

Inkoppertje blijkt instinker

De vmbo-examens geschiedenis van 2012

Weer meer examenkandidaten dan de afgelopen jaren. Voor het vijfde jaar

'Staatsinrichting' en het 'Historisch Overzicht vanaf 1900' als examenonderwerpen.

Veel schrijven, een lang examen, veel inzichtvragen, maar niet al te moeilijk. Of toch wel?

Wat waren de reacties en hoe zijn de examens van het vmbo dit jaar ontvangen?



Huub Kurstjens is toetsdeskundige geschiedenis bij Cito. Huub.Kurstjens@Cito.nl. De auteur dankt Stefan Boom, Niek Nelissen en Delphine Kerkmeester voor hun opmerkingen en bijdragen.

► Zie: Alderik Oosthoek en Anne Hertman, Eindrapportage eindexamenklachtenlijn 2012, Landelijk Aktie Komitee Scholieren, Amsterdam 2012, pag. 9.

► Zie: verslag landelijk VGN-overleg Centraal Examen geschiedenis en Staatsinrichting vmbo TL/GL 2011-2012, eerste tijdvak. (De directe link staat op de Kleio-website.)

► De digitale examens met één vast afnamemoment in GL/TL, Havo en VWO worden voorlopig niet afgenomen; de digitale examens met flexibele afnamemomenten worden wel gewoon afgenomen.

► Een nadere toelichting van toetstechnische begrippen staat op de website van Cito. (De directe link staat op de Kleio-website. Daar staat ook tabel 1.)

Zelden waren de reacties zo uiteenlopend als dit jaar. Direct na de afname van het examen stonden op internet commentaren van leerlingen die varieerden van 'te moeilijk' tot 'makkie', van 'veel kennis' tot 'stond niet in het boek' en van 'snel klaar' tot 'te weinig tijd'. Het LAKS meldde dat er 'over Geschiedenis twee keer zoveel klachten ingediend zijn in vergelijking met voorgaande jaren (2011 buiten beschouwing gelaten). Bijna 70% van de totaal 1.097 klachten gingen over de moeilijkheidsgraad van het examen.'¹ Docenten waren over het algemeen tevreden, maar vonden het examen aan de oppervlakkige kant en (te) makkelijk.² Waardoor is dit grote contrast in belevingen te verklaren?

Herkenbaar

Vmbo-leerlingen willen over het algemeen graag examens met een hoog reproductiegehalte. Wat ze in de klas gehad hebben en wat in het boek staat, moet liefst herkenbaar terugkomen in het examen. Maar de tijd van jaartallen uit het hoofd leren, is voorbij; met geleerde kennis moet wat gedaan worden. Leerlingen die niet met vaardigheden geoefend hebben, kunnen het moeilijk krijgen bij het centraal examen. Er zijn overigens mogelijkheden te over voor leerlingen die wel willen oefenen. Inmiddels wordt er al vijf jaar gewerkt met dezelfde examenonderwerpen. Ieder jaar zijn er twee tijdvakken openbaar, dus na vijf jaar zijn er tien examens beschikbaar als oefenmateriaal. Voor de meeste docenten is de stof dus gesneden koek. Voor leerlingen daarentegen is het examen nieuw. Vmbo-leerlingen onderschatten vaak wat ze kennen en kunnen. Ze worden onzeker als in het examen nieuwe bronnen of gegevens staan. Dit verklaart grotendeels het verschil in perceptie na afname van een examen.

Makkelijk examen

Cito verzamelt na afloop van het examen alle gegevens via Wolf. De Toets- en Itemanalyse (TIA) wees vervolgens uit dat het examen aan de makkelijke kant was (zie tabel 1). Wellicht speelde de aangepaste zak-slaagregeling hierbij een rol, waardoor leerlingen een grotere druk voelden om te presteren. Misschien was er zóveel oefenmateriaal voorhanden dat leerlingen zich beter konden voorbereiden. Mogelijk is ook het examenprogramma na al die jaren te voorspelbaar geworden. Denkbaar is ook dat de beperkte diepgang van het examenprogramma de examenmakers parten gespeeld heeft bij het maken van hun vragen. Wie het weet, mag het zeggen.

