwing. Het gebruik van nieuwe contexten en het herkennen van vakinhoudelijke signaalwoorden horen daarbij. Overigens is in het examen van GL/TL 2017-1 een soortgelijke vraag gesteld (vraag 44) maar dan over de geschiedenis van Albert Heijn. Chronologie is onlosmakelijk verbonden met het vak geschiedenis, maar in dit examen stonden inderdaad wel erg veel chronologievragen. In de komende examens zullen er wat minder worden opgenomen. De plaats waar deze vragen in het examen staan, is gekoppeld aan inhoud en tijd. Omdat het examen chronologisch is opgebouwd, staan deze vragen op de plek waar de laatste verwijzing in tijd in de vraag betrekking op heeft. Dat is vaak aan het eind van een examen, waar teruggeblikt wordt op de voorafgaande perioden. Daar zal dus niet veel aan veranderen. Ook de opbouw van de vraag en de vraagformulering zullen niet aangepast worden. Het betreft immers een historische vaardigheid en die is onderdeel van de vernieuwing.

Digitale examens

Zoveel feedback, zowel positief als negatief, als er gegeven wordt op het examen GL/TL, zo weinig reacties zijn er op de digitale examens van de leerwegen BB en KB. Het gaat hier om een groep van slechts een paar honderd kandidaten. Toch zal er bij deze examens ook wel ruimte zijn voor verbetering en daarom is het soms jammer dat er wat betreft het vak geschiedenis zo weinig bekend is over deze doelgroep. Bijna alle kandidaten voor

het vak geschiedenis van de leerwegen BB en KB maken inmiddels de digitale variant van het examen. Het zou dus interessant zijn om te onderzoeken wat voor leerlingen de voor- en nadelen zijn van het maken van een digitaal examen ten opzichte van een papieren examen. Een voordeel lijkt te zijn dat digitale vragen directer zijn en eenvoudiger te corrigeren. In een digitaal examen is verder vaak meteen duidelijk wat een leerling moet doen, zoals het in volgorde slepen van afbeeldingen. Voor een vergelijkbare vraag in een papieren examen is meer uitleg nodig, waardoor de vraag al snel taliger wordt. Nog een voordeel is mogelijk dat digitale examens levendiger worden door het gebruik van afbeeldingen in kleur en filmfragmenten. Onderzoek moet uitwijzen welke voordelen leerlingen daadwerkelijk hebben bij het maken van digitale examens vergeleken met papieren examens.



Docenten gezocht

Tot slot een oproep. Docenten zijn erg betrokken bij het wel en wee van hun leerlingen en hun vak. Dat komt in het bijzonder bij het centrale examen tot uiting. Deze examens worden in opdracht van CvTE en onder supervisie van Cito gemaakt door docenten die ervaring hebben in examenklassen. Het produceren van goede examens kan alleen plaatsvinden als docenten, op basis van een financiële vergoeding of detachering, hieraan meewerken. Helaas merken wij dat de animo voor dit werk de laatste jaren terugloopt. Ook is er nog weinig belangstelling voor itemconstructie op afstand, de pilot docentenparticipatie. Mocht je interesse hebben en denk je een constructieve bijdrage te kunnen leveren aan de geschiedenis-examens in de toekomst, meld je aan bij een van de auteurs. Dat komt ons vak en de kwaliteit van de examens ten goede. ■

Winkelpersoneel bekijkt de destijds net nieuwe compact disc (cd), mede ontwikkeld door Philips. Foto: Marcel Antonisse, Anefo/ Nationaal Archief, 1 maart 1983.