

## Rapport pilotproject CHMV-examen geschiedenis havo 2006

→ Instituut voor Geschiedenisdidactiek IVGD  
Instituut voor leerplanontwikkeling SLO  
Centraal Instituut voor toetsontwikkeling CITO  
Instituut voor de Lerarenopleiding ILO/UvA

Amsterdam, oktober 2006



# Inhoudsopgave

<b>Voorwoord</b>	<b>4</b>
<b>Samenvatting van conclusies en aanbevelingen</b>	<b>5</b>
<b>Hoofdstuk 1 Doel, vraagstelling en opzet</b>	<b>7</b>
1.1 Voorgeschiedenis	7
1.2 Doel van de pilotstudie	8
1.3 De probleemstelling	10
1.4 De opzet van het pilotproject	10
1.5 Begeleidingsgroep	12
1.6 Voorlichting over het project	12
<b>Hoofdstuk 2 Het onderwijs gedurende twee projectjaren</b>	<b>13</b>
2.1 Beschrijving van het onderwijs	13
2.1.1 De interpretatie van het examenprogramma	13
2.1.2 De beginsituatie van de leerlingen	15
2.1.3 De werkwijze van de verschillende scholen	16
2.1.4 Samenwerking tussen en ondersteuning van deelnemende docenten	17
2.2 Evaluatie van het onderwijs	19
2.2.1 Ervaringen van docenten	19
2.2.2 Ervaringen van leerlingen	23
2.2.3 Eisen aan didactiek en organisatie	26
<b>Hoofdstuk 3 Het examen</b>	<b>28</b>
3.1 Totstandkoming van het examen en toetstechnische analyse van het resultaat	28
3.1.1 Uitgangspunten	28
3.1.2 Validiteitsproblemen bij de constructie	30
3.1.3 De toetstechnische gegevens	31
3.1.4 Samenvatting	35
3.2 Evaluatie van het examen door docenten en leerlingen	36
3.2.1 Was het examen zoals verwacht?	36
3.2.2 Voldeed het examen aan de eindtermen CHMV?	39
3.2.3 Was het examen deugdelijk, haalbaar en eerlijk?	40
3.2.4 Correctie van het examen	41
3.2.5 Landelijke implementatie	42
3.3 Evaluatie van het examen door experts	42
3.3.1 Geschiedenisdocenten	42
3.3.2 Vakdidactici	43
3.3.3 Hoogleraren in binnen- en buitenland	43

<b>Hoofdstuk 4 Vooruitzichten</b>	<b>46</b>
4.1. Ontwikkelingen binnen de pilotscholen	46
4.1.1. De pilotdocenten en de scholen	46
4.1.2. Voornemens voor het vervolgproject	46
4.2. Te verwachten effecten bij implementatie	47
4.2.1. De vraag naar scholing	47
4.2.2. De curricula van de scholen	47
4.2.3. De relatie tot vakinhoudelijke wetenschappelijke expertise	48
<b>Hoofdstuk 5 Conclusies</b>	<b>50</b>
5.1. Beantwoording van de deelvragen	50
5.2. Conclusies	51
5.3. Aanbevelingen	54
<b>BIJLAGEN</b>	<b>57</b>
Bijlage 1: Examenprogramma's	59
Bijlage 2: Bij de projectleiding en evaluatie betrokkenen	71
Bijlage 3: Stukken gebruikt bij werving en voorlichting van projectscholen	72
Bijlage 4: De deelnemende scholen	86
Bijlage 5: Leden van de begeleidingsgroep	87
Bijlage 6: Lijst van in de tekst van het examenprogramma voorkomende begrippen	88
Bijlage 7: Vragenlijst docenten afgenomen in april 2006, voor het CE	89
Bijlage 8: Vragenlijst afgenomen direct na het CE 2006 bij leerlingen van de pilotscholen en bij leerlingen van vijf niet-pilotscholen die het reguliere examen maakten.	96
Bijlage 9: Richtvragen voor interviews met docenten van de pilotscholen na het CE	100
Bijlage 10: De tijdens periode 1 van de pilot gebruikte examenopgaven, het begeleidende bronnenboekje en het correctievoorschrift.	101

## Voorwoord

In de periode van de zomer van 2004 tot het najaar van 2006 heeft op een achttal havo-afdelingen van Nederlandse scholen een proefneming plaatsgevonden met een nieuw examen geschiedenis, met als doel te onderzoeken of landelijke implementatie hiervan mogelijk zou zijn. De proef werd in opdracht van de Minister van Onderwijs, Cultuur en Wetenschappen uitgevoerd onder leiding van een projectgroep samengesteld uit experts van IVGD (Instituut voor Geschiedenisdidactiek), SLO (Instituut voor leerplanontwikkeling) en CITO (Instituut voor toetsontwikkeling). De opdracht was in het najaar van 2006 over de resultaten van de proefneming te rapporteren. Met de aanbieding van dit rapport aan de Minister van OC&W wordt aan die opdracht voldaan.

Om tot een grondige evaluatie van de resultaten van de pilot te komen heeft de SLO besloten deze te doen uitvoeren door een tweetal onafhankelijke onderzoeksters van het Instituut van de Lerarenopleiding aan de Universiteit van Amsterdam. Hun bijdragen vormen dan ook een belangrijk onderdeel van dit rapport.

### Het rapport is als volgt opgezet:

- Een samenvatting van conclusies en aanbevelingen gaat aan de tekst vooraf. Deze worden nader toegelicht in het slothoofdstuk, hoofdstuk 5.
- Hoofdstuk 1 beschrijft de voorgeschiedenis van de pilot, de opzet van de pilot, het doel en de probleemstelling. In dit hoofdstuk wordt uiteengezet waarom is afgezien van de eerder voorgestelde casustoetsing. De probleemstelling wordt vervolgens meer in algemene zin (zonder casustoetsing) geformuleerd en uitgewerkt in deelvragen.
- Hoofdstuk 2 beschrijft het onderwijs in de pilotscholen. Het project ging weliswaar om een nieuw centraal examen, maar een proef daarmee kon niet plaatsvinden zonder een ingrijpend aangepast curriculum. Daarmee werd dat curriculum mede deel van het project. Het hoofdstuk beschrijft eerst de ontwikkelingen op de scholen, en bevat vervolgens een evaluatief deel waarin de ervaringen van docenten en leerlingen aan de orde komen.
- Hoofdstuk 3 gaat vervolgens in op de kern van de zaak: het examen. Beschreven wordt hoe het pilotexamen tot stand kwam en wat de problemen daarbij waren. Daarna volgt het resultaat van de toetstechnische analyse die is uitgevoerd na afloop van het examen. In het tweede, evaluatieve deel van dit hoofdstuk wordt ingegaan op de ervaringen van leraren en leerlingen met het examen, alsmede op het oordeel van externe deskundigen over het examen.
- Hoofdstuk 4 blikt vooruit op mogelijke ontwikkelingen na afloop van dit pilotproject. Daarbij wordt ingegaan op het vervolgproject binnen de pilotscholen en op mogelijke ontwikkelingen bij landelijke implementatie.

De verantwoordelijke auteurs voor de verschillende onderdelen zijn:

Drs. A.H.J. Witschut (IVGD) voor hoofdstuk 1, hoofdstuk 2.1 en 4.

Drs. S. Boom (Cito) voor hoofdstuk 3.1.

Dr. C.A.M. van Boxtel en dr. J.P. van Drie (ILO/UvA, in opdracht van SLO) voor hoofdstuk 2.2, 3.2. en 3.3.

De inleiding, conclusies en aanbevelingen komen voor rekening van de gezamenlijke projectgroep, die ook verantwoordelijk is voor algehele redactie van het rapport.

Amsterdam, oktober 2006.

## Samenvatting van conclusies en aanbevelingen

(Toelichtingen in hoofdstuk 5)

### Algemene conclusie:

Uit de pilot is gebleken dat een examen-havo op basis van het eindexamenprogramma CHMV mogelijk is. Landelijke invoering zal mogelijk zijn, mits goed begeleid en ondersteund.

### Detailconclusies:

- 1 Uit het onderzoek van de beginsituatie van de leerlingen is gebleken dat ze de algemene oriënterende kennis van tien tijdvakken in het najaar 2004 slechts in geringe mate beheersten. Naar de mening van de pilotdocenten is het mogelijk gebleken deze kennis in de loop van twee jaar bij een grote meerderheid van de leerlingen op een aanvaardbaar peil te brengen.
- 2 Een adequate voorbereiding van de leerlingen op het examen heeft kunnen plaatsvinden via verschillende didactische werkwijzen en invullingen van het curriculum
- 3 Over de mate waarin verdere specificatie van de tekst van het examenprogramma gewenst is, bestaat verschil van mening onder de pilotdocenten. Er is kritiek op bepaalde zeer open geformuleerde 'kenmerkende aspecten' van tijdvakken. Aan de andere kant worden vrijheid en eigen verantwoordelijkheid op prijs gesteld.
- 4 De verschillende aard van formulering en detaillering in de tekst van de 'kenmerkende aspecten' levert problemen op voor een valide toetsconstructie.
- 5 De formulering van stof in alleen 'kenmerkende aspecten' van tijdvakken geeft weinig houvast voor toetsopgaven die doorsneden door de tijd, ontwikkelingen en causale relaties beschrijven.
- 6 Pilotdocenten zijn in het algemeen van mening dat het vak geschiedenis met de nieuwe aanpak van het curriculum beter tot zijn recht komt dan met de gebruikelijke aanpak.
- 7 Leerlingen van de pilotscholen vinden net als leerlingen van niet-pilotscholen het geschiedenisonderwijs in havo-4 en havo-5 belangrijk en interessant. Wat dit laatste betreft scoren ze lager dan leerlingen van niet-pilotscholen.
- 8 Leerlingen van de pilotscholen zijn van mening dat ze met de nieuwe aanpak nuttige overzichtskennis hebben opgebouwd, die ze kunnen gebruiken en niet snel zullen vergeten. In dit opzicht scoren ze hoger dan leerlingen van niet-pilotscholen.
- 9 Leerlingen van pilotscholen vinden in mindere mate dan leerlingen van andere scholen dat ze kritisch hebben leren denken en de wereld van vandaag beter hebben leren begrijpen.
- 10 Uit toetstechnische gegevens blijkt dat de moeilijkheidsgraad van het pilotexamen hoger ligt dan die van de gebruikelijke examens. Deze moeilijkheidsgraad komt dichterbij de buurt van het streefniveau.

- 11 De toetstechnische betrouwbaarheid van het pilotexamen is ongeveer even groot als die van gebruikelijke examens.
- 12 Bij de correctie van het pilotexamen zijn geen grotere problemen ontstaan dan bij gebruikelijke examens.
- 13 De docenten van de pilotscholen zijn van mening dat het examenprogramma zoals zij dat hebben uitgetest landelijk kan worden ingevoerd.
- 14 Van de kant van experts bestaat naast tevredenheid ook kritiek op het examen. Deze kritiek richt zich met name op de geringe aandacht voor toetsopgaven die doorsneden door de tijd, ontwikkelingen en causale relaties beschrijven - of in het algemeen toetsopgaven die gaan over historisch denken en redeneren ('vaardigheden').
- 15 Over de vraag hoe een nieuw examen op vwo-niveau kan worden geïmplementeerd bestaat nog onduidelijkheid. Onzekerheid hieromtrent wordt gevoed door in conclusie 14 genoemde redenen.

