

'NA 85 JAAR KLACHTEN OVER BEURSKRACH'

De vmbo-examens geschiedenis van 2014

Op zaterdag 24 mei stond de intrigerende kop boven dit artikel prominent op teletekst. De bron bleek het LAKS te zijn. Er was door kandidaten veel geklaagd over moeilijke woorden. Ook het woord 'beurskrach' (het begin van de economische crisis in 1929) in het vmbo-examen geschiedenis en staatsinrichting leverde veel klachten op. De kop is illustratief voor de aard van de klachten. Die gingen vooral over taal en inhoud. Tijd voor een nadere analyse.

Huub Kurstjens is toetsdeskundige geschiedenis bij Cito (Huub.Kurstjens@Cito.nl). De auteur bedankt Stefan Boom, Gerard Steenhagen, Lody Smeets en Babette Veen voor hun opmerkingen.

Dit artikel bevat alleen een bespreking van het papieren examen geschiedenis GTL 2014-1. De papieren en digitale examen voor BB en KB zijn buiten beschouwing gelaten.

Er was dit jaar na afloop van het examen iets vreemds aan de hand. Docenten en leerlingen waren eerst van mening dat het goed te doen was: niet te moeilijk, niet te lang, representatieve vragen over de stof. Het aantal klachten bij het LAKS was beperkt te noemen: 766 (tegen ruim 2800 klachten vorig jaar, bij nagenoeg dezelfde hoeveelheid kandidaten). Bij het College voor Examens (CvE) was welgeteld één klacht van een docent binnengekomen. Maar tijdens het nakijken sloeg menig docent de schrik om het hart. Niet eerder was de gemiddelde score de afgelopen jaren zo laag geweest. Wat lag hieraan ten grondslag? Er is geen eenduidig antwoord te geven, de materie is complex. Hierbij een voorzichtige poging tot een analyse langs de lijn van taal, vraagstelling en inhoud.

SLORDIG EN SNEL LEZEN

Ieder jaar duikt taal op als een complicerende factor bij het examen. De taalproblematiek is divers van aard. Dat leerlingen slordig en (te) snel lezen is algemeen bekend, niet alleen bij geschiedenis. De lengte van de zinnen en de lengte van het examen lijken niet het grootste probleem te zijn. Het begrijpend lezen, het (historisch) redeneren en de woordenschat van leerlingen daarentegen des te meer. Zo snapten leerlingen het woord 'strijdig' niet in de betekenis van 'in tegenspraak met' (vraag 23). Ook het in eigen woorden opschrijven van een ant-

woord wordt door kandidaten als moeilijk ervaren. Ze willen graag meer meerkeuzevragen, blijkt uit reacties op internet. Afgaande op de resultaten van dit examen lijkt dit laatste echter geen goed idee. Van de open vragen, meerkeuzevragen en voorgestructureerde vragen werden juist de meerkeuzevragen dit jaar het slechtst gemaakt. Een docent op het forum merkte op: 'Begrijpend lezen en veel tekst bij de vragen blijft een groot probleem.' En ook: 'Begrijpend lezen, enorm veel toepassing² en inzicht.¹ Dat hebben deze leerlingen gewoon NIET en zeker niet zoveel.' En: 'Heb het idee dat ik lees- en schrijfvaardigheid aan het beoordelen ben in plaats van historische kennis of vaardigheden. Als je niet goed bent in begrijpend lezen ben je bij voorbaat kansloos.' 'Een erg talig examen! Leerlingen in Rotterdam-West hebben hier erg veel moeite mee.' Een andere docent merkte op: 'Waar ik bang voor was, begint waarheid te worden en dat is dat leerlingen onnodig fouten maken doordat ze de vraag niet goed lezen en daarbij te makkelijk antwoord geven.'² Blijkbaar wordt taal als een steeds groter probleem gezien in het geschiedenisexamen en wordt de vraagstelling soms als lastig ervaren.

