

'Niet gemakkelijk, maar wel goed te doen'

De vmbo-examens geschiedenis van 2011

De examenonderwerpen 'Staatsinrichting' en 'Historisch overzicht vanaf 1900' blijken voor het vmbo gesneden koek te zijn. Docenten en leerlingen weten waar ze aan toe zijn. Toch is het altijd weer afwachten of de examens voldoende aansluiten bij de diverse onderwijsniveaus en bij het onderwijs in de klas.

Huub Kurstjens is toetsdeskundige geschiedenis bij Cito. Met dank aan Stefan Boom, Niek Nelissen, Jan van den Berg en Lody Smeets voor hun opmerkingen en bijdragen.

Nog geen twee uur na afname van de examens verschenen de eerste reacties op het forum van de VGN. 'Ik vond dit examen niet gemakkelijk, maar wel goed te doen. De leerlingen die geleerd hadden en het hele jaar goed meegedaan hadden, kunnen hier echt een voldoende bij halen. Ik vond dat er enkele erg mooie vragen bij zaten,' aldus een docente. Als we kunnen afgaan op de site www.examenklachten.nl liepen de meningen van de leerlingen sterk uiteen. Zo schreef een kandidaat: 'Ik vreesde het ergste, maar het viel mee. Als je had geleerd hiervoor was het appeltje eitje!' Een ander mopperde: 'Pfff Vragen niet fatsoenlijk gesteld. Was heel onbegrijpelijk vond ik. Ik heb die antwoorden eens bekeken en ik heb vaak helemaal anders antwoord gegeven.' De recensent voor Kleio concludeerde: 'Dat het examen aan het einde wat 'zwaar' was doet geen afbreuk aan de prima kwaliteit ervan.'¹

Was kort na de afname van het examen het oordeel van veel docenten dat het examen goed te doen was, na correctie van de examenwerken werd dat beeld enigszins genuanceerd. Een docent verzuchtte: 'Ik schrok van de scores, omdat ik het een goed examen vond.' Een ander moest zijn frustratie even kwijt: 'Hoe krijgen we die leerlingen zover dat ze gewoon doen wat wij vragen? Alles is aan bod geweest in de lessen zelfs de vragen waarvan ik denk: inkoppertje ... Helaas! Blijft af en toe een frustrerend beroep.'

De door Cito verzamelde gegevens (de Wolf-gegevens) na afloop van het examen en de op basis daarvan door de CvE vastgestelde N-term bevestigen het beeld van een niet al te moeilijk examen, maar met een goede spreiding tussen de goede en de minder goede kandidaten (zie kader).

Ontwikkelingen in het vmbo

De meeste vmbo-leerlingen die geschiedenis-examen doen, volgen de theoretische leerweg. Net als in de afgelopen zes jaar is het aantal kandidaten duidelijk gestegen. Dankzij deze stijgende lijn doet nu meer dan eenderde van alle kandidaten in de theoretische leerweg het geschiedenisexamen, het hoogste percentage in de afgelopen 25 jaar.²

Neemt het aantal kandidaten geschiedenis gestaag toe, de samenstelling van de leerlingenpopulatie als geheel verandert. Een deel van de leerlingen die vandaag de dag de TL volgen, zou dat niveau een jaar of tien geleden waarschijnlijk niet gehaald hebben. Er is sprake van een 'opwaartse leerlingenstroom'. Tien jaar geleden zat 60 procent van alle leerlingen in het voortgezet onderwijs in het vmbo en 40 procent in havo/vwo, nu zijn die getallen respectievelijk 50 procent en 50 procent. Opvallend is het scoreverschil tussen de kaderberoepsgerichte leerweg (KB) en de gemengde en theoretische leerweg (GTL). De helft van de examenvragen is identiek. Op deze vragen scoorden GTL leerlingen een gemiddelde p-waarde van 67,6, terwijl KB-leerlingen voor dezelfde vragen een gemiddelde p-waarde van 49,6 scoorden, een verschil dat niet eerder zo groot was.