#### **Aanbevelingen**

- 1 Het verdient aanbeveling te bezien welke nadere uitwerking van of toelichting op de kenmerkende aspecten van tijdvakken gewenst is. Dit met behoud van de vrijheid van docenten om eigen concrete invullingen te kiezen.
- 2 Bezien moet worden hoe met een nadere toelichting op de examenstof diachronische ontwikkelingen en causale relaties meer accent kunnen krijgen.
- 3 Onderzocht moet worden in hoeverre op de eindtermen betreffende 'historisch denken en redeneren' (4 t/m 7) een nadere toelichting gewenst is.
- 4 Bij de verdere toetsontwikkeling ten behoeve van experimentele examens verdienen de aspecten van 'betekenis voor het heden' en 'historisch denken en redeneren' speciale aandacht.
- 5 Bij de verdere toetsontwikkeling moet worden bezien hoe recht kan worden gedaan aan het niveauverschil tussen havo en vwo.
- 6 De landelijke invoering van het nieuwe examen dient ondersteund te worden door een uitvoerig programma van voorlichting en nascholing.
- 7 De tijdens de uitvoering van de pilot opgedane ervaringskennis van docenten uit pilotscholen moet bij de implementatie worden benut voor voorlichting en scholing van andere docenten.
- 8 Landelijke examens gebaseerd op domein A en B van het nieuwe examenprogramma kunnen in het havo vanaf 2009 en in het vwo vanaf 2010 plaatsvinden, mits bovenstaande aanbevelingen worden opgevolgd en tijdig oplossingen voor gesignaleerde problemen worden gevonden.

# 1 | Doel, vraagstelling en opzet

## 1.1. Voorgeschiedenis

Het pilotproject met een nieuw geschiedenisexamen is voortgevloeid uit de werkzaamheden van de Commissie Geschiedenisonderwijs ('De Wit'), die in maart 1998 het rapport *Het Verleden in de Toekomst* uitbracht, en van de Commissie Historische en Maatschappelijke Vorming (CHMV, 'De Rooij'), die in februari 2001 het advies *Verleden, Heden en Toekomst* publiceerde.

De Commissie-De Wit had een onderzoek gedaan naar de verwachtingen die de samenleving, waaronder het vervolgonderwijs, ten aanzien van het geschiedenisonderwijs koesterde, en daarbij een oordeel gegeven over de aspecten kennis en vaardigheden, thematische verdieping en chronologisch overzicht, en hun onderlinge verhouding. In haar rapport constateerde deze commissie dat het nodig was een 'nieuwe balans van kennis, inzicht en vaardigheden' tot stand te brengen. Ze constateerde 'een zekere verwijdering tussen wat in een bredere maatschappelijke kring wenselijk wordt geoordeeld, en de gangbare praktijk in het geschiedenisonderwijs.'<sup>1</sup> In de samenleving werd het wenselijk geacht - aldus de commissie - dat leerlingen ten minste beschikken over een bepaalde op de chronologie gebaseerde overzichtskennis. In het geschiedenisonderwijs was daarentegen het accent verschoven naar een op 'historische vaardigheden' gebaseerde, overwegend thematische, benadering. Het examen bestond uit min of meer los van elkaar staande thema's.

De Commissie-De Rooij kreeg de opdracht op basis van de adviezen van de Commissie-De Wit voorstellen te doen voor nieuwe kerndoelen en eindtermen voor geschiedenis in alle vormen van onderwijs, van basisonderwijs tot en met de bovenbouw van het havo en vwo. De kern van alle door de commissie voorgestelde programma's vormde een stelsel van tien tijdvakken met aansprekende namen die de ontwikkeling van het chronologisch besef van leerlingen geacht werden te ondersteunen. In haar beleidsreactie op het rapport van de Commissie-De Rooij constateerde de toenmalige staatssecretaris van Onderwijs, Cultuur en Wetenschappen drs. K.Y.I.J. Adelmund dat het evenwicht tussen kennis en vaardigheden in het voorstel van de commissie werd 'hersteld door de introductie van in alle sectoren van het funderend onderwijs (basisonderwijs, voortgezet onderwijs) terugkerende historische oriëntatiekennis in de vorm van tien met een korte karakteristiek aangeduide tijdvakken. Dit wordt in de meeste adviezen een geslaagde opzet gevonden, en deze zal het kader moeten vormen voor de verdere uitwerking in alle betrokken onderwijssectoren. Het geschiedenisonderwijs zal daardoor beter door de samenleving herkend en meer gewaardeerd worden.'<sup>2</sup>

In dezelfde beleidsreactie werd echter ook geconstateerd dat de voorstellen betreffende een nieuw eindexamen 'nadere doordenking' behoeft: 'Bij de nadere doordenking (waarbij toets- en vakdeskundigen moeten zijn betrokken) moet worden bezien hoe ruimte kan worden gegeven aan de beoogde vernieuwing, ook in de toetsing, zonder te blijven in de situatie van een vak met een centraal examen waarop steeds kritiek is.'<sup>3</sup>

In september 2002 gaf Minister van OC&W Mw. M.J.A. van der Hoeven een 'informele werkgroep' - samengesteld uit vertegenwoordigers van VGN, CITO, CEVO, SLO, het Tweede Fase Adviespunt en voormalige leden van de Commissie-De Rooij (CHMV) - de opdracht de examenproblematiek nader te onderzoeken. Aan de orde was de vraag of het examenprogramma zoals het in het rapport-CHMV voorkomt op zichzelf duidelijk genoeg was, of dat nog een nadere specificatie nodig was van de globaal aangeduide 'oriëntatiekennis' van de

<sup>1</sup> Het verleden in de toekomst. Advies van de commissie geschiedenisonderwijs, Den Haag, maart 1998. p. 35-36.

<sup>2</sup> Brief van Staatssecretaris K.Y.I.J. Adelmund aan de Tweede Kamer, d.d. 13 december 2001, kenmerk VO/BOB/01/42016, pag. 4.

<sup>3</sup> Ibidem, pag. 5.



tijdvakken. Verder werd de vraag besproken of de in het rapport-CHMV geopperde 'casustoetsing' een geschikte vorm van examinering zou opleveren.

Tijdens de werkzaamheden van de informele werkgroep werd op zeer bescheiden schaal enig onderzoek verricht, door op één school een aantal leerlingen van havo- en vwo-eindexamenklassen wat opgaven te laten maken die 'algemene oriëntatiekennis' betroffen. De leerlingen werd vervolgens naar hun ervaringen gevraagd. Deze proef was echter veel te beperkt en in zijn opzet ongeschikt om er conclusies aan te verbinden omtrent de bruikbaarheid van het examenprogramma en de haalbaarheid van bepaalde toetsvormen. De informele werkgroep concludeerde dan ook dat nader onderzoek nodig was. In januari 2003 rapporteerde de groep met de volgende aanbevelingen:

- Er blijft verschil van inzicht over de mate waarin specificatie van de oriëntatiekennis op vmbo-, havo- en vwo-niveau noodzakelijk is voor valide en betrouwbare toetsing.
- Casustoetsing is geen doel op zich. Het gaat er om te toetsen op manieren die recht doen aan de uitgangspunten van het rapport van de CHMV.
- We hebben niet voldoende kennis verzameld om principiële uitspraken te kunnen doen over de (on)mogelijkheid van het kunnen toetsen van 'historisch besef' zoals omschreven in het rapport van de CHMV.
- De pogingen die in de herfst van 2002 werden ondernomen om toetsen te ontwerpen, maken duidelijk dat er meer nodig is om tot beter gefundeerde uitspraken te komen.
- Er dient daarom een ruim opgezette en goed ondersteunde pilot te komen.
- Het veld moet zo snel mogelijk op de hoogte worden gebracht van de activiteiten van de informele werkgroep en de (eventuele) pilot.

Op 9 juli 2003 stuurde de Minister het verslag van de informele werkgroep naar de Tweede Kamer, vergezeld van een brief waarin zij meedeelde voornemens te zijn de programma's van de CHMV per 2007 in te voeren: 'In het rapport wordt voorgesteld, voor het voorgestelde centraal examen een proef te starten op een beperkt aantal scholen (havo/vwo). Ik neem dat voorstel over. Een dergelijke proef moet goed worden voorbereid en kan niet eerder starten dan per 1 augustus 2004 (ik heb het rapport in maart ontvangen). Er is echter inhoudelijk een brede overeenstemming over het voorgestelde examenprogramma. Daarom wil ik het per 1 augustus 2007 in het havo/vwo invoeren: tegelijk met de voorgestelde aanpassingen in de profielen. Ik zal na de zomer in overleg met VGN, SLO, CITO en CEVO bezien welke examenvorm dan per die datum mogelijk is.'<sup>4</sup>

Aldus werd besloten tot de uitvoering van een pilot met een examen-nieuwe stijl, waarover dit rapport verslag uitbrengt.

### 1.2. Doel van de pilotstudie

Het doel van de pilot was meer zekerheid te verschaffen over de vraag of een examen gebaseerd op de voorstellen van de CHMV mogelijk zou zijn. Om deze vraag te beantwoorden werd in eerste instantie gebruik gemaakt van de algemeen gebruikelijke criteria voor adequate centrale toetsing: validiteit, betrouwbaarheid, bruikbaarheid (efficiëntie) en aanvaardbaarheid. Met validiteit wordt bedoeld dat een toets een goede afspiegeling vormt van wat leerlingen geacht worden te leren, m.a.w. dat getoetst wordt wat de bedoeling

<sup>4</sup>Brief van Minister van OC&W M.J.A. van de Hoeven aan de Tweede Kamer, d.d. 9 juli 2003, kenmerk VO/01/03/12922.

was. Hierbij kan het bijvoorbeeld gaan om een goede afspiegeling in de opgaven van de inhoud van de examenstof. Met betrouwbaarheid wordt bedoeld dat de meting op zodanige wijze gebeurt, dat de toetsresultaten een reëel beeld geven van wat kandidaten weten en kunnen. Een toets is bruikbaar en efficiënt als het technisch mogelijk is deze binnen de gestelde tijd en omstandigheden goed uit te voeren. Een toets wordt bijvoorbeeld onbruikbaar als hulpmiddelen worden gevraagd die niet algemeen beschikbaar zijn (bijv. een computer), of als er zóveel leeswerk aan te pas komt, dat binnen de gestelde tijd het beantwoorden van vragen niet meer mogelijk is. De aanvaardbaarheid van een toets duidt op de mate waarin deze door het onderwijsveld en de samenleving als fair en redelijk herkend wordt.