VRAAGSTELLING

Wanneer is een examen een goed examen en wanneer is een vraag een goede vraag? Vooraf wordt nagedacht over een spreiding van de vragen over de hele stof, over het aantal scorepunten, over vraagvormen en vraagtypen, over de moeilijkheidsgraad, (historische) vaardigheden, bronnen, enz.³ Het geheel moet een afgewogen mix van dit alles zijn en recht doen aan de stof die omschreven is in de syllabus.⁴ Ook een vmbo-leerling zal open vragen moeten kunnen beantwoorden die toepassing en inzicht van de stof toetsen en een beroep doen op (historisch) redeneren. Dat kan geoefend worden,

¹ Gerefereerd wordt hier aan het RTTI-model, waarbij toepassing² verwijst naar toepassing in een andere context.

² Zie ook het forum van de VGN.

³ Bij vraagvormen moet gedacht worden aan open vragen, meerkeuzevragen en voorgestructureerde vragen; bij vraagtypen kan gedacht worden aan het gebruik van verschillende handelingswoorden, zoals uitleggen, verklaren, beredeneren, beargumenteren, maar ook aan ordeningsvragen, zoals in een juiste historische volgorde zetten, e.d.

⁴ Zie de syllabus van 2014, te downloaden via www.examenblad.nl (doorklikken naar vmbo-geschiedenis-syllabus 2014).

bijvoorbeeld aan de hand van eerder afgenomen examens. Dat er soms vragen zijn die ook voor de examenmakers anders uitpakken dan gedacht, blijkt uit het volgende voorbeeld:

(vraag 3, uit: GTL 2014-1)

In de Grondwet van 1848 werd het censuskiesrecht ingevoerd.

→ Noem een argument dat in 1848 gebruikt werd door de voorstanders van het censuskiesrecht.

De vraag stuurt niet genoeg in de richting van het door ons gewenste antwoord. Want het jaar 1848 kan op twee manieren opgevat worden, zowel positief (mensen krijgen meer politieke invloed **na** 1848!) als negatief (mensen hadden te weinig politieke invloed **vóór** 1848!). En als de nadruk op censuskiesrecht gelegd wordt, kan het ook tweeledig uitgelegd worden (er *is* in ieder geval een vorm van kiesrecht, maar ook: het is voorbehouden aan de rijken en de mensen met een opleiding). Dat is waarschijnlijk ook wat een docent bedoelde op het forum met de opmerking: 'Het valt mij op dat veel leerlingen de vragen net iets anders interpreteren.' Het correctievoorschrift voorziet wel in meerdere juiste antwoorden, maar de sturing van de vraag is niet eenduidig. Het was beter geweest óf naar een argument te vragen van mensen die vóór het censuskiesrecht zijn (zonder jaartal) óf naar een reden te vragen waaruit blijkt dat de invoering van het censuskiesrecht in 1848 de invloed van burgers op de politieke besluitvorming vergroot heeft of de macht van de koning beperkt heeft.

IRONIE HERKENNEN

Een jaarlijks terugkerend probleem is het toetsen van kennis en vaardigheden door middel van spotprenten. Vaak bevat een spotprent een bepaalde ironie die door kandidaten niet herkend wordt of anders geïnterpreteerd wordt dan de tekenaar oorspronkelijk bedoelde. Bovendien hebben kandidaten vaak moeite een spotprent te plaatsen in de tijd waarin deze gemaakt is. Vragenmakers proberen dit probleem zoveel mogelijk te tackelen door het jaartal te vermelden wanneer de prent gemaakt is of (extra) informatie bij de prent te geven en/of een vraag voor te structureren met 'Doe het zo'. Dat dit niet altijd lukt, blijkt uit vraag 6 uit GTL 2014-1 (zie kader op de volgende bladzijde).

In de tekening komen drie personen voor: een bisschop, een katholieke arbeider en een fabrikant. Wie wordt afgebeeld als slachtoffer en wie als boeman, is duidelijk. Tot zover begrijpen de kandidaten de prent. Dan komt de tekenaar in beeld. ►

DE GEGEVENS VAN DE TOETS- EN ITEMANALYSE 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 CvE.