Taal

Steeds meer blijkt dat zowel het begrijpend lezen als de woordenschat van de kandidaten niet overeenkomt met het gewenste beheersingsniveau. En juist deze beheersing van het Nederlands is cruciaal voor het vak geschiedenis. Bij het maken van examenvragen wordt geprobeerd rekening te houden met het taalniveau van de doelgroep. Bronnen worden bewerkt om ze toegankelijk te maken voor

► Uit: Kleio 4, juli 2011, pag. 60 (bespreking vmbo-examen door Dennis Koolmoes)

► Redenen voor deze sterke toename staan in mijn artikel in Kleio nr. 6, sept. 2009, pag. 29.

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

In deze bijlage staan schematisch de gegevens die bij Cito via het Wolfsysteem 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 2011 in cijfers

	GL/TL	KB
gemiddelde score / schaallengte	37,70 / 61	26,7 / 54
gemiddelde p' waarde*	61,8	49,5
vastgestelde N-term*	0,4	1,6
% onvoldoende (na vaststelling door CvE)	31,9%	37,6%
betrouwbaarheid*	0,75	0,69
standaarddeviatie*	7,19	6,59
aantal kandidaten	18341	166
steekproefgrootte	14922	133

N.B. 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 het dagelijks bestuur van het CvE.

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. In een histogram is de grafische weergave van deze spreiding te zien.

► Onduidelijk is of dat te maken heeft met de omvang van de 10 tijdvakken of dat er geen nadere specificaties zijn, dus geen nader houvast is voor vmbo-leerlingen, bij de kenmerkende aspecten. Nader onderzoek hiernaar is gewenst.

► De *quick-scan* is een korte vragenlijst die ingevuld kan worden bij het opsturen van de scoreresultaten via het WOLF-systeem naar Cito. De gemiddelde score voor het examen bedroeg 6,73 (bij N 637 = 100). In lerarenjargon: een 7-.

► Uit: Kleio 4, juli 2011, pag. 60

vmbo-leerlingen. Niet elke vraag is echter te presenteren in Jip-en-Janneke-taal. Bronnen kunnen wel vereenvoudigd weergegeven worden, maar er zijn grenzen aan wat historisch verantwoord is. Een docent vatte dit taalprobleem kernachtig samen met de woorden: 'Leerlingen lezen gewoon slecht.'

Examenonderwerpen

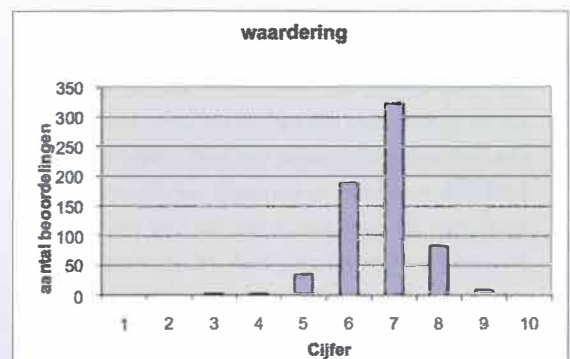
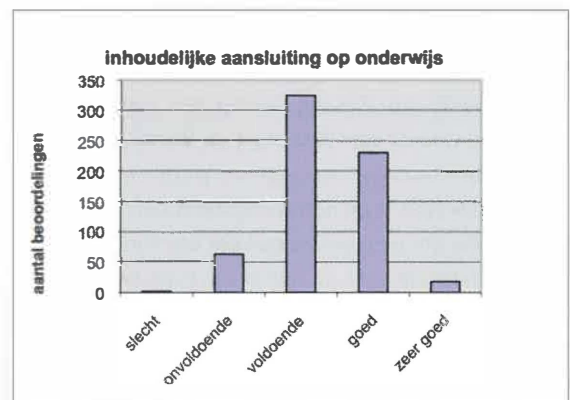
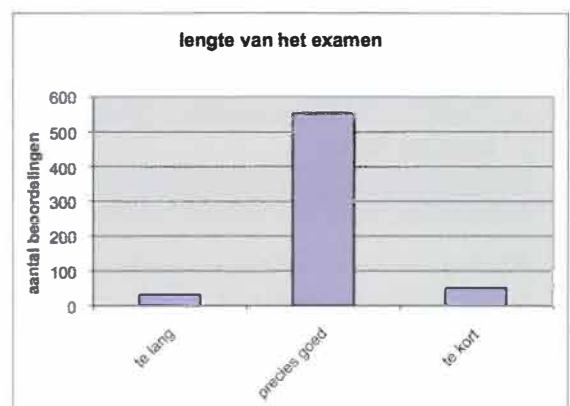
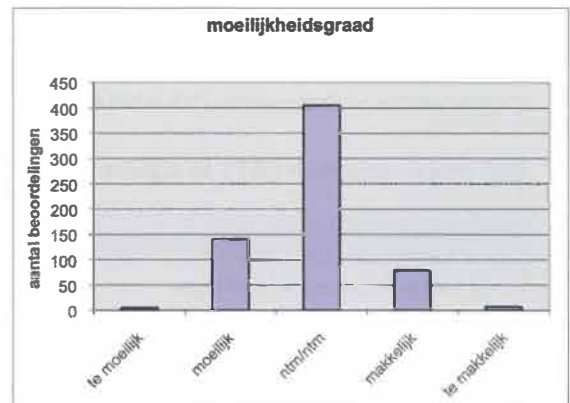
De examenthema's 'Staatsinrichting' en 'Historisch Overzicht vanaf 1900' blijken bij zowel docenten als leerlingen goed aan te slaan. De invulling van de onderwerpen is echter, zo blijkt uit enkele jaren ervaring, aan enige revisie toe. Tijdens de vmbo-studiedag op 5 november 2010 in Utrecht bleek 90 procent van de 130 aanwezigen de huidige onderwerpen relevant en betekenisvol voor vmbo-leerlingen te vinden. Tijdens dezelfde bijeenkomst bleek daarentegen bij een peiling over de tien tijdvakken met bijbehorende kenmerkende aspecten ('De Rooy') dat een nóg hoger percentage van de aanwezigen géén voorstander is van centrale examinering van de tien tijdvakken in het vmbo.³