Op grond van de eerste twee criteria werd al spoedig duidelijk dat een casustoetsing, zoals voorgesteld in het rapport-CHMV, niet de aangewezen vorm was voor een valide en betrouwbare toetsing. Ook met andere toetsvormen zou recht gedaan kunnen worden aan het door de CHMV voorgestelde examenprogramma. De criteria van bruikbaarheid en aanvaardbaarheid betekenden zelfs een belangrijk bezwaar tegen de casustoetsing. In een brief, gedateerd 14 maart 2005, aan de Minister van OC&W gaf de projectleiding dan ook aan te willen afzien van casustoetsing. Zij motiveerde deze keuze in die brief als volgt:

'Als deze eisen [bedoeld zijn de eisen van validiteit, betrouwbaarheid, efficiëntie en aanvaardbaarheid] nader worden omschreven en worden gelegd op het voorstel van de CHMV om in een examen van vier uur (één uur voorbereiding) op grond van een omvangrijke serie bronnen een betoog te laten schrijven, dan zien we dat:

- de inhoudsvaliditeit moeilijk te waarborgen is: als de te toetsen stof bestaat uit de 49 kenmerkende aspecten in tien tijdvakken, dan kan daaraan moeilijk in één overkoepelend thema worden voldaan.
- de begripsvaliditeit evenmin gewaarborgd kan worden: er wordt een te grote nadruk gelegd op de lees- en schrijfvaardigheid van de kandidaten.
- de betrouwbaarheid van de toets nauwelijks te garanderen is: niet alleen zou het zeer moeilijk zijn gelijkwaardige opgaven te produceren, maar de beoordeling ervan zou tot grote ongelijkheid kunnen leiden. Buitenlandse voorbeelden van casustoetsing (zoals het IB-examen, of het Engelse eindexamen) werken met een beoordelingssystematiek die achteraf wordt opgesteld: testing by outcome. Daarbij wordt al het examenwerk beoordeeld door een groep professionele 'markers' die op basis van het ingeleverde werk de beoordeling aanpassen. Dat is een breuk met het systeem van de Nederlandse centrale examens, dat uitgaat van vooraf geformuleerde correctievoorschriften die door de eigen docent worden toegepast.
- de toets waarschijnlijk niet specifiek genoeg zal zijn en onvoldoende kan differentiëren: hoewel dit afhankelijk is van de manier van beoordelen lijkt het in het Nederlandse systeem van beoordeling door de eigen docent met een vooraf vastgesteld antwoordmodel zeer moeilijk. Buitenlandse voorbeelden van casustoetsing (Engeland) werken met meer of minder juiste antwoorden - op grond van een testing by outcome-procedure. Dat is een beoordelingssysteem dat in Nederland niet wordt toegepast in de centrale examens.
- een toets van vier uur niet efficiënt is, en organisatorisch onhaalbaar binnen de huidige organisatie van de centrale examens.
- de aanvaardbaarheid van het examen met de casustoetsing niet zal toenemen. Als het in de huidige situatie al zo is dat de thema's gezien worden als een te beperkte hoeveelheid stof voor een centraal examen, dan zal één casuonderwerp (bijvoorbeeld een debat over plaatsing van een slavernijmonument) dit probleem eerder verergeren dan verbeteren.'

Met het afzien van casustoetsing werd aangesloten bij de constatering van de eerdere informele werkgroep dat casustoetsing geen doel op zichzelf was. Het ging erom een vorm van toetsing te ontwikkelen die recht

deed aan het nieuwe examenprogramma. Daarvoor leken goede mogelijkheden te bestaan zonder gebruik te maken van casustoetsing.

### 1.3. De probleemstelling

De vraag die met de pilotstudie moest worden beantwoord, kon nu in algemene zin als volgt worden geformuleerd:

'Welke vorm/vormen van examinering is/zijn passend en praktisch haalbaar, uitgaand van de examenprogramma's geschiedenis (uitgezonderd het programma voor het combinatievak geschiedenis en maatschappijleer) zoals omschreven in het rapport van de Commissie Historische en Maatschappelijke Vorming?'

Deze vraagstelling impliceerde dat de examenprogramma's zoals voorgesteld door de CHMV het uitgangspunt vormden van het onderzoek. Hiermee werd bedoeld op de eindexamenprogramma's voor havo en vwo, wat in de praktijk werd toegespitst op alleen het programma havo, omdat de pilot alleen daar plaatsvond (zie 1.4. voor de motivering van deze keuze). De vraagstelling liet de mogelijkheid open van verschillende soorten 'passende' examens. Tijdens de uitvoering van het pilotproject is het examenprogramma enkele malen aangepast in het kader van de voorbereiding van een vernieuwde tweede fase havo/vwo die met ingang van de cursus 2007-2008 zou worden ingevoerd. De laatste versie van het examenprogramma kwam in mei 2006 tot stand<sup>5</sup>. In de gewijzigde versies ging het echter voornamelijk om de inhoud van het schoolexamen en de verdeling van de stof over schoolexamen en centraal examen. Voor de pilotstudie betekende dat geen wezenlijk verschil. Bij de pilot werd uitgegaan van domein A en B (historisch besef en oriëntatiekennis) als stof voor het centraal examen. De omschrijving van deze domeinen is nauwelijks gewijzigd ten opzichte van de CHMV-voorstellen. De omschrijving van de 'oriëntatiekennis' in 49 kenmerkende aspecten van tien tijdvakken bleef geheel ongewijzigd.

De vraagstelling van het pilotproject werd uitgewerkt in de volgende deelvragen:

- 1 Kan een valide examen worden ontwikkeld dat voldoet aan de eindtermen?
- 2 Kan een examen worden ontwikkeld dat voldoet aan de eisen van deugdelijke toetsing?
- 3 Kan een examen worden ontwikkeld dat voldoet aan de verwachtingen van docenten en leerlingen?
- 4 Kan een examen worden ontwikkeld dat voldoende betrouwbaar te corrigeren is?
- 5 Kan een examen worden ontwikkeld dat in de ogen van docenten en leerlingen haalbaar en eerlijk is?
- 6 Wat zijn de verschillen tussen een examen nieuwe stijl en de reguliere examens?
- 7 Slagen scholen en docenten erin om een aanvaardbaar curriculum en een didactiek te realiseren waarmee leerlingen goed voorbereid worden op het beoogde examen?
- 8 Hoe beoordelen experts het pilotexamen?

### 1.4. De opzet van het pilotproject

De opdracht om een pilotproject uit te voeren werd gegeven aan het Instituut voor Geschiedenisdidactiek IVGD, het instituut voor leerplanontwikkeling SLO en het instituut voor toetsontwikkeling CITO. Een projectgroep van drie deskundigen uit de drie betrokken instituten kreeg gezamenlijk verantwoordelijkheid voor de leiding van het project, waarvan de algemene leiding kwam te berusten bij het IVGD. Voor de evaluatie van het

<sup>5</sup> Zie bijlage 1 voor het oorspronkelijke programma volgens de CHMV en het programma van mei 2006.

project is in een later stadium experts van de Universiteit van Amsterdam (Instituut voor de Lerarenopleiding ILO/UvA) verzocht een bijdrage te leveren.<sup>6</sup>

Om de schaal van de pilot niet te groot te laten worden, werd besloten deze vooralsnog te beperken tot een achttal scholen van het havo. Overwogen werd dat de examinering op het vmbo nog tamelijk recent gewijzigd was, zodat vernieuwing daarvan later aan de orde zou zijn dan op het havo en vwo. Een proefneming op het havo leek urgenter dan een proef op het vwo, gezien de moeilijkheidsgraad van wat van leerlingen gevraagd zou worden. Een geslaagde proef op het havo zou voldoende garanties inhouden dat ook op het vwo volgens de nieuwe opzet gewerkt zou kunnen worden. Op de projectleiding is desondanks in de loop van de pilot enige druk uitgeoefend om toch tot een uitbreiding van de pilot naar vwo te komen, maar daarvoor werd door de Minister van OC&W geen toestemming gegeven.

In de loop van de maanden april, mei en juni 2004 werd een achttal landelijk gespreide scholen geworven waarop de pilot in het najaar van 2004 zou starten. Sommige scholen hebben zich spontaan aangemeld nadat zij door de berichtgeving over de werkzaamheden van de 'informele werkgroep' over de voorgenomen pilot gehoord hadden. Andere scholen zijn door de verschillende leden van de projectleiding benaderd. De aanmelding gebeurde op basis van vrijwilligheid, waarbij zeker een rol heeft gespeeld dat de betrokken docenten er wel wat in zagen mee te werken aan een vernieuwde examenopzet. De docenten die meededen, deden dat dus vanuit een positieve grondhouding. Hiermee moet bij de interpretatie van de onderzoeksgegevens rekening gehouden worden. Bij de werving en eerste voorlichting van deelnemende scholen hebben vier documenten een rol gespeeld: 1 Het eindexamen geschiedenis - informatie voor leerlingen, 2 Informatie voor scholen, 3 Het project - voornaamste karakteristieken van het nieuwe onderwijs en 4 Karakteristiek van het nieuwe examen.<sup>7</sup>

Op de acht scholen ging het zonder uitzondering om havo-afdelingen van bredere scholengemeenschappen. Ze zijn gevestigd in Groningen, Haaksbergen, Zeist, Delft, Papendrecht, Goirle, Tilburg en Nijmegen.<sup>8</sup> In totaal twintig docenten waren bij de pilot betrokken, variërend van twee tot vier per school. Het aantal leerlingen dat uiteindelijk het eerste pilotexamen heeft afgelegd bedraagt 659. Hiertoe behoorde ook een aantal in het examenjaar 2005 gezakte kandidaten. Deze kregen de keus alsnog mee te doen met de pilot (wat voor hen een omschakeling betekende) of opnieuw een examen af te leggen volgens het 'oude regime' - wat betekende dat ze individueel zouden moeten studeren, afgezonderd van de overige kandidaten in hun havo-5-groepen. De meesten kozen ervoor mee te doen aan de pilot.

Gemikt werd op een eerste pilotexamen in het jaar 2006. Om de leerlingen de gelegenheid te geven zich hierop goed voor te bereiden, werd in de cursussen 2004-2005 en 2005-2006 respectievelijk in havo-4 en havo-5 een aangepast onderwijsprogramma aangeboden. De rapportage over het pilotexamen dat in mei 2006 zou plaatsvinden, zou in het najaar van 2006 worden opgeleverd, waarna daarop gebaseerde besluitvorming zou kunnen plaatsvinden over de wijze waarop de nieuwe examenprogramma's per 2007 zouden worden geïmplementeerd.

Invoering van de programma's in 2007 impliceert dat het eerste landelijke examen nieuwe stijl op het havo in 2009, en op het vwo in 2010 plaatsvindt. Voor de acht pilotscholen zou dat inhouden dat in de examenjaren

<sup>6</sup> Zie bijlage 2 voor de bemensing vanuit de diverse betrokken instituten.

<sup>7</sup> Deze stukken zijn in bijlage 3 opgenomen.

<sup>8</sup> Zie bijlage 4 voor de lijst van deelnemende scholen.

2007 en 2008 tijdelijk zou moeten worden teruggeschakeld naar het oude programma. De docenten van de pilotscholen vonden dat niet wenselijk. Daarom drongen zij aan op een vervolgproject, waarin ook gedurende de examenjaren 2007 en 2008 met een nieuw examen zou kunnen worden geëxperimenteerd. Dit zou het bijkomende voordeel hebben van een ruimere ervaring alvorens de nieuwe examens landelijk zouden worden ingevoerd. In haar brief van 2 juni 2005 aan de projectleiding onderschreef de Minister deze beide argumenten en stond verlenging van het experiment op de acht pilotscholen toe.

Deze rapportage heeft echter uitsluitend betrekking op de eerste fase van het project die met het examen in 2006 werd voltooid, hoewel inmiddels op de projectscholen in de cursus 2005-2006 begonnen is met de opleiding van een tweede pilotgeneratie in havo-4. Wel zal in dit rapport aandacht worden besteed aan de aandachtspunten die in het vervolgproject speciaal betrokken kunnen worden (zie hoofdstuk 4).