HET VMBO-EXAMEN 2014 IN CIJFERS (ALLEEN DE PAPIEREN EXAMENS):

	GL/TL	KB
gemiddelde score / schaallengte	34,63 / 64	27,64 / 54
gemiddelde p' waarde*	54,1	51,2
vastgestelde N-term*	1,5	1,4
% onvoldoende (na vaststelling door CvE)	20,50%	27,30%
gemiddelde cijfer	6,4	6
betrouwbaarheid*	0,74	0,78
standaarddeviatie*	7,48	7,46
aantal kandidaten	20.011	23
steekproefgrootte	17.395	22

NB De gegevens voor KB zijn veel minder betrouwbaar dan voor GL/TL vanwege het geringe aantal kandidaten dat een papieren examen doet en als gevolg van een veel kleinere steekproef. Het grootste deel van de kandidaten geschiedenis in KB doet mee aan de digitale examens.

* 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.

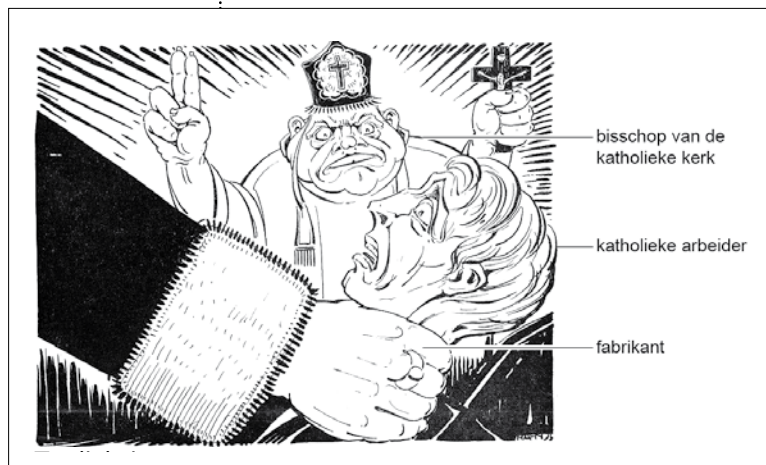
* De **betrouwbaarheid** is de mate waarin de scores consistent, nauwkeurig en reproduceerbaar zijn, kortom vrij van meetfouten.

* 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.

VRAAG 6 UIT GTL 2014-1

BRON 1

Een politieke tekening (rond 1900), met als titel:
De bisschop van de katholieke kerk:
‘Schreeuw niet zo hard!
De mensen zouden denken dat je ontevreden was.’



5 Kritiek in een prent hoeft niet te betekenen dat dit ook de kritiek van de tekenaar zelf is (hij/zij kan de prent bijvoorbeeld in opdracht gemaakt hebben van iemand anders), vandaar de formulering dat de tekenaar de kritiek laat zien, of: dat in de tekening kritiek te zien is.

6 Zie in de syllabus GS/K/3: de leervaardigheden in het vak geschiedenis en staatsinrichting.

7 Zie: <http://tiny.cc/werkvormen>. Bjorn Wansink en Hanneke Tuithof geven hier tien voorbeelden van activerende werkvormen. Ook het associëren van woorden door gebruik te maken van mindmappen is een mooie toepassing van een werkvorm hiervoor.

8 De RIR-waarde heeft betrekking op de kwestie in hoeverre een vraag discrimineert t.o.v. de hele toets. Een RIR lager dan 16 discrimineert te weinig, een RIR hoger dan 25 discrimineert erg goed.

Toelichting

Een bisschop is een leider van de katholieke kerk.

Gebruik bron 1.

2p 6 De tekenaar laat kritiek zien op de houding van de katholieke kerk rond 1900.

→ Leg uit, met behulp van een onderdeel uit de bron, welke kritiek de tekenaar laat zien op de houding van de kerk.

Doe het zo:

De kritiek is: ... (noem kritiek).

Dat blijkt uit: ... (noem onderdeel van de bron).

