

Het examen nieuwe stijl

DE STAND VAN ZAKEN

Dit schooljaar is het nieuwe examenprogramma voor geschiedenis ingevoerd. De centrale eindexamens volgen voorlopig nog de oude systematiek van de wisselende onderwerpen, maar in een pilot wordt ervaring opgedaan met een centraal examen nieuwe stijl. De eerste vier tijdvakken van een nieuwe stijl havo-examen zijn afgenomen, een vwo-pilot is gestart. Wat is op dit moment de stand van zaken? Waarmee kunnen geschiedenisdocenten nu al rekening houden?

Iedere toets, of het nu gaat om een overhoring, een mondeling of een eindexamen, moet voldoen aan kwaliteitscriteria. De belangrijkste daarvan zijn *validiteit*, *betrouwbaarheid* en *aanvaardbaarheid*. Ook een centraal eindexamen volgens de eindtermen van het nieuwe examenprogramma moet aan deze criteria voldoen. Validiteit wil zeggen: meet de toets wat het moet meten? Vaak wordt binnen het begrip validiteit onderscheid gemaakt tussen *begripsvaliditeit* en *inhoudsvaliditeit*. Bij de begripsvaliditeit gaat het erom of de begrippen die in de toets naar voren komen wel de begrippen zijn die tot de leerstof voor de toets horen. Bij een geschiedenistoets mag bijvoorbeeld een bron niet in een vreemde taal (Engels bijvoorbeeld) blijven staan. Geschiedenis kennis moet worden getoetst, geen kennis van de Engelse taal.

In het geschiedenisindexamen zoals dat tot nu toe gebruikelijk is, wordt de inhoud van de stof in een stofomschrijving vastgelegd. Op basis van deze stofomschrijvingen maken de uitgevers hun katernen en de vragenmakers hun vragen en vooral de antwoorden daarop. De inhoudelijke begrenzing van het onderwerp wordt door de stofomschrijvingen bepaald.

In het examen nieuwe stijl is er geen stofomschrijving

meer. Uitgangspunt voor de inhoudelijke begrenzing van de stof zijn de tien tijdvakken met hun 49 kenmerkende aspecten die door de Commissie Historische en Maatschappelijke Vorming zijn vastgelegd. Vanaf het komende examenjaar zal de Cevo in een jaarlijks aangepaste syllabus (zie www.cevo.nl) voor zowel de examens oude als nieuwe stijl de eisen vastleggen.

Het tweede onderdeel van validiteit dat wordt onderscheiden is de inhoudsvaliditeit, de spreiding van de toetsvragen over de stof die geleerd is. Bij een toets over vijf paragrafen van een hoofdstuk uit het leerboek, moeten de vragen evenredig over die vijf paragrafen verdeeld zijn. Als alle vragen uit één paragraaf komen, is de toets niet inhoudsvalide. Bij het examen tot nu toe wordt de spreiding bepaald door de stofomschrijving. Een paragraaf uit de stofomschrijving van een bepaalde lengte (dus deel van de totale stof) komt voor eenzelfde deel in het examen terug. In het examen van dit jaar bijvoorbeeld zal aan de Eerste Wereldoorlog ongeveer 16 punten (van de circa 78) worden besteed. Zonder stofomschrijving moet er een andere oplossing voor de spreiding komen. In theorie zou ieder van de tien tijdvakken evenveel punten kunnen krijgen. Maar dat zou betekenen dat de prehistorie met drie

Stefan Boom is werkzaam bij het Cito als toetsdeskundige geschiedenis en staatsinrichting, afdeling vo-havo/vwo.

kenmerkende aspecten even zwaar mee zou tellen als de eerste helft van de twintigste eeuw met acht kenmerkende aspecten. Daarom is gekozen voor spreiding op grond van het relatieve aantal kenmerkende aspecten (van de 49) die in een tijdvak voorkomen.