#### **1.5. Begeleidingsgroep**

De uitvoering van het project werd begeleid door een groep van deskundigen uit de vakwetenschap, de eerste- en tweedegraads vakdidactiek, de CEVO, het Tweede Fase Adviespunt en het bestuur van de VGN.<sup>9</sup> In deze klankbordgroep werden problemen en resultaten besproken en werd de projectleiding geadviseerd over de verdere voortgang.

#### **1.6. Voorlichting over het project**

Gedurende de looptijd van de pilot is op een aantal manieren voorlichting gegeven over de achtergronden van het nieuwe examen en het verloop van de pilot. De behoefte aan informatie was groot, zowel bij docenten in het veld als bij educatieve uitgeverijen die duidelijkheid wensten over het soort leermiddelen dat zij met het oog op de nieuwe examens van 2009 en daarna geacht werden te ontwikkelen. De moeilijkheid van de projectleiding was hierbij dat niet op de resultaten van de pilot vooruitgelopen kon worden en dat evenmin via een al te grote mate van invulling en interpretatie van de tekst van het eindexamenprogramma daarin zaken zouden worden gelegd die er strikt genomen niet in te lezen waren. De mate van duidelijkheid die kon worden gegeven was daardoor een tijd lang beperkt.

Niettemin heeft de projectleiding geprobeerd betrokkenen op de hoogte te houden door het publiceren van een aantal artikelen in het blad *Kleio* van de Vereniging van docenten Geschiedenis en staatsinrichting in Nederland (VGN) - in totaal zes artikelen die in enigerlei verband stonden met de nieuwe ontwikkelingen. Met het bestuur van de VGN bestond voortdurend contact via de begeleidingsgroep. Ook is de projectleiding twee maal op bezoek geweest in de voltallige VGN-bestuursvergadering. Aan docenten in het veld werd, behalve via de *Kleio*-artikelen voorlichting gegeven op conferenties: op de VGN-didactiekconferenties in 2005 en in 2006, op gebruikersbijeenkomsten van methoden van de uitgeverijen Wolters-Noordhoff en ThiemeMeulenhoff en op twee Tweede Fase-conferenties van de SLO. Op deze wijze zijn enige honderden docenten (naar schatting 300 à 400) met het project in aanraking gekomen. Met voor geschiedenis verantwoordelijke uitgeverij van Wolters-Noordhoff en ThiemeMeulenhoff hebben contacten plaatsgevonden op verzoek van deze uitgeverijen. Ten slotte werden op de website van het IVGD alle relevante stukken met betrekking tot het project gepubliceerd.

---

<sup>9</sup> Zie bijlage 5 voor de bemensing van deze groep.

## 2 | Het onderwijs gedurende twee projectjaren

Dit hoofdstuk valt in twee delen uiteen: een beschrijvend deel (2.1) waarin wordt uiteengezet hoe het onderwijs op de acht projectscholen verlopen is, en een evaluatief deel (2.2) waarin wordt beschreven hoe de deelnemende docenten en leerlingen achteraf op hun ervaringen met betrekking tot het gegeven onderwijs terugkijken en deze waarderen. (De ervaringen met betrekking tot het examen worden in hoofdstuk 3 besproken).

### 2.1. Beschrijving van het onderwijs

#### 2.1.1. De interpretatie van het examenprogramma

De tekst van het programma voor het centraal examen havo bestaat uit twee belangrijke onderdelen:

- Beheersing van vormen van historisch denken en redeneren (voornamelijk beschreven in domein A: Historisch besef);
- Beheersing van een kader van oriëntatiekennis (voornamelijk beschreven in domein B: Oriëntatiekennis).

Met het eerste onderdeel was in het geschiedenisonderwijs al ervaring opgedaan sinds de introductie van vernieuwde examens in 1993/1994. De toen ingevoerde 'structuurbegrippen en vaardigheden' waren praktisch gelijklopend aan de eindtermen 4 t/m 7 uit het nieuwe eindexamenprogramma. Ook in de programma's van de Tweede Fase havo/vwo van 1998/'99 werden de 'structuurbegrippen en vaardigheden' vrijwel ongewijzigd gehandhaafd. Wat dat betreft hoefde er dus weinig te veranderen. Niettemin heeft de projectleiding in de loop van het pilotproject met de deelnemende docenten nog wel enige aandacht besteed aan dit deel van het examenprogramma, waarbij nu en dan een zeker verschil van interpretatie bleek te bestaan over deze 'oude' eindtermen.

Nieuw was het kader van oriëntatiekennis. Men kan stellen dat dit in de plaats kwam van het beheersen van twee (wisselende) eindexamenthema's, waarvoor jaarlijks een uitvoerige stofomschrijving werd vastgesteld. Maar dat betekende niet dat gewoonweg de ene stofomschrijving werd vervangen door een andere. Het ging om een andere vorm van kennis. In het nieuwe examenprogramma worden bij tien tijdvakken 'kenmerkende aspecten' in algemene bewoordingen aangeduid, zonder daarbij te specificeren welke voorbeelden van gebeurtenissen, personen of concrete verschijnselen leerlingen daarbij moeten kennen - wat in stofomschrijvingen van examenthema's tot dusverre wel gebruikelijk was. Een vraag die al in de informele werkgroep werd besproken, was of het specificatieniveau van de kenmerkende aspecten in het rapport-CHMV voldoende duidelijkheid en houvast bood. Volgens de daarop betrekking hebbende eindterm (nr 8) dienden kandidaten in staat te zijn 'bij elk kenmerkend aspect van een tijdvak een passend voorbeeld te geven van een gebeurtenis, ontwikkeling, verschijnsel of handeling dan wel gedachtegang van een persoon en dit voorbeeld te gebruiken om het betreffende aspect te verduidelijken'. De vraag was welke voorbeelden dit zouden kunnen zijn en of het gewenst was dergelijke voorbeelden als een vast curriculum aan de projectscholen voor te schrijven.

Niet specificeren van namen en data had volgens de projectleiding voordelen voor het leren van de leerlingen. Zij moesten aan de hand van concrete voorbeelden zaken als feodalisme, industrialisatie of koude oorlog kunnen uitleggen. De vraag was niet langer 'Heb ik van James Watt gehoord', maar: 'Wie of wat zal ik gebruiken om de industriële revolutie uit te leggen?' Leerlingen moesten zich dus goed realiseren wat de betekenis was van de feiten die ze leerden. Docenten van de pilotscholen vonden het een voordeel dat zij en hun leerlingen op feitelijk niveau de vrijheid hadden zelf keuzes te maken. Er werd geen landelijk keurslijf voorgeschreven,

terwijl iedereen het toch in grote lijnen over hetzelfde had. Dat was ook de essentie van de voorstellen van de CHMV, die was uitgegaan van de opvatting dat een overheid in een democratische, pluriforme samenleving niet tot in detail kan voorschrijven welke historische kennis als canonic moet worden beschouwd.

Het ging echter wel om een landelijke, betrouwbare en valide toets. Zowel voor de uitvoerders van het onderwijscurriculum als voor de ontwikkelaars van de toetsopgaven betekende de geringe graad van specificatie van het programma een probleem. Bij elke opgave die een examenconstructiegroep ontwerpt, speelt historische kennis een rol. Voortdurend rees daarbij de vraag wat bij de examenkandidaten bekend mocht worden verondersteld en wat niet. Ook de docenten in het onderwijs worstelden steeds met het probleem wat ze nu wel of niet moesten behandelen. Daardoor ontstond in de groep van pilotdocenten aanvankelijk de neiging om aan te dringen op grotere duidelijkheid en specificatie. De projectleiding meende echter dat bij nadere specificatie van het programma niet meer aan de doelstelling van de pilot zou kunnen worden voldaan, namelijk om uit te proberen of een examen op basis van het CHMV-programma mogelijk was. Indien dit programma werd voorzien van een meer of minder uitgebreide specificatie, zou het niet meer hetzelfde programma zijn. Bovendien zou dan a priori zijn vastgesteld dat het CHMV-programma op zichzelf onvoldoende was. De projectleiding besloot daarom - in overleg met de pilotdocenten - het 'pure' programma een kans te geven zonder nadere uitleg, om te zien of onder die omstandigheden centrale examinering mogelijk zou blijken.

De projectleiding gaf aan het examenprogramma wel een zo groot mogelijke mate van duidelijkheid door de tekst van de 'kenmerkende aspecten' zorgvuldig te analyseren op de daarin voorkomende begrippen en inhoudselementen. In sommige gevallen leverde dat aanknopingspunten op, zoals in het geval van kenmerkend aspect 27 dat gaat over rationeel optimisme en verlicht denken toegepast op godsdienst, politiek, economie en sociale verhoudingen. In andere gevallen, zoals dat van kenmerkend aspect 40 ('Het voeren van twee wereldoorlogen') bleef de onduidelijkheid groot, omdat de tekst geen enkele nadere aanduiding bevatte. Op grond van de tekst van de kenmerkende aspecten stelde de projectleiding een lijst van 91 begrippen op waarvan duidelijk was dat de leerlingen die zouden moeten beheersen, omdat deze begrippen letterlijk in de tekst van het examenprogramma voorkomen.<sup>10</sup>

In de loop van de uitvoering van het project maakte de aanvankelijke onzekerheid bij de docenten over wat ze nu wel of niet in hun curriculum moesten opnemen plaats voor een grotere slagvaardigheid en vertrouwen de juiste keuzes te kunnen maken. Door het bespreken van sjabloonvragen in de plenaire groep werd duidelijker wat op het examen van de leerlingen verwacht zou worden. Later kwamen ook proefopgaven gereed die meer duidelijkheid verschaften. Bovendien leek bij de docenten de waardering toe te nemen voor het feit dat ze de vrijheid hadden zelf met concrete invullingen van hun curriculum te komen. Over hun bevindingen daarbij wordt verwezen naar het evaluatieve deel van dit hoofdstuk. Daar komt ook de vraag aan de orde of de omschrijving van het programma naar het oordeel van de docenten nu duidelijk genoeg is of niet.

---

<sup>10</sup> Zie bijlage 6 voor de begrippenlijst.

### 2.1.2. De beginsituatie van de leerlingen

De leerlingen die in het najaar van 2004 aan het onderwijstraject voor het pilotexamen begonnen, waren voor de overgrote meerderheid afkomstig uit de driejarige havo-onderbouw. Een kleinere groep bestond uit instromers met een vmbo-t-diploma. Daarnaast waren er enkele doubleurs. Voor alle leerlingen gold dat ze nooit eerder met het stelsel van tien tijdvakken, zoals voorgesteld door de CHMV, hadden kennisgemaakt.