6 MAXIMUMSCORE 2

Voorbeeld van een juist antwoord is:

- De kritiek is: de (bisschop van de) katholieke kerk werkt samen met de fabrikanten tegen de arbeiders / negeert de harde aanpak van de fabrikant / gebruikt (zijn) haar invloed om de arbeider rustig te houden (de arbeider moet luisteren) 1

- Dat blijkt uit: de bestraffende houding/het niet ingrijpen van de (bisschop van de) katholieke kerk (ten aanzien van de fabrikant/ten gunste van de arbeider) / de neerbuigende/snuwende uitspraak ‘Schreeuw niet zo hard! De mensen zouden denken dat je ontevreden was’ 1

Opmerking: Het tweede scorepunt mag alleen toegekend worden als het eerste deel van het antwoord juist is.

Hij laat in zijn tekening kritiek zien op de houding van de kerk, in de persoon van de bisschop, die de kant kiest van de fabrikant en niet van de arbeider.⁵

Hier zit het eerste probleem: de gelaagdheid van de tekening. Kandidaten gaan in hun antwoord in op de onderdrukking van de arbeider en niet op de bisschop (en de directeur) die hem onderdrukt. De vraag kent een ‘Doe het zo’ om de leerling te helpen, maar stuurt desondanks nog te weinig.

In plaats van ‘De kritiek is ...’ was het vanwege de sturing van de vraag beter geweest de kritiek te herhalen en te specificeren. Dus: ‘De kritiek op de houding van de kerk is ...’. Voor sommigen blijft de ironie van de tekenaar echter een brug te ver.

Vervolgens was er ook nog kritiek op het correctievoorschrift. Dat zou te beperkt zijn en oneerlijk.

Te beperkt omdat meerdere antwoorden juist zijn. Daarom begint het antwoord ook met de zin ‘voorbeeld van een juist antwoord’, waardoor ook andere antwoorden, mits historisch correct, goed kunnen zijn. Het oneerlijke aspect van de puntentoekenning zat volgens enkele critici vooral in de opmerking dat het tweede scorepunt alleen toegekend mag worden als het eerste deel van het antwoord juist is. Maar hoe kan een onderdeel van de bron ter illustratie van het noemen van de kritiek goed gerekend worden als de kritiek zelf niet juist is? Dan zou een onderdeel van de bron goed gerekend moeten worden, zonder dat er enig inhoudelijk verband met de kritiek gelegd is. Wat

‘HET LIJKT EROP DAT DIT EXAMEN VOOR VEEL KANDIDATEN MOEILIK WAS. DAT WIL NIET ZEGGEN DAT HET EEN SLECHT EXAMEN WAS’

helaas wel voorkwam, is dat kandidaten vaak in het tweede deel van het antwoord lieten blijken dat ze de prent begrepen hadden door dáár pas de kritiek te verwoorden en niet, zoals de bedoeling was, bij het eerste deel van het antwoord. Het hulpmiddel ‘Doe het zo’ blijkt nu bij enkele kandidaten in hun nadeel uit te pakken. Dat is inderdaad vervelend. Maar het weglaten ervan zou weer problemen opleveren bij de meeste andere kandidaten die wél behoefte hebben aan een voorstructurering van het antwoord.

BEGRIPPEN IN JUISTE SAMENHANG

In ieder examen zit minstens één vraag waarbij kandidaten aan de hand van enkele begrippen in een paar zinnen moeten kunnen uitleggen wat die met elkaar te maken hebben. Daarbij zijn meerdere antwoorden mogelijk en dus lastig te corrigeren. Een voorbeeld hiervan is vraag 39 (zie kader).

Een verband leggen tussen de begrippen ‘vrijheid van meningsuiting’ en ‘communistische partij’ lukt nog wel. Maar met het begrip ‘nationalisme’ erbij wordt het leggen van een verband voor kandidaten wel moeilijk. Toch is een dergelijke open vraag, waarbij in eigen woorden iets moet worden opgeschreven en waarbij verbanden moeten worden gelegd, een vaardigheid die ook vmbo-leerlingen moeten beheersen.⁶ Maar misschien zijn drie abstracte begrippen en het ontbreken van een concrete gebeurtenis waar die begrippen aan gerelateerd moeten worden net iets te veel. Dat zou betekenen dat het met een abstract begrip minder en/of met behulp van een concrete gebeurtenis misschien wél zou kunnen lukken. En dan met de aanvulling dat het in het antwoord om enkele zinnen of een kort verhaaltje moet gaan (dus geen losse woorden of losse definities).