Waardering examen

Het positieve oordeel over het examen blijkt eveneens uit de *quick-scan* na afloop van het examen (zie kader).⁴ De moeilijkheidsgraad, de lengte van het examen, de aansluiting op het gegeven onderwijs en de waardering van het examen zijn over de hele linie beoordeeld als voldoende tot goed. Daarmee zijn we op de goede weg, maar zijn we er nog niet, want ook dit examen kende enkele onvolkomenheden.

De problemen in het examen zaten dit jaar niet bij het onderwerp 'Staatsinrichting'. Op deze vragen was over het algemeen weinig kritiek. Ze waren aan de makkelijke kant met een gemiddelde p-waarde van 64,8 (terwijl het Historisch Overzicht een gemiddelde p-waarde had van 61,8). Een docent merkte op: 'Hoewel het steeds lastiger lijkt te worden om nieuwe vragen te bedenken voor dit onderdeel is het dit jaar goed gelukt om de stof te verdelen over de vragen en deze in afwisselende vraagvormen aan te bieden.'⁵ Het probleem zat vooral bij het Historisch Overzicht na 1900. Er was sprake van een cumulatie van moeilijke vragen en problemen aan het eind. Dat was een gevolg van drie keuzes in de opbouw van het examen. Ten eerste de plaatsing van overall-vragen/chronologisch

overzicht vragen aan het eind van het examen. Ten tweede enkele moeilijke verrijkingsdelen die aan het eind bevraagd werden. Ten derde het voor leerlingen taai onderwerp Europa en de Europese samenwerking.



N 637 = 100%

De kandidaten kwamen niet in tijdnood, maar de combinatie van deze drie factoren leidde ertoe dat de spanningsboog bij veel leerlingen voor het eind van het examen al geknapt was. Dat blijkt uit de afnemende scores aan het eind van het examen. Aangezien de chronologie over het algemeen de leidraad is bij het samenstellen van het examen en deze vragen niet anders dan aan het eind van het examen opgenomen kunnen worden, lijkt het erop dat hierop meer geoefend zal moeten worden.