Om enig zicht te krijgen op de beginsituatie van deze leerlingen m.b.t. het eindexamenprogramma werd onmiddellijk na het begin van de cursus 2004-2005 een tweetal onderzoeken onder hen uitgevoerd. Het eerste was een onderzoek naar de associaties die de leerlingen hadden met de tien tijdvakken als ze geconfronteerd werden met de naam en tijdsaanduiding ervan. Dit onderzoek was belangwekkend aangezien een veronderstelling in het CHMV-rapport is dat de tijdvakken 'beelden' bij leerlingen oproepen en op die wijze ondersteunend kunnen werken bij het creëren van een tijdsbesef.<sup>11</sup> In het CHMV-rapport wordt voor deze veronderstelling echter geen wetenschappelijke grond aangevoerd. Het tweede onderzoek dat werd uitgevoerd naar de beginsituatie van de leerlingen bestond uit een vijftigtal meerkeuzevragen die nauw aansloten bij de 49 kenmerkende aspecten (één vraag per kenmerk, m.u.v. kenmerk 8 over het ontstaan van christendom en jodendom, dat over twee vragen werd verdeeld). Op die wijze kon enige indruk worden verkregen over de bekendheid bij de leerlingen met de door de 49 kenmerkende aspecten gerepresenteerde leerstof. De meerkeuzetoets moest gemaakt worden ná de associatieve test omdat de vragen in de meerkeuzetoets de associaties van de leerlingen anders zouden kunnen beïnvloeden.

Het resultaat van de meerkeuzetoets liet zien dat de gemiddelde 'juist'-score uitkwam op 35%, de gemiddelde 'onjuist'-score op 43% en de gemiddelde 'weet niet'-score op 22%. Voor zover de vragen representatief waren voor de examenstof, zou dat betekenen dat aan het begin van het onderwijstraject gemiddeld ongeveer tweederde van de stof niet door de leerlingen beheerst werd. Een uitsplitsing naar tijdvak kan moeilijk gemaakt worden omdat er maar enkele vragen over ieder tijdvak gesteld werden.

Bij het afnemen van de associatieve toets werden aan de leerlingen formulieren voorgelegd waarop alleen de namen en de tijdsgrenzen van de tijdvakken voorkwamen, bijvoorbeeld 'Tijd van monniken en ridders, 500 tot 1000 n.Chr. (vroeg Middeleeuwen)'. Op deze formulieren moesten de leerlingen aangeven: 1) waar ze spontaan aan dachten bij een tijdvak, 2) welke gebeurtenissen uit een tijdvak ze zich herinnerden, 3) welke personen ze bij een tijdvak vonden passen, en 4) welke begrippen ze typerend vonden voor een tijdvak. Per rubriek mochten ze maximaal drie dingen noemen.<sup>12</sup>

De spontane associaties van leerlingen bleken dicht bij de namen van de tijdvakken te blijven. In tijdvak 1 bijvoorbeeld hielden van de 639 associaties 238 (37%) direct verband met jagers en boeren. Bij tijdvakken met minder beeldende namen was het aantal leerlingen dat zich er niets bij kon voorstellen dan ook het grootst: bij de tijd van regenten en vorsten schoot 37% van de leerlingen niets te binnen, bij de tijd van steden en staten 23%. Vrij associëren ging de leerlingen gemakkelijker af dan het noemen van historische personen of gebeurtenissen. Gemiddeld over de tien tijdvakken had slechts 14% van de leerlingen geen enkele associatie, maar 52% noemde geen enkele persoon. De gebeurtenissen vormden de moeilijkste categorie. Gemiddeld

<sup>11</sup> Verleden, heden en toekomst, pag. 20.

<sup>12</sup> Over de resultaten van de associatieve toets betreffende de beginsituatie van de leerlingen wordt uitvoerig gerapporteerd in: D. van Straaten, 'Berenvellen en markiespakjes'. Het tijdvakkenkader en beelden van tijd. In: Kleio 2006 nr 1, p. 2-7



over de tien tijdvakken noemde 60% van de leerlingen geen gebeurtenis.

De beelden van de leerlingen berustten zowel op schoolse als op buitenschoolse kennis. Ze waren vaak algemeen en stereotiep. De tijdvaknamen bleken te kunnen functioneren als 'triggers'. Woorden als 'ridders', 'pruiken' en 'Romeinen' appelleerden aan aanwezige kennis. Hoe concreter de naam, hoe meer er boven leek te komen. Ook al was de kennis van leerlingen beperkt, de tijdvaknamen zorgden ervoor dat ze zich tenminste iets bij de verschillende tijden voorstellen, hoe onjuist die voorstellingen soms ook waren.

De meeste pilotdocenten waren geneigd teleurgesteld te reageren op de in hun ogen magere resultaten van het onderzoek naar de beginsituatie van de leerlingen. Het onderwijs in de onderbouw had minder blijvend resultaat opgeleverd dan ze verwachtten. Een conclusie was dan ook dat er nog het nodige gedaan moest worden om de algemene kennis van de leerlingen op peil te brengen. De vrees voor een 'herhaling van de onderbouw' - behandelen van wat leerlingen al lang wisten en wat dus tot verveling zou leiden - werd erdoor verminderd. Ook kregen de docenten enig zicht op de wijze waarop leerlingen op de tijdvakbenamingen reageerden en hoe ze associaties zouden kunnen gebruiken bij het opbouwen van een tijdsbeeld,

### 2.1.3. De werkwijze van de verschillende scholen

Mede doordat vanuit de projectleiding geen curriculum of nadere interpretatie van het examenprogramma werden voorgeschreven, liep de aanpak van het onderwijs in de verschillende projectscholen zeer uiteen. Hierop was ook van invloed dat gedurende de looptijd van de pilot geen op het nieuwe examenprogramma gerichte schoolboeken aanwezig waren. In een aantal gevallen werd in deze leemte voorzien door een bestaand boek te gebruiken dat een algemeen historisch overzicht bevatte, maar daarin waren de tien tijdvakken en de 49 kenmerkende aspecten nog niet verwerkt. Docenten maakten overzichten van te leren stof per tijdvak, of lieten de leerlingen zulke overzichten maken.

Bij de aanpak van het curriculum zijn grofweg drie modellen te onderscheiden geweest:

- 1 Een volledig nieuw, chronologisch geordend curriculum, aan het begin van havo-4 startend met tijdvak 1 en in havo-5 eindigend met tijdvak 10 - met tussentijdse herhalingen en/of een herhaling onmiddellijk voorafgaand aan het CE. In deze situatie werd het programma van toetsing en afsluiting ingericht volgens deze aanpak, wat betekende dat telkens een aantal tijdvakken voor een schooltoets moest worden bestudeerd. Aan het eind van het programma werd een toets gegeven over alle tien de tijdvakken samen. Deze aanpak werd gehanteerd op het H.N. Werkman College in Groningen en op het Assink Lyceum in Haaksbergen.
- 2 Een curriculum dat zich in havo-4 eerst bezighield met (diachronische) thema's uit de gebruikelijke voor de havo-bovenbouw bestaande boeken, om pas later expliciet in te gaan op de tien tijdvakken. Tijdens de behandeling van de thema's werden wel die elementen eruit gelicht die verband hielden met het algemene referentiekader, zoals omschreven in de kenmerkende aspecten. In een enkel geval werd de expliciete behandeling van de tien tijdvakken zelfs uitgesteld tot havo-5 (Papendrecht). De schooltoetsen in het PTA waren gedeeltelijk gericht op thema's (zoals ook voorheen gebruikelijk), gedeeltelijk op 'tijdvakkenstof'. Ook hier werd afgesloten met een toets over alle tijdvakken samen. Deze aanpak werd gehanteerd op het Grotius College in Delft en op Scholengemeenschap De Lage Waard in Papendrecht.
- 3 Een curriculum waarbij een accent lag op zelfstandige oriëntatie door de leerlingen op de tien tijdvakken door het verzamelen van documentatie. Na een aanvankelijke introductie van het gehele stelsel van tijdvakken werden leerlingen aan het werk gezet om voorbeelden te verzamelen in documentatiemappen en zich

op die wijze langzamerhand een steeds beter ingevuld beeld van de tijdvakken te vormen. Leerlingen maakten samenvattingen en hielden presentaties. Tijdens de lessen werd ook aandacht besteed aan thema's. Zo'n aanpak werd gehanteerd op de Scholengemeenschap 'De Breul' in Zeist en op 2College Cobbenhagen in Tilburg. Hierbij werden soms creatieve oplossingen ingezet, zoals het in verband brengen met het tijdvakkenkader van buitenschoolse ervaringen (vakantie, een gelezen boek, een museum, een bijgewoonde toneelvoorstelling), of het door leerlingen laten voorbereiden van een grote quiz-avond over historische kennis, die in het bijzijn van ouders en andere belangstellenden daadwerkelijk plaatsvond.

Het curriculummodel was op iedere school weer anders, soms in een mengvorm tussen de beschreven aanpakken. Zo hanteerde men in Nijmegen eigen documentatiemappen van leerlingen per tijdvak, waarbij het bezwaar ontstond dat de mappen erg dik en onoverzichtelijk werden. De leerlingen konden eigenlijk geen juiste selectie maken en er overzicht in brengen, zodat ingrijpen van de docent nodig was. Iets dergelijks deed zich ook voor op Mill Hill College. Daarbij was bovendien een probleem dat door leerlingen vervaardigde tijdvakkenoverzichten als leerstof voor 'de anderen' functioneerden - men was dus van elkaar afhankelijk. Niet iedere leerlingengroep kweet zich in dit opzicht even goed van zijn taak.

De aanpakken werden in de loop van de tijd aangepast als bleek dat een bepaald model niet het resultaat opleverde dat men ervan verwachtte. Zo was de strategie van het 2College Cobbenhagen aanvankelijk erg leerlinggestuurd, maar deze werkwijze werd in de loop der tijd steeds meer vervangen door een docentgestuurde aanpak. Op andere scholen, zoals in Haaksbergen, was de werkwijze van meet af aan docentgestuurd.

Over de ervaringen van de docenten bij de uitvoering van de verschillende curricula wordt gerapporteerd in het evaluatieve deel van dit hoofdstuk.

#### **2.1.4. Samenwerking tussen en ondersteuning van deelnemende docenten**

In de opzet van het project was voorzien dat docenten van de deelnemende scholen zouden optreden als materiaalontwikkelaars waarbij aanpakken voor curricula en pta's, lesmateriaal, toetsen, e.d. zouden worden uitgewisseld. Met het oog hierop werden regelmatig plenaire bijeenkomsten gehouden voor alle deelnemende docenten.

Op de bijeenkomst van november 2004 werd besloten een digitaal uitwisselingsforum te creëren waarop docenten hun materiaal konden plaatsen. Op die wijze kon les- en toetsmateriaal tussen de scholen worden uitgewisseld. Er werd gebruik gemaakt van een BSCW-omgeving (Basic Support for Cooperative Work) van de Hogeschool van Amsterdam. Alle betrokken docenten van de pilotscholen hadden toegang tot deze omgeving.

Op de plenaire bijeenkomsten passeerde een reeks van onderwerpen de revue, deels op aanvraag van de deelnemende docenten. Zo werd in de loop van de twee projectjaren gesproken over:

- de algemene opzet van de pilot en de manier van omgaan met het nieuwe examenprogramma, vooral met het oog op het opstellen van programma's van toetsing en afsluiting (PTA);
- de manier waarop docenten bij de voorbereiding op het nieuwe examen gebruik zouden kunnen maken van bestaande boeken;
- de resultaten van de peiling van de beginsituatie (zie 2.1.2);

- de diverse werkwijzen van de scholen (zie 2.1.3);
- de te verwachten examenopgaven;
- de wijze waarop pilotscholen zelf geschikte toetsvragen konden ontwikkelen door gebruik te maken van de sjabloonvragen;
- de wijze waarop de 'kenmerkende aspecten' van tijdvakken voor leerlingen geconcretiseerd konden worden;
- de aanpak van historisch denken en redeneren (eindterm 4 t/m 7);
- de resultaten van het proefexamen;
- hoe op de valreep voor het examen het best kon worden geopereerd om leerlingen zoveel mogelijk overzicht te geven;
- hoe diende te worden omgegaan met correctie en tweede correctie van het examen.