Gevolg daarvan is dat elk kenmerkend aspect in principe voor ongeveer anderhalve punt meetelt (uitgaande van een scoreschaal van circa 75 punten). Dat gaat bijvoorbeeld ook op voor kenmerkend aspect (ka) 40, 'het voeren van twee wereldoorlogen'. De zestien punten die in het reguliere examen aan de Eerste Wereldoorlog worden besteed, kunnen in het examen nieuwe stijl nooit meer zijn dan anderhalve punt, dat is hooguit één vraag. Zo strikt hoeft de verdeling niet te zijn. In sommige vragen kunnen meerdere kenmerkende aspecten bevroegd worden. Bovendien hoeven niet alle 49 kenmerkende aspecten in elk examen aan bod te komen. De examenmakers (Cevo en Cito) zouden ook de tien tijdvakken als uitgangspunt hebben kunnen nemen voor de indeling van het examen nieuwe stijl. Maar de verdeling per tijdvak zou een te strak keurslijf vormen.

Daarom is gekozen voor spreiding op grond van het relatieve aantal kenmerkende aspecten (van de 49) die in een tijdvak voorkomen.

De tien tijdvakken zijn door de examenmakers altijd beschouwd als een didactisch hulpmiddel, geen toetsdoel. Daarom wordt in het examen de meer gangbare traditionele (en internationaal geaccepteerde) indeling in prehistorie en oudheid, middeleeuwen, vroegmoderne tijd en moderne tijd toegepast. Dit wordt zichtbaar in de kopjes waarmee de onderdelen van het examen worden gescheiden. Prehistorie/oudheid en de middeleeuwen krijgen ieder ongeveer 15% van de te verdelen punten, de vroegmoderne tijd 25% en de moderne tijd de rest, 45%. Hierbij zijn vragen die 'door de tijd heen' gaan niet als aparte categorie opgenomen. Een exacte indicator vormt deze verdeling dus niet.

Betrouwbaarheid

Een goed examen is voldoende betrouwbaar. Dat wil zeggen dat de scores die behaald worden nauwkeurig, consistent en reproduceerbaar zijn, kortom vrij van meetfouten. Binnen dit begrip wordt onderscheid gemaakt tussen de objectiviteit, specificiteit en de moeilijkheidsgraad van de toets. De objectiviteit van (de

beoordeling van) een toets is afhankelijk van de helderheid en eenduidigheid van het correctievoorschrift.

Een toets met uitsluitend gesloten vragen wordt niet anders beoordeeld door twee verschillende beoordelaars. Over de beoordeling van de open vragen bij havo/vwo geschiedenisexamens is elk jaar veel debat tussen docenten en eerste en tweede correctoren. Door een meer stringent correctievoorschrift (uit het antwoord moet blijken) en door het streven elke prestatie van de kandidaat met één punt te waarderen (one bit one point) in plaats van het niet onderscheiden van verschillende stappen, wordt in de reguliere examens geprobeerd daar wat aan te doen. Maar ook bij de vastgelegde stofom-





schrijvingen geldt onverkort dat meerdere varianten van antwoorden door de aard van het vak geschiedenis mogelijk (moeten) zijn.

Bij de leg-uit-vragen van het geschiedenisexamen komt dit veel voor. Dat blijkt onder andere uit de debatten tussen correctoren op het VGN-forum (www.vgnkleio.nl) in de examentijd en de normbesprekingen van de VGN die in alle regio's worden gehouden. Daaruit wordt duidelijk dat ook de stofomschrijvingen altijd voor meerdere interpretaties vatbaar zijn, dat sommige passages verschillend uitgelegd kunnen worden, en dat, met wetenschappelijke argumenten soms meer antwoorden juist zijn dan uit de stofomschrijving valt af te leiden.

In het examen nieuwe stijl worden de juiste antwoorden niet meer afgeleid uit een stofomschrijving. De kenmerkende aspecten schrijven geen centraal vastgestelde personen, gebeurtenissen of data voor. Kandidaten moeten 'bij elk kenmerkend aspect van een tijdvak een passend voorbeeld geven van een gebeurtenis, ontwikkeling, verschijnsel of handeling dan wel gedachtegang van een persoon en dit voorbeeld gebruiken om het betreffende aspect te verduidelijken'. Deze voorbeelden zijn door de kandidaat zelf gevonden, meestal in samenspraak met zijn of haar docent, die op grond van de eigen vakkennis de gekozen voorbeelden geaccepteerd heeft. Voor het correctievoorschrift levert dat het probleem op, dat de eigen voorbeelden van de kandidaten nog meer ruimte laten voor interpreta-

