

Inpakken en uitpakken

De havo- en vwo-geschiedenisexamens van 2011

De examens stonden dit jaar in het teken van incidenten. Het inpakken bij de drukker ging fout waardoor een verkeerde bijlage uitgedeeld werd. Bij het uitpakken had dit gesignaleerd moeten worden, maar bij één school ging het mis.

Door de nieuwe 'sociale media' werd onder scholieren snel bekend dat de geheimhouding van het havo-geschiedenisexamen in het geding was, omdat het incident plaatsvond op de vrijdag voor de afnamedatum. Dit ongelukje leidde tot een erratum waarmee de eerste twee vragen van het havo-examen ongeldig werden verklaard; de kandidaten kregen op voorhand alle punten voor deze vragen. Jammer, vooral voor die kandidaten die de juiste antwoorden wisten en daarop hadden kunnen scoren.

De productie van een examen is een langdurig en intensief traject, dat begint bij het examenprogramma en de stofomschrijving. De leden van een constructiegroep (cg), zelf ervaren geschiedenisleraren, maken conceptvragen, die in de cg diverse malen worden besproken en telkens herzien. De toetsdeskundige van Cito stelt hieruit de verschillende tijdvakken samen, die enkele malen worden besproken door de vaksectie Geschiedenis van het College voor Examens (CvE). Voor de vaststelling van het examen door de vaksectie wordt het gescreend door hoogleraren geschiedenis en door een neerlandicus. Na de vaststelling vinden leesrondes plaats op fouten, de drukproefcontrole en ten slotte het drukken, inpakken en verzenden naar de scholen. Daar moet het juiste examen op het juiste moment bij de juiste kandidaten terechtkomen. In deze keten van mensenwerk speelt Murphy een rol: wat mis kan gaan, gaat ooit eens mis.¹

De ontvangst

In de pers werd dit jaar aan deze zaken meer aandacht besteed dan aan de inhoud van het geschiedenisexamen. De kranten kiezen elk jaar voor een andere aanpak van de examenrecensies; dit jaar was er betrekkelijk weinig

aandacht voor geschiedenis.

Het aantal klachten bij het LAKS was relatief gering. De trend van vorig jaar zet zich door: er wordt door de kandidaten minder geklaagd. Vorig jaar vermoedden wij een verband met het examenrooster waarin geschiedenis vlak voor het lange Hemelvaartsweekend was geplaatst, dit jaar kon dat geen verklaring zijn. Wellicht speelt de bekendheid van de thema's een rol. Het onderwerp 'Dynamiek en stagnatie' werd voor de derde keer geëxamineerd, 'Vietnam' zelfs voor de vierde keer, want dit thema was hetzelfde als in 2005, 2006 en 2010. Er was dus voor de kandidaten meer oefenmateriaal in de vorm van oude examens beschikbaar dan bij een nieuw thema. Dat zorgde ervoor dat de kandidaten zich beter konden voorbereiden; zij konden zich beter voorstellen wat er van hen werd verwacht op het examen. Voor hun docenten geldt uiteraard hetzelfde: ook zij waren beter in staat aan te geven welke richting het examen op zou gaan en hadden de thema's beter in de vingers.

Stefan Boom is toetsdeskundige geschiedenis bij het Cito. Met dank aan Huub Kurstjens en Niek Nelissen.

► Er bestaat ook een wet van Murphy (met opzet fout gespeeld): bij drukproefcontroles zorgen verbeteringen voor het ontstaan van nieuwe fouten. In het examen van 2010 heeft dit ons parten gespeeld.

Veel kritiek was er op de vragen bij het schilderij uit de Leidse Lakenhal dat zowel in het havo- als in het vwo-examen voorkwam.





In het vwo-examen was veel waardering voor vraag 22 (de prenten van Op-landt en Behrendt).

Voor de examenmakers levert het langdurig bevragen van een beperkt thema echter problemen op. Het wordt steeds moeilijker onbekend, geschikt bronmateriaal op te graven dat past bij de stofomschrijving, zonder in details te vervallen. Wij streven naar een examen dat geen herhaling is van eerdere examens, maar sommige paragrafen van de stofomschrijving zijn zo beperkt, dat herhaling bijna niet te vermijden is.