De behoefte aan duidelijkheid omtrent de te verwachten examenopgaven was onder de docenten begrijpelijkerwijs groot. De moeilijkheid was hier dat het ontwikkeltraject van het nieuwe examen parallel liep aan de looptijd van de uitvoering van het onderwijs. Het geven van meer duidelijkheid over examenopgaven zou betekenen dat de constructiegroep zich vroegtijdig zou vastleggen en daarmee speelruimte zou verliezen om naar zo geschikt mogelijke opgaven te zoeken. Wel kon reeds in april 2005 aan de hand van voorbeelden duidelijk worden gemaakt volgens welke beginselen de constructiegroep bij het maken van het examen te werk ging. Nog voor de zomervakantie van 2005 kwamen ook enige voorbeeldopgaven beschikbaar die duidelijkheid schiepen. Op één van de plenaire vergaderingen werd een verzameling door de deelnemende docenten ingediende toetsopgaven besproken zoals zij meenden dat die op het examen zouden kunnen voorkomen. Door het commentaar hierop werd duidelijker wat wel of niet zou kunnen worden verwacht. In maart 2006 kwam tenslotte een volledig proefexamen beschikbaar dat scholen konden gebruiken als laatste test. Van dit proefexamen is door de projectscholen op verschillende manieren gebruik gemaakt: sommige zetten het in als officiële toets op een schoolexamen, andere deelden het alleen uit aan de leerlingen om zich een idee te vormen van wat zij konden verwachten.

Wat betreft de duidelijkheid van de kenmerkende aspecten en de concrete voorbeelden die daarbij zouden kunnen worden gebruikt: ook hier was onder de pilotdocenten behoefte aan informatie. Werd er aanvankelijk gesproken over een tekst die de kenmerkende aspecten zou kunnen specificeren, na enige tijd gaven de meeste docenten aan daaraan geen behoefte meer te hebben. Ze stemden ermee in dat een goede proef-neming bij een diversiteit aan curricula en aanpakken zou moeten plaatshebben. Bovendien begon men de vrijheid om zelf een curriculum te kunnen vormgeven te waarderen. Zie voor meer uitgewerkte gegevens hierover paragraaf 2.2.1. Van de kant van het IVGD werden in april 2005 een aantal voorbeelden van concretisering gepresenteerd (gebeurtenissen, personen) in de vorm van afbeeldingen met commentaar, gepresenteerd in powerpointdia's. Deze verzameling werd in juni 2005 gecompleteerd voor alle kenmerkende aspecten, resulterend in de cd-rom 'Tijdvakken Concreet'. Deze kon als oefenmateriaal voor de leerlingen worden ingezet. Op de cd-rom was ook een 'lege structuur' te vinden van deze presentatie, die leerlingen konden gebruiken om er zelf hun eigen invullingen en voorbeelden in te plaatsen. Mede hierdoor werd duidelijk dat de invullingen op de cd slechts mogelijke voorbeelden waren, geen voorgeschreven inhouden met een canonic karakter. De cd werd via de IVGD-website en via een melding in het blad 'Didaktief' (november 2005) ook voor andere belangstellenden beschikbaar gesteld. In de loop van 2005 en 2006 hebben ca 30 docenten in den lande de cd besteld.

De ondersteuning van de projectleiding heeft verder bestaan uit bezoeken aan alle pilotscholen door twee leden: Wilschut (IVGD) voor zaken betreffende het inhoudelijke curriculum en Marshall (SLO) voor zaken betreffende de begeleiding van het veranderingsproces. Tijdens deze bezoeken werd een intensievere gedachteswisseling mogelijk dan op de plenaire projectbijeenkomsten. De zaken van iedere school afzonderlijk werden grondig besproken. Hierdoor ontwikkelde de projectleiding ook een scherp beeld van de situatie op de afzonderlijke projectscholen.

De ondersteuning is door de docenten van de verschillende scholen naar eigen inzicht op verschillende wijze gebruikt. Een enkeling maakte intensief gebruik van de materialen op de BSCW-omgeving, maar uit de aantallen 'hits' op deze site bleek dat dit lang geen algemeen verschijnsel was. Misschien liep de situatie op de verschillende scholen te zeer uiteen om goed van elkaars materiaal gebruik te kunnen maken. Wellicht was men ook te zeer bezig met de implementatie op de eigen school om aan de situatie elders gericht aandacht te kunnen besteden. In de eerste fase van implementatie van een vernieuwing is dat vaak het geval. Pas als men zelf ervaringen heeft opgedaan met mogelijke problemen, gaat men gericht kijken naar wat anderen precies doen. Het vormen van een 'community of learners' onder de pilotdocenten is dan ook maar gedeeltelijk gelukt, nog het meest via de plenaire pilotbijeenkomsten. Discussie vond op de BSCW-omgeving in het geheel niet plaats, hoewel daarvoor wel mogelijkheden waren. Ook de inzet van de cd Tijdvakken Concreet en die van proefopgaven en proefexamen verschilden per school. Op sommige scholen is de cd in lessituaties gebruikt, op andere alleen uitgedeeld, en op sommige is hij niet uitgedeeld. Alleen de diagnostische toetsen en associatieve diagnose aan het begin van de cursusjaren zijn door alle scholen ingezet - dit op verzoek van de projectleiding.

## 2.2. Evaluatie van het onderwijs

### 2.2.1 Ervaringen van docenten

Een belangrijke vraag is of de scholen en docenten die deelnamen aan de pilot erin slaagden om een aanvaardbaar curriculum en een didactiek te realiseren waarmee leerlingen goed voorbereid worden op het beoogde examen. In hoeverre zijn docenten tevreden over het nieuwe curriculum en de gehanteerde didactiek en welke problemen hebben ze ervaren? De docenten die deelnamen aan de pilot is gevraagd naar hun mening over:

- de lijst met kenmerkende aspecten;
- hun manier van lesgeven;
- de leeropbrengsten;
- veranderingen in didactiek en organisatie die noodzakelijk worden geacht voor implementatie;
- de voor- en nadelen van het examenprogramma.

Gegevens zijn verkregen uit een vragenlijst die de docenten van de pilotscholen vlak voor het CE hebben ingevuld en interviews na het CE. De vragenlijst is door 15 van de 20 pilotdocenten ingevuld. Op elke pilotschool is een interview afgenomen waarbij tenminste één, maar doorgaans twee of drie pilotdocenten aanwezig waren.<sup>13</sup>

---

<sup>13</sup> De vragenlijst en de richtvragen voor de interviews zijn opgenomen in bijlagen 7 en 9.

### **De stofomschrijving in 'kenmerkende aspecten'**

De oriëntatiekennis die in het nieuwe centraal examen wordt getoetst, is in het examenprogramma omschreven in een kader van tien tijdvakken met 49 bijbehorende kenmerkende aspecten. Bij aanvang van het project vond ongeveer de helft van de docenten de stof te weinig gespecificeerd. Vlak voor het CE was dit iets gewijzigd: ruim de helft van de docenten was tevreden over de mate van specificatie. Uit de interviews die na het CE zijn afgenomen blijkt dat een aantal docenten die voor het examen de stofomschrijving nog te weinig gespecificeerd vond, die stofomschrijving na het examen wel acceptabel vond. Toch blijven er verschillen van mening bestaan over de vraag of de kenmerkende aspecten duidelijk genoeg zijn. Terwijl de meesten de kenmerkende aspecten voldoende duidelijk en gespecificeerd vinden, zorgt de globale omschrijving in combinatie met het ontbreken van een methode bij een aantal docenten toch wel voor enige onzekerheid. Enkele docenten gaven aan dat het ook een kwestie van wennen was om met de kenmerkende aspecten om te gaan. Of zoals één docent het verwoordde: 'Inmiddels is de generatie geschiedenisdocenten zo gewend aan de stofomschrijving dat elke verandering mogelijk bedreigend of tenminste onwennig is. Dat heb ik bij mijzelf geconstateerd.'

De kritiek op de kenmerkende aspecten richt zich met name op de formulering van de kenmerkende aspecten en de grote verschillen daarin. Sommige zijn heel uitgebreid en andere heel eenvoudig of zelfs onvolledig geformuleerd. Er is met name kritiek op de formulering van de kenmerkende aspecten van de laatste twee tijdvakken. Zo wordt het kenmerkende aspect 'het voeren van twee wereldoorlogen' (nr. 40) door veel pilot-docenten te vaag gevonden. Daarnaast is er onzekerheid over hoe diep je moet gaan bij het behandelen van de kenmerkende aspecten. Twee docenten gaven in dit kader aan dat ook de leerlingen behoefte hebben aan duidelijkere en concretere formuleringen. Er zouden dan meer concrete namen en plaatsen genoemd moeten worden. Anderen vinden de weinig concrete omschrijving van de kenmerkende aspecten geen probleem omdat hierin juist een taak voor de docent ligt en het leerlingen uitdaagt om verdieping te zoeken.

Als belangrijk voordeel van de kenmerkende aspecten werd door veel docenten genoemd dat ze ruimte bieden voor eigen invulling van de lessen. De kenmerkende aspecten bieden ruimte voor het behandelen van diverse onderwerpen en tot het betrekken van de actualiteit in de lessen. Er is altijd wel een verband te vinden met één van de aspecten.

### **De lessen**

Vooraf deelden de meeste docenten de verwachting dat de leerlingen in dit programma zelfstandig zouden kunnen werken en samenwerken, met name door hen zelf een tijdvakdossier te laten samen stellen. De rol van de docent zou dan meer komen te liggen in het begeleiden van de leerlingen bij het samenstellen van hun eigen dossier. De rol van het schoolboek zou kleiner zijn, omdat er op het moment van de pilot geen methode beschikbaar was die aansloot bij het programma. Eén docent had duidelijk andere verwachtingen: hij verwachtte meer te gaan doceren en vertellen.

In de praktijk heeft een deel van de docenten overwegend docentgestuurd gewerkt - een werkwijze die gekenmerkt werd door veel uitleg (soms in de vorm van 'hoorcolleges') en het aanbieden van informatie-materiaal. Een ander deel van de docenten realiseerde gedurende de pilot een overwegend leerlinggestuurde werkwijze, waarbij de leerling in belangrijke mate zelf verantwoordelijk werd gehouden voor het verwerven van de nodige tijdvak kennis en activerende werkvormen werden gehanteerd. Voorbeelden van zelfwerkzaam-

heid zijn: leerlingen zelf spotprenten of afbeeldingen laten zoeken die bij een kenmerkend aspect passen en de relatie laten uitleggen; leerlingen in groepjes kenmerkende aspecten laten uitwerken en aan elkaar presenteren; leerlingen zelf een begrippenlijst laten maken. Docenten lieten hun leerlingen gedurende twee jaar een tijdvakdossier aanleggen waarin alle belangrijke informatie over de tijdvakken verzameld werd. Ook in de praktische opdrachten werden leerlingen aan het werk gezet met het tijdvakkenkader, bijvoorbeeld door een historische film te koppelen aan kenmerkende aspecten, een onderzoek te verrichten naar een onderwerp dat zich over meer tijdvakken uitstrekt, een poster te maken van een tijdvak, of een markant persoon of gebeurtenis in de tijd te plaatsen.