tieverschillen tussen de beoordelaars. Vandaar dat in de correctievoorschriften bij het nieuwe stijl examen nog minder 'Uit het antwoord moet blijken' zal staan dan nu al het geval is. De formuleringen 'Voorbeeld van een juist antwoord is' (waarna een voorbeeld van een juist antwoord volgt) of 'Een juist antwoord bevat de volgende elementen' waarbij ook voorbeelden (tussen haakjes) worden gegeven, zullen vaker gebruikt worden. Een betrouwbare toets is voldoende specifiek. Dat wil zeggen dat de toets alleen goed gemaakt kan worden door kandidaten die de leerstof bestudeerd hebben. In de reguliere examens wordt die specificatie door de stofomschrijving gegeven. Zonder de stof bestudeerd te hebben (volgens sommigen is een weekendje van te voren al genoeg) zou er geen voldoende gehaald mogen worden.

Het examen nieuwe stijl gaat meer over algemene kennis. Oriëntatiekennis is per definitie niet specifiek. Dat is wellicht de reden dat uit steekproeven die in de pilot genomen zijn, blijkt dat vwo-leerlingen die de nieuwe stof niet bestudeerd hebben beter of net zo goed scoren op het nieuwe stijl examen als havisten die er wel les in hebben gehad. Een weekendje van te voren alle tien tijdvakken en de kenmerkende aspecten uit het hoofd leren is voor een kandidaat met een goede algemene ontwikkeling misschien nog steeds mogelijk, maar voor vrijwel alle havo-kandidaten zal dat waarschijnlijk niet genoeg zijn om een voldoende te halen voor het examen nieuwe stijl.

Moeilijkheidsgraad

Uit de toets- en itemanalyse die van elk examen met de gegevens van de versnelde correctie door Cito gemaakt wordt, is de gemiddelde p-waarde af te leiden. De p-waarde is een maat voor de moeilijkheidsgraad van het examen. Gestreefd wordt naar p-waarden tussen de 40 en 70. Daaronder (kleiner dan 40) is de toets (te) moeilijk, daarboven (boven 70) is de toets (te) gemakkelijk.

Gemiddelde p-waarden havo-examen oude en nieuwe stijl

	Regulier examen	Nieuwe stijl
2006	63,3	56,2
2007	48,1	50,5

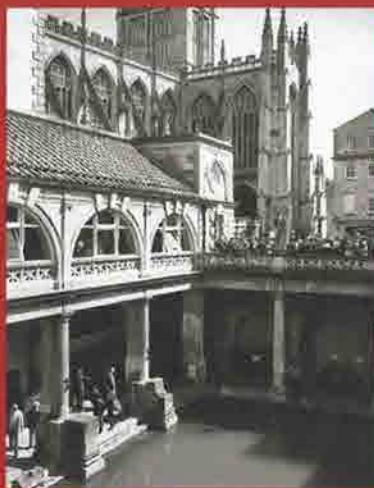
Het examen nieuwe stijl blijkt dus in 2006 wat moeilijker en in 2007 iets gemakkelijker te zijn geweest dan de reguliere examens. Maar voor een goed onderbouwde conclusie hebben we nog te weinig gegevens kunnen verzamelen. We hebben nog net voldoende inzicht in de

Voorbeeldvraag examen havo nieuwe stijl, eerste tijdvak 2007.

Twee redeneringen:

- 1 Dit badhuis wordt vaak gebruikt als bewijs dat de verovering van Engeland door de Romeinen samengaat met de verspreiding van de Romeinse cultuur onder de bevolking.
- 2 Deze foto toont aan dat er continuïteit bestaat tussen de Romeinse tijd en de middeleeuwen.

- Noem een bezwaar tegen elk van deze redeneringen en licht je antwoord telkens toe.



Een foto van een Romeins bad in Bath, een stad in Engeland. Het bad is opgegraven naast een middeleeuwse gotische kathedraal in het centrum van de stad.