Na het invoeren van hun gegevens in Wolf hebben docenten de mogelijkheid om een korte indruk van het GL/TL-examen te geven. Deze reacties worden verzameld in de *quick-scan*, die laat zien dat 69% van de docenten het examen niet te moeilijk en niet te makkelijk vond, 81% vond de lengte precies goed, 92% vond de inhoudelijke aansluiting van het examen bij het gegeven onderwijs voldoende tot goed en de gemiddelde waardering voor het examen was 6,82, wat hoog te noemen is. Toch kijken we met gemengde gevoelens terug op het examen: het examen discrimineerde onvoldoende en was te makkelijk. De N-term van 0,3 zegt genoeg; een van de laagste van alle vakken op GL/TL (gemengde en theoretische leerweg). Overigens: de N-term wordt niet door Cito, maar door de CvE bepaald. Cito construeert en CvE normeert.

BB- en KB-digitaal

Voor de basisberoepsgerichte leerweg (BB) en de kaderberoepsgerichte leerweg (KB) worden digitale examens afgenomen. Papierexamens komen daar bijna niet meer voor.

Het aantal kandidaten voor geschiedenis in BB en KB is gering (zo'n 200 per leerweg), zeker vergeleken met GL/TL met zo'n 19.000 kandidaten (zie tabel 2). Toch verloopt de invoering van de digitale examens trager dan oorspronkelijk gedacht. In technisch opzicht zijn er bij de afname nog enkele onvolkomenheden. Docenten en leerlingen die ermee gewerkt hebben, zijn over het algemeen zeer te spreken over deze examens. Vooral de toepassingsmogelijkheden bij computerexamens zijn vele malen groter dan bij papieren examens (denk maar aan het gebruik van kleur en filmpjes). Gezien de recente afnameproblemen zal het nog wel enige jaren duren voordat ook in de andere leerwegen de digitale examens breed zullen worden ingevoerd.³

GL/TL: enkele indrukken

Het onderwerp Staatsinrichting was in GL/TL dit jaar (opnieuw) makkelijker (p-waarde 69,8⁴) dan het Historisch Overzicht (p 63,5). De chronologievragen (p 37,1) en de Europa-vragen (p 28,5) scoorden, zoals we overigens al jaren gewend zijn, beduidend lager dan het gemiddelde (p 65,4). In tegenstelling tot andere jaren scoorden de verrijgingsdeelvragen dit jaar hoger (p 65,3). Sommige vragen getuigden volgens docenten van een zekere creativiteit, maar schoten hun doel voorbij, omdat ze niet voldoende discrimineerden. Zo gaven bij vraag 40 bijna alle kandidaten een juist antwoord op de vraag om een reden te

'De invoering van digitale examens verloopt trager dan gedacht'

geven voor het behouden van de laatste stukken van de Berlijnse Muur (p 89). En bij vraag 45 konden leerlingen een persoon kiezen uit het Historisch Overzicht en aangeven waarom zij vonden dat die van grote historische betekenis is geweest (p 86). Er zit altijd wel een persoon bij waarover een leerling historisch gezien iets zinvol weet te melden. Aangezien Staatsinrichting als eerste onderwerp aan bod komt en het examen chronologisch is opgebouwd, gaan de eerste vragen bijna altijd over de Grondwet van 1848. De als makkelijk bedoelde instapvraag pakte anders uit. Leerlingen interpreteerden een schematische tekening verkeerd. Vermoedelijk dacht een groot deel van de leerlingen op basis van wat ze geleerd hadden over 1848, dat het ging ►

Tabel 1: De gegevens van de Toets- en Item Analyse op basis van de versnelde correctie

In deze tabel staan schematisch de gegevens die bij Cito via het Wolfssysteem zijn binnengekomen op basis van de versnelde correctie. Deze gegevens zijn onmisbaar 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 CvE.