Het VGN-forum begint zijn waarde te bewijzen. Op het forum wordt door een beperkt aantal docenten gediscussieerd over varianten in de antwoorden, maar hun debat wordt door een groot aantal collega's gevolgd. Meer en meer wordt de vakspecifieke regel uit het correctievoorschrift (een antwoordvariant is alleen juist indien die door vakliteratuur wordt gesteund) daarbij toegepast. En zo hoort het ook.

Ook de centrale voorbespreking bleef zakelijk. Zowel bij de havo- als vwo-bespreking werden opmerkingen gemaakt over de streepjes in de vragen. De streepjes hebben als doel meer nadruk te leggen op de instructie, die de kandidaat sneller ziet en hopelijk beter leest. Bijvoorbeeld in vraag 3 uit het havo-examen:

De Hollandse landbouw was een belangrijke factor in de economische groei van de Republiek tot ongeveer 1670. Tussen 1670 en ongeveer 1740 droeg deze landbouw echter nauwelijks meer bij aan de economische groei van de Republiek.

4p 1 Geef hiervoor een verklaring door:

- uit te leggen waardoor de Hollandse landbouw tot 1670 bijdroeg aan de economische groei van de Republiek en
- twee problemen voor de Hollandse landbouw tussen ongeveer 1670 en 1740 te noemen, die aan deze bijdrage een einde maken.

Gelezen moet dan worden: "Geef hiervoor een verklaring door uit te leggen waardoor" en daarna "geef hiervoor een verklaring door twee problemen ... te noemen". Als hier een a-vraag en een b-vraag van zou worden gemaakt (zoals werd gesuggereerd), zou het te veel gaan lijken op een meerkeuzevraag. In de lay-outregels van Cito wordt daarom voor streepjes gekozen als een vraag met meerdere onderdelen geen meerkeuzevraag is.

Bij het eerste streepje wordt een uitleg gevraagd. Daarvoor worden twee scorepunten toegekend. In het antwoordmodel worden die twee punten niet onderverdeeld in de twee stappen die moeten worden gezet. Bij deze vraag luidde de eerste antwoordmogelijkheid: 'De Hollandse landbouw was al vroeg (eerste

'Het wordt steeds moeilijker onbekend, geschikt bronmateriaal op te graven, zonder in details te vervallen'

stap) marktgericht/commercieel / gericht op de export, (tweede stap) waardoor er veel geld mee verdiend werd.' Stap (1) en (2) zijn niet onderverdeeld, omdat met name het tweede antwoorddeel zonder het eerste geen volledige uitleg geeft. De kandidaten zullen veelal de tweede stap niet hebben gezet. Dat blijkt ook uit de gegevens bij deze vraag, 28 procent van de kandidaten haalt 3 punten, 19 procent alle vier.

Dit jaar is bij geschiedenis ook de 'Quickscan' uitgevoerd. Uit deze digitale enquête blijkt dat de docenten redelijk tevreden zijn over het examen; er werd een gemiddeld cijfer van 6,3 gegeven door de vwo-docenten en een 6,4 door de havo-docenten. Van de vwo-docenten geeft 80 procent het geschiedenisexamen een zes of hoger, bij de havo-docenten is dat 85 procent.

Uit verder gemaakte opmerkingen blijkt dat veel vwo-docenten zeer veel moeite hadden met de lage N-term. De normering is een gecompliceerde operatie die door het dagelijks bestuur van de CvE op basis van de toets- en itemanalyse van Cito (waarvoor Wolf de gegevens levert) wordt uitgevoerd. Voor toelichting daarop kan de VGN-site worden geraadpleegd, waar een link is opgenomen waarin veel vragen over het geschiedenisexamen worden besproken.² De gegevens waarop de normering zijn gebaseerd (van alle examenvakken) zijn te vinden in het examenverslag op www.cito.nl.

Het pilotexamen (op de examensite van Cito te vinden) verliep voor zowel havo als vwo zonder de incidenten die zich bij het reguliere examen hebben voorgedaan. Wij raden docenten aan deze examens in te zien en zo mogelijk in hun klassen uit te proberen. Dat is een goede voorbereiding op de invoering van oriëntatiekennisvragen in het nieuwe eindexamen.