Enkele losse vragen onder de loep

Vraag 18 leverde nogal wat discussie op in het veld. De leerlingen moesten twee kenmerken van de Sovjet-Unie noemen waaruit bleek dat die tussen 1917 en 1941 een totalitaire staat was. Dat leverde de nodige problemen op bij de correctie omdat de vraagstelling en dus ook het correctievoorschrift erg open waren. Er stond geen opsomming van kenmerken in het correctievoorschrift (bij de vakspecifieke regels in het correctievoorschrift staat *voorbeelden* van juiste antwoorden en geen opsomming van alle denkbare juiste antwoorden). De interpretatie van de antwoorden op basis van vakdeskundigheid moest dus bij de beoordeling de doorslag geven. Het probleem was vooral: wat te doen als er kenmerken van dictatuur genoemd worden terwijl kenmerken van totalitair gevraagd worden? Mag je van een vmbo-leerling verwachten dat hij onderscheid kan maken tussen het begrip totalitair en het algemenere begrip dictatuur?

Een tweede bijzonder geval vormde vraag 21, de Jesse Owens-vraag. Het is volgens sommigen een hardnekkige historische mythe dat Hitler is weggelopen bij de medaille-uitreiking vanwege de zwarte huidkleur van Owens. Met de vraag was echter niets mis: op basis van zijn politieke ideeën wilde Hitler inderdaad liever niet bij de huldiging van Owens zijn. Dat was het meest voor de hand liggende antwoord, ook volgens de kandidaten, want het was de best gemaakte vraag van het examen met een p'-waarde van 0,88!

In vraag 28 leverde de omschrijving Indonesië of Nederlands-Indië problemen op. Opgemerkt werd dat 'Indonesië als historisch begrip historisch onjuist zou zijn'. De vraag gaat echter over het bestuur van Indonesië, het gebied dus, en over die aanduiding is geen verschil van mening. Bovendien wordt in het verrijksdeel van het examenprogramma ook consequent gesproken over 'Indonesië' en

die omschrijving wordt aangehouden in het examen.

Het correctievoorschrift van vraag 35 bleek te strikt te zijn. Alles wat te maken had met vredesonderhandelingen die uiteindelijk geleid hebben tot het Camp David-akkoord van 1978 mocht goed gerekend worden.

De vraag over Europa (vraag 39) zou multi-interpretabel zijn. Sommige omschrijvingen zouden bij meer dan één instelling kunnen horen. Dat was inderdaad een zwakte van de vraag, maar uiteindelijk konden de kandidaten door deductie wel tot het juiste antwoord komen.

BB en KB-digitaal

De geschiedenisexamens voor de basisberoepsgerichte leerweg (BB) zijn in 2011 volledig digitaal afgenomen. Er is dit jaar bij geschiedenis geen gebruik meer gemaakt van papieren examens. Voor de tweede keer werden ook digitale examens afgenomen voor KB. Een kleine honderd kandidaten deed daar voor geschiedenis aan mee. Voor KB blijft daarnaast, zo lang deze pilot duurt, de mogelijkheid bestaan om ook papieren examens af te nemen. In beide gevallen gaat het in deze leerwegen om kleine aantallen kandidaten (tussen de 200 en 300). Dat neemt niet weg dat het prettig is, dat er een mogelijkheid bestaat om in BB en KB geschiedenisexamens af te nemen.⁶ Bovendien wordt veel ervaring opgedaan bij het maken van digitale examens voor ons vak. Het is alleen jammer dat er weinig respons komt op deze examens. Dat komt niet alleen door het geringe aantal kandidaten. Ook door de flexibele afname-momenten en het gebruik van een specifiek afnameprogramma kan niet direct na afname kennis genomen worden van deze examens. Overigens zijn er wel voorbeeldexamens op te vragen bij Cito.

Tot slot

Het vak geschiedenis in het vmbo blijft in beweging. De ontwikkeling naar digitale afname van examens zet door. Er is brede steun vanuit het veld voor de huidige examenonderwerpen, maar het examenprogramma zelf lijkt aan revisie toe. En, *last but not least*, de geschiedenisdocent doet er goed aan niet alleen zijn of haar geschiedenismethode als voorbereiding op het examen te gebruiken. Het oefenen met (volop beschikbare) examens en de syllabus levert goede resultaten op. ■

► Vóór de invoering van het vmbo in 2001, met de eerste examens in 2003, in het toenmalige vbo, was die keuze niet mogelijk.