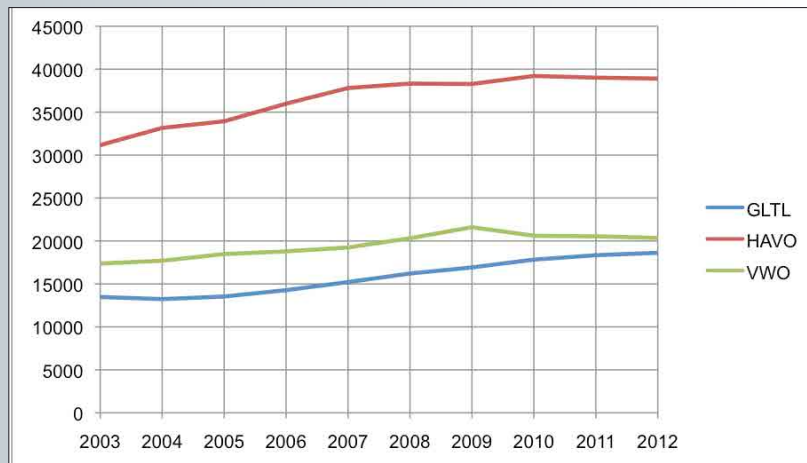
het vmbo-examen 2012 in cijfers (alleen de papieren examens):

	GL/TL	KB
gemiddelde score / schaallengte	41,22 / 63	31,77 / 57
gemiddelde p - waarde*	65,4	55,7
vastgestelde N-term*	0,3	1,2
% onvoldoende (na vaststelling door CvE)	24,0%	21,2%
betrouwbaarheid*	0,71	0,73
standaarddeviatie*	6,75	6,57
aantal kandidaten	19397	218
steekproefgrootte	15529	66

NB De gegevens voor KB zijn veel minder betrouwbaar dan voor GL/TL door het geringe aantal kandidaten en een veel kleinere steekproef.

- De p'-waarde is een getal tussen 0 en 1 dat de moeilijkheidsgraad van een opgave weergeeft. De p'-waarde wordt berekend door de gemiddelde score op een opgave te delen door de maximaal haalbare score op die opgave. Een opgave met een p'-waarde van .10 is erg moeilijk; een opgave met een p'-waarde van .90 is erg gemakkelijk. In de tabel zijn de p'-waarden aangegeven als percentages.
- De N-term is de normeringsterm, een ijkvariabele die ervoor moet zorgen dat aan kandidaten in verschillende jaren en tijdvakken even zware eisen worden gesteld. De besluitvorming over de normering verloopt in drie fasen:
 - 1 De vakmedewerker van Cito geeft een advies aan de vaksectie van het CvE, op basis van de analyses van de afnamegegevens.
 - 2 De vaksectie voegt aan dit advies een inhoudelijke beoordeling van het examen toe. Die komt tot stand op basis van de reacties en commentaren van bijvoorbeeld vakverenigingen, platformen en het LAKS, maar ook rechtstreeks van docenten en leerlingen. De combinatie van het Cito-advies en de inhoudelijke beoordeling leidt tot een normeringsadvies van de vaksectie aan de CvE-leiding.
 - 3 Het dagelijks bestuur van het CvE verzamelt en beoordeelt de adviezen en stelt uiteindelijk de definitieve normering vast. Het resultaat is te zien in de (bovenstaande) vastgestelde N-term.
- De betrouwbaarheid is de mate waarin de scores consistent, nauwkeurig en reproduceerbaar zijn, kortom vrij van meetfouten. Bij een betrouwbare meting zal het resultaat dus niet beïnvloed mogen zijn door storende factoren met betrekking tot de toets, kandidaat en beoordelaar, zoals het tijdstip van de toetsafname, de specifieke vormgeving van de toets, pech of geluk met de opgaven. De mate waarin een meting ongevoelig is voor bepaalde, versturende factoren, kan worden geschat door berekening van een betrouwbaarheidscoëfficiënt. De waarde van een betrouwbaarheidscoëfficiënt ligt tussen 0 en 1. Hoe dichter het cijfer bij 1, hoe betrouwbaarder het examen is.
- De standaarddeviatie zegt iets over de spreiding van het hele examen: liggen de resultaten dicht bij elkaar (dan hebben de kandidaten ongeveer dezelfde resultaten) of is er sprake van spreiding van de resultaten. Hoe groter het cijfer, hoe beter de spreiding.