Bijzondere vragen

De makkelijkste vraag in het havo-examen was vraag 4, met een p-waarde van 81.³ Blijkbaar konden de kandidaten goed uit de voeten met het passen van begrippen bij de levensloop van Jan Poppen. De moeilijkste havo-vraag zat in Vietnam. Vraag 15 had een p-waarde van 25, blijkbaar hadden de kandidaten het begrip guerrilla niet meer paraat. De havisten hebben 'Vietnam' iets slechter gemaakt ($p=58$) dan 'Dynamiek en stagnatie' ($p=62$).

In het vwo-examen was vraag twee ($p=91$) het makkelijkst. De vwo-kandidaten waren goed op de hoogte van het begrip malthusianse spanning. Het moeilijkst bleek vraag 20 ($p=48$), de volgordevraag uit het Vietnam-deel. Overigens hadden de docenten daar weinig moeite mee. Veel kritiek was er op de vragen bij het schilderij uit de Leidse Lakenhal dat zowel in het havo- als het vwo-examen voorkwam.

In het vwo-examen was de waardering voor vraag 22 (de prenten van Opland en

Behrendt) en bij de havo vonden de regio-voorzitters de matchingvraag (21) prima. Meer gesloten vraagvormen blijken te worden gewaardeerd. Dat hangt samen met de correctiedruk die onverminderd hoog blijft bij geschiedenis. Veel docenten pleiten dan ook

'Uit verder gemaakte opmerkingen blijkt dat vwo-docenten zeer veel moeite hadden met de lage N-term'

voor het plaatsen van geschiedenis in het begin van de examenweken en niet in de tweede week zoals dit jaar (en volgend jaar).

Per tijdvak bevatten de havo- en vwo-geschiedenisexamens één of enkele vragen die identiek zijn, de overlapvragen. Examenmakers kiezen hiervoor om meer inzicht te krijgen in het niveauverschil tussen havo en vwo. De psychometrische gegevens van de overlapvragen leveren een grote hoeveelheid data op, die op langere termijn onder andere meer zicht kunnen geven op wat havisten en vwo'ers moeilijker dan wel makkelijker vinden.

Hoewel de examens dit jaar goed zijn verlopen, blijven er nog veel aandachtspunten over. De aansluiting bij het taalgebruik van havisten in de vragen, de helderheid van de correctievoorschriften, de correctiedruk, de te lage moeilijkheidsgraad van de vwo-examens enzovoort, blijven onze aandacht houden. En het in- en uitpakken van de examens natuurlijk... ■

► <http://www.vgnkleio.nl/4-1-313/onderwijs/examen/veelgestelde-vragen-rond-het-centraal-examen-geschiedenis-voor-havo-en-vwo> bevat een FAQ (veel gestelde vragen) document waarin door de Commissie voor Examens uitspraken worden gedaan over veel kwesties die met de correctie te maken hebben. Een aanrader voor elke geschiedenisdocent!

► Zie www1.cito.nl/oenw/info/ttb/eind_fr.htm voor de toetstechnische begrippenlijst van Cito.

Schooltype/Leerweg	VWO	VWO	HAVO	HAVO
Soort examen	CSE	CSE	CSE	CSE
Vaknaam	geschiedenis	geschiedenis (pilot)	geschiedenis	geschiedenis (pilot)
Totaal aantal kandidaten	18.929	416	34.022	655
Steekproefgrootte	12.212	414	22.474	609
Aantal vragen	27	29	29	27
P ¹ -waarde totale steekproef	0,66	0,67	0,60	0,59
P ¹ -waarde, alleen jongens	0,68	0,67	0,61	0,60
P ¹ -waarde, alleen meisjes	0,65	0,67	0,59	0,57
Maximumscore	77	78	73	76
Gemiddelde score	51,1	52,4	43,7	44,6
Normeringsterm	0,1	0,1	0,8	0,9
Gemiddeld cijfer	6,1	6,2	6,2	6,2
Percentage onvoldoendes (<5,5)	26,2	22,5	24,9	25,3
Standaardafwijking	9,3	8,2	9,7	8,9
Betrouwbaarheid (asympt. GLB)	0,77	0,76	0,79	0,79
Standaardmeetfout	4,4	4,0	4,4	4,1

© Stichting Cito, Instituut voor Toetsontwikkeling Arnhem (2011)