Er zijn genoeg mogelijkheden om dit soort vraagtypes met kandidaten te oefenen. Niet alleen door vragen uit oude examens te gebruiken, maar ook op een meer speelse manier. Tijdens de jaarlijkse vmbo-studiedagen in Utrecht wordt hier regelmatig in workshops mee geoefend.⁷

Overigens geldt voor alle drie de genoemde vraagvoorbeelden dat het correctievoorschrift voldoende ruimte biedt voor meerdere juiste antwoorden.

Alle drie de vragen geven bovendien een duidelijk verschil aan tussen de meer en minder vaardige kandidaten en discrimineren dus goed (RIR-waardes van respectievelijk 23, 21, en 36).⁸

MOEILIJK?

Vóór de definitieve vaststelling worden examens gescreend op vakinhoud, Nederlands, toetstechniek, e.d. Daarbij wordt uiterste zorgvuldigheid betracht. Maar garantie dat daarmee een vraag goed is, kan helaas niet gegeven worden. Het lijkt erop dat dit examen voor veel kandidaten moeilijk was, vooral als gevolg van taalaspecten, de vele denkstappen, het schakelen, historische vaardigheden of een combinatie daarvan. De N-term van 1,5, vergelijkbaar met die van 2010, is hiervan een duidelijk gevolg. Toch wil dat niet zeggen dat het een slecht examen was. Ook docenten vonden het goed te doen, maar wel aan de moeilijke kant. Veel kandidaten hebben relatief makkelijke vragen slecht gemaakt. Ook dat speelt een rol. Er is dan ook geen reden om het volgend jaar een heel ander soort examen te verwachten. Maar over enkele jaren zullen er wel wat veranderingen plaatsvinden met betrekking tot de inhoud en de vorm van de stofomschrijving in de syllabus.

Na de zomervakantie krijgt een commissie in opdracht van het College van Examens (CvE) de taak

om de bestaande syllabus door te lichten, waarbij ‘Staatsinrichting’ en ‘Historisch Overzicht vanaf 1900’ anders worden geordend en hier en daar ook anders worden ingevuld. Inhoudelijke overlap tussen de verschillende onderwerpen, zoals die er nu op onderdelen is, moet zoveel mogelijk voorkomen worden. Het verrijkingsdeel zal ook onder de loep genomen worden. De totale omvang van de examenstof blijft wel min of meer gelijk. De bedoeling is dat er over enkele jaren examens zullen komen op basis van een aangepaste syllabus. Tot die tijd zal de huidige syllabus uitgangspunt blijven voor de examenproductie. Blijf op de hoogte en houd *Kleio* en andere relevante informatiebronnen, zoals de websites van *Examenblad.nl*, het CvE en de nascholingsconferenties, in de gaten. ■

Werkloze mannen staan in de rij voor een gaarkeuken in Chicago. De eigenaar is Al Capone. 1931, National Archives and Records Administration.



VRAAG 39 UIT GTL 2014-1

2p **39** Een bewering:

De glasnost van Gorbatsjov heeft bijgedragen aan het uiteenvallen van de Sovjet-Unie.

→ Geef een historische verklaring waaruit blijkt dat de bewering juist is.

Gebruik daarbij de volgende drie begrippen in een juiste samenhang:

vrijheid van meningsuiting, communistische partij en nationalisme.

39 MAXIMUMSCORE 2

Voorbeeld van een juist antwoord is:

Door de glasnost kwam er (meer) **vrijheid van meningsuiting**.

Dat betekende verlies van macht voor de **communistische partij**.

Het **nationalisme** kreeg daardoor (in verschillende deelrepublieken) steeds meer ruimte. Vervolgens viel de Sovjet-Unie uiteen.

Indien drie begrippen juist worden gebruikt 2

Indien twee begrippen juist worden gebruikt 1

Opmerking: De begrippen moeten in een juiste onderlinge samenhang en betekenis worden gebruikt.