Tabel 2: aantal examenkandidaten geschiedenis in het voortgezet onderwijs in vmbo-GL/TL, havo en vwo (2003-2012)



► Zie ook: Huub Kurstjens, *Niet gemakkelijk, maar wel goed te doen*. De vmbo-examens geschiedenis van 2011 (in: Kleio, nov. 2011, pag. 44). Ook de resultaten van een peiling tijdens een druk bezochte vmbo-studiedag in Utrecht in november 2011 laten niets aan duidelijkheid te wensen over.
► Zie de syllabus van 2013. Te downloaden via www.examenblad.nl (aldaar doorklikken naar vmbo-geschiedenis-syllabus 2013).

over de grondwetswijziging en dus over invoering van de ministeriële verantwoordelijkheid. Maar de situatie sloeg op de situatie vóór 1848: de minister die verantwoording aflegt aan de koning (vandaar de buiging). De als in-koppertje bedoelde vraag bleek moeilijk, een onbedoelde 'instinker': 70% van de leerlingen beantwoordde deze vraag fout.

Bij vraag 19 en 20 hoorde een communistische prent uit Rusland (1920) met als titel: 'De rode ploeger'. De vragen werden door de docenten positief gewaardeerd, maar de vindingrijkheid van de leerlingen bij de beantwoording overtrof alle verwachtingen. Daardoor hadden docenten een probleem bij de puntentoekenning van deze vragen. Bij vraag 19 was het mogelijk om voor ieder juist onderdeel van het antwoord van de vraag een scorepunt toe te kennen. Bij vraag 20 kon dit niet, omdat ieder gegokt onderdeel van de bron een scorepunt zou hebben opgeleverd. Vandaar dat bij deze vraag uitsluitend 0 of 2 punten (bij een juiste combinatie van bron met uitleg) toegekend konden worden. Wolf geeft standaard alle antwoordmogelijkheden aan, in dit geval 0-1-2, maar één punt toeken-

nen zou dus niet hebben gemogen, terwijl dat in 7% van de gevallen wel gebeurde. In de toekomst moet het correctievoorschrift hiervoor nog duidelijker instructies geven.

Overigens mogen docenten zelf ook wat kritischer zijn in wat ze wel en niet goed rekenen. Een docent merkte hierover op het forum op: 'De antwoorden moeten wat mij betreft historisch gezien wel kloppen. De Nederlandse taal kan een goed antwoord helaas in de weg zitten, maar onze geloofwaardigheid als corrector is ook belangrijk.'

Diepgang

Sommige docenten vonden het examen oppervlakkig. Het nieuwe van het examenprogramma is er een beetje af en diep doorvragen is er niet bij. Voorbeeldexamens zijn in ruime mate voorhanden. Hoe te voorkomen dat het examen een verzameling trucjes wordt? De vorige wijziging van het examenprogramma, de invoering van het 'Historisch Overzicht vanaf 1900', dateert van vijf jaar terug. Staatsinrichting gaat zelfs al mee vanaf 2002. De keuze van beide onderwerpen wordt breed gedragen. Docenten vinden de onderwerpen relevant en betekenisvol voor vmbo-leerlingen.⁵ Toch lopen toetsenmakers tegen grenzen aan: wat kan wel en wat kan niet worden bevraagd? Hoeveel voorkennis mag er verondersteld worden bij leerlingen? Niet de schoolboeken, maar het examenprogramma is hierbij maatgevend.

'De vindingrijkheid van de leerlingen overtrof alle verwachtingen'

Tegelijkertijd mogen we leerlingen niet overvragen: de onderwerpen zijn omvangrijk, zeker voor vmbo-leerlingen, en de behandeling in de klas vergt veel tijd. Maar het gaat dan ook wel om de helft van het eindcijfer.

De komende jaren zijn enkele veranderingen te verwachten. Om te beginnen bevat het examenprogramma van 2013 aanvullingen en verduidelijkingen. Het gaat nadrukkelijk niet om extra stof, maar om een nadere toelichting van de bestaande examenstof.⁶ In de syllabus zijn de wijzigingen geel gemarkeerd. Daarnaast wordt het examenprogramma de komende jaren aangepast. Niet de onderwerpen wijzigen, maar de ordening en de structuur wordt aangepast. Wanneer dit allemaal zijn beslag gaat krijgen, is nog onbekend, maar dat zal zeker niet vóór het examenjaar 2014 zijn. ■

Een communistische prent uit Rusland (1920) met als titel: 'De rode ploeger'.