4. Waarnaar richten de Grieken vanaf 800 v. Chr. op verrijkelende plekken?
- a. Geef een overzicht van de goederen (spullen) waarmee mensen de potten en
haar kolonies werd gekend.
- c. Welke gevolgen had de kolonisatie voor de Griekse cultuur?

mag door dyslectici worden overgeslagen,
n onze atleten niet af om de eerste prijs te behalen, want dat
en enkoling heraken. Nou, ons doel is veel belangrijker.
De prijs die onze atleten moeten behalen is vrijheid en
ik geluk voor de polis. En dat geluk bestaat uit rijkdom en
maar bovenal uit de vrijheid te voeren en in de veiligheid
de eigen familie te mogen leven.
Onze mannen zullen dit door hun training voor iedereen in onze polis
mogelijk maken. Andere polis zullen doelebhang van ons zijn in
oorlogstijd. Maar ook in vrede zullen onze atleten andere burgers
zijn.

moeilijkheden die de oriëntatiekennis oplevert voor ha-
visten. Hoe abstract zijn de kenmerkende aspecten? Hoe
moeilijk is het met eigen voorbeelden iets uit te leg-
gen? Welke eisen worden aan die uitleg gesteld? Welke
eisen worden aan het gekozen eigen voorbeeld gesteld?
Heeft de tweede corrector dezelfde opvatting als de
eigen docent over het passende voorbeeld en uitleg?
Er zal, nog meer dan nu het geval is, veel aanspraak
worden gemaakt op de professionaliteit en de vakken-
nis van de geschiedenisdocenten. Het bijspijkeren van
leemtes in die vakkennis, het gaat niet meer om een
beperkt onderwerp maar over de 'hele geschiedenis', zal
een langdurig proces zijn.

eens kritiek los vanuit de hoek van de specialisten. De
academische kritiek op de stofomschrijving Indonesië
van 2006/2007 is daarvan een duidelijk voorbeeld. Over-
igens werd in die kritiek ervan uitgegaan dat de meest
recente stand van zaken in de wetenschap naar voren
zou moeten komen in een stofomschrijving. Dat is nog
maar de vraag. Vaak is het academische debat daarover
nog niet voldoende uitgekristalliseerd om een duidelijke
stofomschrijving op te leveren. In veel gevallen is het
debat alleen te begrijpen als de 'traditionele' benade-
ring van het onderwerp bekend is. En dat is voor de
examenkandidaten niet het geval. Voor hen is ook zo'n
'oude' benadering nieuw.

**Het bijspijkeren van leemtes
in die vakkennis, het gaat niet
meer om een beperkt onder-
werp maar over de 'hele ge-
schiedenis', zal een langdurig
proces zijn.**

Aanvaardbaarheid

Een goede toets is redelijk aanvaardbaar. Dat wil zeg-
gen dat, in het geval van het centraal examen, ouders,
docenten, kandidaten en - voor havo en vwo - uni-
versitaire historici, de toets acceptabel vinden. Het
geschiedenisexamen nodigt elk jaar weer uit tot het
ventileren van verschillende meningen. Niet alleen rui-
men veel kranten in de examentijd ruimte in voor een
'recensie' van het examen door kandidaten, vakspecia-
listen of bekende Nederlanders, maar ook als de nieuwe
stofomschrijvingen worden gepubliceerd komt er nogal

Voorbeeldvraag examen havo nieuwe stijl, eerste tijdvak 2007.

Uit deze bron kun je afleiden dat de bevolking van Oer omstreeks 2600 v.
Chr. in een stedelijke gemeenschap leeft.

Leg uit, telkens met een gegeven uit de bron, dat er in die tijd rond Oer een
stedelijke gemeenschap bestaat

- in economisch en
- in sociaal opzicht.

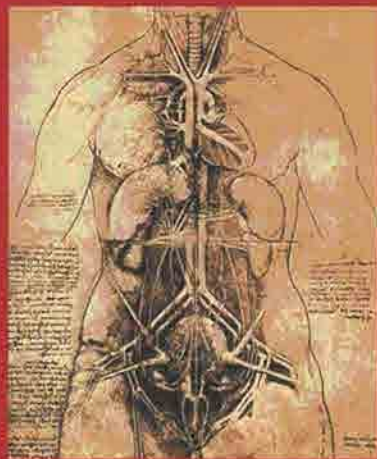


Houten tableau gevonden in een Soemerisch graf bij Oer (Uruk) in het tegenwoordige Irak
ca. 2600 v. Chr. Het bord heeft twee kanten 'oorlog' en 'vrede'. Deze kant heet 'oorlog'.

**Voorbeeldvraag examen
havo nieuwe stijl, eerste
tijdvak 2007.**

Deze prent van Leonardo da Vinci wordt vaak gebruikt om te laten zien dat de Middeleeuwen voorbij zijn. Leg dit uit, door:

- eerst met de tekening uit te leggen dat deze tekening typerend is voor een geleerde uit de Renaissance en
- daarna uit te leggen welke aanpak een middeleeuwse geleerde zou kiezen.