De docenten zijn over het algemeen redelijk tevreden over het curriculum en de didactiek zoals die door hen in havo 4 en 5 gerealiseerd zijn. Vijf docenten gaven aan tevreden te zijn en één docent was zelfs zeer tevreden. Drie docenten waren ontevreden en zes niet tevreden maar ook niet ontevreden. Tevredenheid was er met name over de didactiek, de themakeuze, en over het uiteindelijke doel dat men bereikt had bij de leerlingen. De lessen waren volgens deze docenten levendiger, meer historisch en minder thematisch. Leerlingen een eigen dossier laten samenstellen werkte motiverend en daarnaast waren er veel mogelijkheden om ICT te gebruiken. Eén docent merkte op weer het gevoel te hebben geschiedenisles te geven. Een andere docent noemde de mogelijkheid om zowel een 'klassieke docent' te kunnen zijn als ook een coachende rol te kunnen aanleren. Waarover waren de docenten minder tevreden? Een veel genoemd punt was de onzekerheid aan het begin van de pilot en het zoeken naar een goede werkwijze en geschikt materiaal. Dit kostte veel tijd. Deze onzekerheid was er ook bij de leerlingen. Er werd op een andere manier lesgegeven dan in de onderbouw en dat was ook voor de leerlingen wennen. Dat laatste is overigens in het algemeen het geval bij de overgang van havo-3 naar havo-4. Een tweede punt waar meerdere docenten ontevreden over waren is dat ze nu minder diepgang in de lessen konden aanbrenge(n). Een aantal docenten was ontevreden over de geringe afwisseling in didactische werkvormen. Met name in het tweede jaar - het examenjaar - zijn een aantal docenten wat 'traditioneler' gaan werken door meer uitleg te geven en minder variatie in werkvormen toe te passen. Andere punten waarover docenten ontevreden waren zijn: te laat begonnen met het tijdvakdossier, leerlingen die te weinig verantwoordelijkheid namen voor hun tijdvakdossier en te weinig contacturen. Bij de scholen waar met thema's werd gewerkt, gaven enkele docenten aan het lastig te vinden de overzichtskennis te koppelen aan de thema's uit de methodes, met name omdat de gebruikte methodes niet op zo'n werkwijze waren afgestemd en het bij leerlingen veel onduidelijkheid schiep over wat nu geleerd moest worden. Ruim een kwart van de docenten gaf in de vragenlijst aan dat zij door de nadruk op het overzicht van tijdvakken minder aandacht aan historische vaardigheden hebben besteed dan zij gewend waren te doen. Een aantal van hen was hierover ook ontevreden.

Over het algemeen zijn de docenten van de pilotscholen dus redelijk tevreden. Ontevredenheid staat vooral in verband met de omstandigheid dat in het begin voor veel docenten niet echt duidelijk was waar het naar toe ging en hoe het examen er uit zou zien. Het was uitproberen. Na het CE was er over het algemeen een duidelijker idee ontstaan over hoe men de lessen wilde aanpakken. In de vragenlijst en interviews is docenten van de pilotscholen gevraagd wat ze, op basis van de ervaringen in 2005 en 2006, bij een nieuwe groep leerlingen anders zouden doen. Genoemd werden: meer nadruk leggen op overzichtskennis en meer verwijzen naar kenmerkende aspecten; meer controle op de tijdvakdossiers die leerlingen samenstellen; leerlingen meer laten samenwerken; minder stof behandelen en meer de diepte in; meer structuur aanbieden; gevarieerder lesgeven; meer tijd besteden aan de voorbereiding op het CE; leerlingen beter begeleiden bij hoe het aan te

pakken en te leren; eerder beginnen met de tijdvakken en meer aandacht voor historische vaardigheden.

#### **De leeropbrengsten**

Er was zowel bij de docenten als bij de leerlingen onzekerheid over hoe het examen er uit zou zien. Er waren behalve één oefenexamen geen voorbeelden van (oude) examenvragen beschikbaar. Vlak voor het CE (april 2006) twijfelden veel docenten van de pilotscholen eraan of de leerlingen wel goed voorbereid waren op het examen. Als redenen voor deze twijfel werden genoemd dat een deel van de leerlingen niet, of te laat was gaan werken; dat ze relatief weinig geoefend hadden met bronnen en vaardigheden; dat er geen schoolboek was als houvast; ten slotte waren er twijfels ten aanzien van lees- en schrijfvaardigheden van de leerlingen. Volgens de docenten vonden de leerlingen het lastig dat er geen boek was waardoor ze minder houvast hadden.

De docenten dachten voorafgaand aan het CE verschillend over de bereikte leeropbrengsten. Zes docenten vonden die tegenvallen, voor zeven docenten waren ze zoals verwacht en één docent vond ze beter dan verwacht. Tevredenheid was er over het feit dat de leerlingen die regelmatig werkten ook beter scoorden dan leerlingen die dat niet deden. Eén docent merkte op 'Ik heb niet het idee dat de leerlingen nu het vak beter beheersen. Ze hebben er echter wel met meer plezier aan gewerkt (ze vinden het leuker) en ze hebben ook een beter overzicht van wat er in de geschiedenis is gebeurd.' De docenten die de resultaten vonden tegenvallen, merkten op dat leerlingen nog te weinig overzicht hadden, dat er te weinig diepgang werd bereikt en dat de goede leerlingen niet genoeg aan hun trekken kwamen. De scores op het CE waren over het algemeen zoals de docenten verwacht hadden. Soms scoorde een hele klas beneden verwachting. Ook scoorden sommige leerlingen anders dan verwacht, soms beter, soms slechter. Maar, zo merkt een docent op, dit is ook het geval bij het reguliere examen.

#### **Voor- en nadelen**

Wat vinden de docenten van de pilotscholen de voordelen en nadelen van dit programma? Voor de meeste docenten had dit programma veel voordelen, voor een aantal zelfs alleen maar voordelen. De nadelen hadden voor een deel te maken met de al eerder genoemde onzekerheid die gepaard ging met deze pilot. Een enkele docent noemde als nadeel de vaagheid van enkele kenmerkende aspecten en de namen van enkele tijdvakken. Eén docent benadrukte het belang van het beginnen met dit programma in de onderbouw. 'Je bent nu eigenlijk een huis aan het bouwen zonder fundament.' Eén docent waarschuwde voor het gevaar dat er allerlei samenvattingen en dergelijke kunnen gaan circuleren op het internet. En dat de kans aanwezig is dat een leerling ook een voldoende kan halen zonder te leren.

De voordelen hadden zowel betrekking op de leerstof als op de eigen rol van de docent. Ten aanzien van de leerstof werd onder andere genoemd: de leerstof heeft een duidelijke structuur en samenhang, en een duidelijk minimum; meer accent op basisleerstof en minder accent op leesvaardigheid; meer mogelijkheden om op zaken terug te grijpen omdat deze als bekend verondersteld kunnen worden; de leerlingen krijgen een overzicht, een kapstok van de geschiedenis en hun kennis is minder fragmentarisch; de leerstof kan meer betekenisvol voor de leerlingen gemaakt worden; meer mogelijkheden voor het betrekken van de actualiteit; en leerlingen vinden deze kennis nuttig en van belang voor hun vervolgopleiding.

Als voordeel ten aanzien van de eigen rol als geschiedenisdocent werd genoemd dat er een grotere keuzevrijheid is om onderwerpen te behandelen zonder af te wijken van het programma. Hierdoor komt de eigen kennis van de docent meer tot zijn recht. Daarnaast werd het behandelen van alle tijdvakken afwisselender gevonden en werd het doel waarbij er meer nadruk ligt op overzichtskennis belangrijk gevonden. Een aantal docenten gaf aan het leuk en uitdagend te vinden om op een andere manier over het vak na te denken en hierin les te geven. Een aantal docenten kreeg het gevoel weer echt met geschiedenis bezig te zijn: 'Ik vind mezelf weer meer geschiedenisleraar' en 'Ik moet er niet aan denken om weer op de oude manier les te moeten geven.'

### 2.2.2 Ervaringen van leerlingen

In hoeverre waren leerlingen tevreden over het nieuwe curriculum en de gehanteerde didactiek en welke problemen hebben zij ervaren? De meningen van de leerlingen zijn op twee manieren geïnventariseerd. Op de eerste plaats is een vragenlijst ontwikkeld waarin leerlingen hun mening konden geven over het vak geschiedenis in 4 en 5 havo en over het CE.<sup>14</sup> De vragenlijst is direct na het CE afgenomen en ingevuld door 561 leerlingen van pilotscholen (85% van het totale aantal leerlingen dat het examen nieuwe stijl heeft gemaakt). Om enig zicht te krijgen op de mate waarin de meningen van de leerlingen gerelateerd zijn aan het nieuwe examenprogramma is dezelfde vragenlijst ook afgenomen bij leerlingen die het reguliere examen (over de thema's Vietnam en Volksopvoeding in Nederland) hebben gemaakt. Deze vragenlijst is door 299 leerlingen van vijf verschillende scholen ingevuld. Met behulp van SPSS (een statistiekprogramma) is met t-toetsen nagegaan of de verschillen tussen de meningen van de twee groepen leerlingen significant waren. Met significant (in de tabellen met een \* aangegeven) wordt bedoeld: een betekenisvol verschil waarbij de kans dat het verschil op toeval berust minimaal is (5%).

Op de tweede plaats is een bijeenkomst georganiseerd waaraan van elke pilotschool twee tot vier leerlingen deelnamen. Er waren 21 leerlingen aanwezig. Met deze leerlingen is afwisselend plenair en in kleine groepen gesproken over hun ervaringen met het nieuwe examenprogramma. De vragen van de vragenlijst waren uitgangspunt voor het gesprek.

#### Kenmerkende aspecten

De kenmerkende aspecten boden houvast aan de leerlingen. In de leerlingenbijeenkomst gaven veel leerlingen wel aan dat ze het moeilijk vonden om onderscheid te maken tussen wat wel en niet belangrijk is en of iets wel of geen goed voorbeeld is van een kenmerkend aspect. De leerlingen werd gevraagd om aan te geven welke drie kenmerkende aspecten ze het moeilijkst vonden. De kenmerkende aspecten die het meest genoemd werden, zijn: ontstaan en verspreiding van de islam (nr. 10), conflict in de Nederlanden dat resulteerde in de stichting van de Nederlandse staat (nr. 22), moderne vorm van imperialisme die verband hield met de industrialisatie (nr. 36) en opkomst van politiek-maatschappelijke stromingen (nr. 31). Deze kenmerkende aspecten komen niet overeen met de kenmerkende aspecten die door een aantal pilotdocenten als te weinig gespecificeerd werden beoordeeld. Opvallend is dat het om kenmerkende aspecten gaat waarin vormen van historische ontwikkeling en oorzaak-gevolgrelaties voorkomen.

<sup>14</sup> De vragenlijst is opgenomen in bijlage 8.