Anatomische tekening van een mens door Leonardo da Vinci (1452-1519), Italiaanse schilder, beeldhouwer, onderzoeker en geleerde.

Op het reguliere examen is ook algemene kritiek. De beperktheid en inhoud van de gekozen thema's, de jaarlijks andere details die moeten worden geleerd en de beperktheid van het programma werden aan de kaak gesteld. Dit leverde (mede) de vraag naar een nieuw geschiedenisprogramma op.

Het gevaar van oppervlakkigheid en gebrek aan diepgang is aanwezig.

Het examen nieuwe stijl

Het eerste examen nieuwe stijl is in een uitgebreid rapport geëvalueerd. Dat levert een genuanceerd beeld op. Er is sprake van verbetering vergeleken met het 'oude stijl' examen als het gaat om het overzicht, maar het gevaar van oppervlakkigheid en gebrek aan diepgang is aanwezig (zie www.ivgd.nl). Docenten hebben hun commentaar gegeven in een door de VGN georganiseerde enquête waaruit een voorkeur voor het toetsen van de oriëntatiekennis in het schoolexamen bleek en handhaving van wisselende onderwerpen voor het centraal examen (zie www.vgnkleio.nl). Of het nieuwe stijl examen meer of minder aanvaardbaar zal zijn dan het examen met wisselende onderwerpen is nu nog niet duidelijk. De Nijmeegse vakdidacticus Jan de Vries noemt het 'een interessante mislukking'. Zijn praktische oplossing sluit aan bij die van de VGN-docenten, zijn meer ideale oplossing neigt naar een combinatie van het toetsen van overzichtskennis en een thematische aanpak (zie zijn artikel in Kleio nr. 6, 2007).

Wat kunnen docenten en hun leerlingen nu verwachten

van het nieuwe examen? Er zijn tot nu toe vier havo-examens afgenomen, de eerste en tweede tijdvakken van 2006 en 2007. Via www.cito.nl zijn die examens op te vragen en als voorbeeld te gebruiken. In het nieuwe programma is de bestudering van oude examens zinvoller dan in het oude systeem met wisselende onderwerpen. Examenvragen zullen op grotere schaal in proefwerken op de scholen gebruikt kunnen worden.

In het nieuwe programma is de bestudering van oude examens zinvoller dan in het oude systeem.

In de examenvragen wordt soms rechtstreeks gevraagd naar het reproduceren van een kenmerkend aspect, soms wordt gevraagd een kenmerkend aspect te identificeren in een bron of een verschijnsel. De negenenveertig kenmerkende aspecten moeten dus gekend worden.

In een aantal vragen moet met een eigen voorbeeld een kenmerkend aspect uitgelegd worden. Dat sluit direct aan op domein A, de vaardigheden en benaderingswijzen die net als domein B (de oriëntatiekennis met de tien tijdvakken en de 49 kenmerkende aspecten) tot de examenstof horen. Vaardigheden worden op dezelfde wijze getoetst als in de laatste jaren op de examens gebruikelijk, vaak met bronnen, soms met een (uitgebreide) intro. Oefenen met bronnen, oefenen met vaardigheden, blijft een belangrijk onderdeel van de voorbereiding op het eindexamen.

De acht scholen die de pilot voor het havo-examen vormen, gaan verder met een pilot-project voor een vwo-examen nieuwe stijl. In 2010 zal er ook een vwo-examen nieuwe stijl worden afgenomen.

Ten slotte

Het nieuwe examenprogramma wordt bepaald door de eisen die in domein A en domein B zijn vastgelegd. Het centrale examen is daarop gebaseerd. Niet op de methodes die nu zijn verschenen, niet op een meer of minder expliciete canon. Oriëntatiekennis - de 49 kenmerkende aspecten zonder concrete namen van personen, gebeurtenissen of jaartallen - vormt de stofinhoudelijke basis. Of daarmee een begin is gemaakt met de toetsing van historisch denken en redeneren (historisch besef) en of deze aanpak van het centraal examen een goede manier is om dat te toetsen, blijft onderwerp van debat. Maar dat is nu eenmaal de aard van het vak geschiedenis. De productie van vragen voor een vwo-examen nieuwe stijl zal aan dit debat een extra impuls kunnen geven.