### De lessen

Zowel de leerlingen van de pilotscholen als de leerlingen van de scholen die het reguliere programma volgden gaven in de vragenlijst aan de geschiedenislessen in havo 4 en 5 interessant te vinden (zie Tabel 1). De leerlingen die het reguliere programma hadden gevolgd vonden de lessen echter interessanter dan de leerlingen van de pilotscholen. Dit kan te maken hebben met de aanwezigheid van het thema 'Vietnam' in het reguliere examen van 2006. De leerlingen van de scholen die het reguliere programma hadden gevolgd vonden het vak wat moeilijker dan de leerlingen van de pilotscholen. Beide groepen vonden dat ze in havo 4 en 5 veel nieuwe dingen geleerd hadden ten opzichte van de onderbouw. De leerlingen die het reguliere programma volgden, waren het wel meer eens met deze bewering dan de leerlingen uit de pilotgroep.

Tabel 1. Mening van leerlingen die het pilotexamen en leerlingen die het reguliere examen hebben gemaakt over de lessen geschiedenis in 4 en 5 havo

Items uit de vragenlijst	Pilot examen			Regulier examen		
	N	Gem.	St.dev.	N	Gem.	St.dev.
Ik vond de geschiedenislessen in havo 4 en 5 over het algemeen interessant [1 = zeer mee oneens – 5 zeer mee eens]	532	3.38	1.04	277	3.55*	.98
Ten opzichte van de geschiedenislessen in de onderbouw, heb ik in havo 4 en 5 [1 = heel veel nieuwe dingen geleerd – 5 = niets nieuws geleerd]	542	2.18*	.74	287	1.98	.63
Ik vind geschiedenis een moeilijk vak	532	2.84	1.07	277	3.06*	1.09

N = aantal leerlingen dat de vraag heeft ingevuld

\* significant verschil tussen pilot en regulier ( $p < .05$ ).

Tijdens de leerlingenbijeenkomst bleek dat leerlingen verschillend dachten over het ontbreken van een schoolboek. Sommige leerlingen misten een schoolboek, andere juist niet, die vonden het prettig om informatie zelf op te zoeken. De leerlingen vonden niet dat de leerstof in de bovenbouw een herhaling van de stof uit de onderbouw was. Hoewel veel gebeurtenissen misschien al wel bekend waren, gaven de leerlingen aan nu meer verbanden te zien en gebeurtenissen beter te kunnen koppelen. In tegenstelling tot de onderbouw was er nu meer kans om over allerlei zaken te discussiëren.

### Leeropbrengsten

Hoe denken leerlingen over de leeropbrengsten van het geschiedenisonderwijs dat ze in 4 en 5 havo genoten hebben? Leerlingen uit de pilotgroep waren het gemiddeld minder eens met de stelling 'volgend jaar zal ik het meeste van wat ik de afgelopen twee jaar geleerd heb weer vergeten zijn', dan leerlingen die het reguliere examen maakten. Ook tijdens de bijeenkomst met leerlingen van de verschillende pilotscholen vlak na het CE kwam naar voren dat de leerlingen het idee hebben dat ze deze leerstof niet snel zullen vergeten. Redenen die leerlingen daarvoor geven zijn dat je met dit programma niet een korte periode leert voor het examen, maar er twee jaar mee bezig bent, en dat de stof steeds weer terug komt.

In de vragenlijst is leerlingen ook gevraagd naar wat ze nu geleerd hebben. Tabel 2 geeft een overzicht van de resultaten. Voor leerlingen die het pilotexamen hebben gemaakt waren een 'goed overzicht van de oudheid tot nu' en 'het in de tijd kunnen plaatsen van gebeurtenissen' de belangrijkste leeropbrengsten. De leerlingen van de pilotscholen scoorden hoger op deze leeropbrengsten dan leerlingen die het reguliere examenprogramma hebben gevolgd. Deze uitkomst is in lijn met de aard en de doelstellingen van het nieuwe examenprogramma. Voor leerlingen die het reguliere examenprogramma hadden gevolgd waren het 'beter begrijpen hoe en waardoor samenlevingen (kunnen) veranderen' en 'geschiedenis interessanter gaan vinden' de belangrijkste leeropbrengsten. De leerlingen die het reguliere programma volgden scoorden hoger dan de leerlingen uit de pilotgroep op 'de wereld van vandaag beter begrijpen' en 'kritischer leren denken'. Deze verschillen tussen de leerlingen van de pilotgroep en de leerlingen van niet-pilot scholen zouden te maken kunnen hebben met de constatering van een aantal pilotdocenten dat ze zich bij de implementatie van het nieuwe examenprogramma vooral op de overzichtskennis hebben gericht, waardoor de andere eindtermen van het examenprogramma minder aandacht hebben gekregen dan ze wellicht verdienen. Hierbij moet wel gezegd worden dat het hier gaat om wat leerlingen zelf zeggen dat ze geleerd hebben. We beschikken niet over toetsgegevens waarmee we kunnen controleren in hoeverre de kennis en vaardigheden van leerlingen uit de pilotgroep daadwerkelijk verschillen van die van leerlingen uit het reguliere examenprogramma.

Tabel 2. Mening van leerlingen die het pilotexamen en leerlingen die het reguliere examen hebben gemaakt over wat ze geleerd hebben.

Items uit de vragenlijst	Pilot examen			Regulier examen		
	N	Gem.	St.dev.	N	Gem.	St.dev.
Volgend jaar zal ik het meeste van wat ik de afgelopen twee jaar geleerd heb weer vergeten zijn [ 1 = zeer mee oneens – 5 = zeer mee eens]	532	2.72	.99	277	3.03*	1.03
Door de geschiedenislessen in havo 4 en 5:						
-heb ik een goed overzicht gekregen van de oudheid tot nu	536	3.73*	.76	288	2.97	.89
-begrijp ik beter hoe en waardoor samenlevingen (kunnen) veranderen	536	3.55	.74	288	3.45	.80
-kan ik gebeurtenissen uit het verleden beter in de juiste tijd plaatsen	536	3.54*	.80	288	3.32	.84
-begrijp ik de wereld van vandaag beter	536	2.91	.85	288	3.06*	.89
-heb ik kritischer leren denken	536	2.87	.84	288	3.12*	.93
-heb ik dingen geleerd die ik kon gebruiken bij andere vakken (bijv. CKV of aardrijkskunde)	536	3.07	.96	288	3.11	1.00

-heb ik dingen geleerd die nuttig zijn voor mijn vervolgstudie of latere beroep	536	2.57	.96	288	2.49	1.04
-heb ik dingen geleerd waarvan ik vind dat iedereen ze eigenlijk zou moeten weten	536	3.32	.94	288	3.34	.95
-ben ik geschiedenis interessanter gaan vinden	536	3.18	.98	288	3.39*	.97
-ben ik geschiedenis een belangrijker vak gaan vinden	536	3.06	.91	288	3.17	.92

N = aantal leerlingen dat de vraag heeft ingevuld

\* significant verschil tussen pilot en regulier ( $p < .05$ ).

Uit de bijeenkomst met de leerlingen kwam het beeld naar voren dat de leerlingen over het algemeen tevreden waren over het nieuwe examenprogramma. Met name dat ze nu een duidelijk overzicht over het verleden hebben gekregen, vonden ze erg prettig. Veel leerlingen noemen dit ook als belangrijkste leeropbrengst, naast meer zicht op allerlei verbanden en oorzaken en gevolgen van ontwikkelingen. Een leerling verwoordde het als volgt: 'nu ken ik het volledige alfabet.' Ze vonden de leerstof nuttig voor hun algemene ontwikkeling. De leerlingen vonden het prettig dat je niet zoveel jaartallen hoefde te leren. Een aantal leerlingen miste de diepgang. Ook vonden sommige leerlingen het jammer dat er weinig aandacht was voor Nederlandse geschiedenis, de laatste vijftig jaar en voor actualiteit. De meningen hierover verschilden erg per school.

### 2.2.3 Eisen aan didactiek en organisatie

Welke verandering in didactiek en organisatie wordt voor de invoering van het nieuwe curriculum noodzakelijk geacht? Vooral het feit dat er tijdens de pilot geen op het tijdvakkenkader geschreven bovenbouwmethode beschikbaar was, maakte dat docenten op een wat andere manier gingen werken. Waar de lessen voor het oude examen vaak meer een trainingskarakter hebben, richt het nieuwe programma zich meer op het opbouwen. Het programma biedt volgens de docenten van de pilotscholen ruimte om de lessen op veel verschillende manieren didactisch vorm te geven. Die vrijheid vonden de docenten die meededen aan de pilot prettig. Sommige docenten hebben hun manier van werken niet echt aangepast. Andere docenten wel. Zij kozen voor een scenario waarbij leerlingen meer zelfstandig en in groepen werken en meer gebruik maken van ICT.

Op drie pilotscholen vond men het geringe aantal contacturen (1,5 of 2) problematisch voor invoering van het nieuwe curriculum. Op de diverse pilotscholen waren niet overal dezelfde aantallen contacturen voor geschiedenis beschikbaar. Het aantal 1,5 of 2 was hierbij het kleinste. Voor meer diepgang zijn volgens de docenten op deze scholen meer contacturen nodig. De docenten van de pilotscholen is gevraagd welke kennis en vaardigheden leerlingen moeten beheersen om in de eindexamenjaren met de leerstof voor het nieuwe examenprogramma om te kunnen gaan. De meeste docenten waren van mening dat het prettig zou zijn als leerlingen al bekend zouden zijn met de tijdvakken en (enkele) kenmerkende aspecten. De leerstof zou dan meer cyclisch opgebouwd kunnen worden, waarbij wat in de onderbouw op de rails gezet wordt, in de bovenbouw uitgebreid wordt. In de onderbouw wordt dan het fundament gelegd voor de bovenbouw, zowel qua inhoud als manier van leren. De pilotdocenten zijn niet bang voor te veel herhaling. Het is 'een illusie dat leerlingen alles onthouden', aldus één van de docenten en dus is enige herhaling op zijn plaats. Voor de

leerlingen zou herhaling niet vervelend zijn, ze vinden het volgens een aantal docenten juist fijn om dingen te herkennen. Eén docent gaf aan dat het kunnen denken in termen van grote lijnen en historisch besef pas echt begint op vijftien-, zestienjarige leeftijd. Als de tijdvakken in de onderbouw al aan de orde zijn geweest, is het mogelijk om de lessen in de bovenbouw meer te richten op het verdiepen en verbreden van de leerstof. Er kunnen dan andere voorbeelden van de kenmerkende aspecten gegeven worden, de actualiteit kan er meer bij betrokken worden, er kan meer aandacht zijn voor verbanden en grote lijnen, en er kan samenhang gezocht worden met andere vakken. In de bovenbouw zou volgens één pilotschool ook meer thematisch gewerkt kunnen worden. De meeste pilotdocenten zien geen spanning tussen het nieuwe examenprogramma en de vernieuwde onderbouw, maar geen van de pilotscholen stapt dan ook over naar een scenario van leergebieden (scenario 3 of 4).

## 3 | Het examen

Dit hoofdstuk valt in drie delen uiteen: een beschrijvend deel (3.1) waarin wordt uiteengezet hoe het examen totstandgekomen is en wat de toetstechnische resultaten ervan waren, een evaluatief deel (3.2) waarin wordt beschreven hoe de deelnemende docenten en leerlingen achteraf op hun ervaringen met het examen terugkijken en deze waarderen, en een evaluatief deel (3.3) waarin het oordeel over het examen van deskundigen in binnen- en buitenland wordt weergegeven.

### 3.1 Totstandkoming van het examen en toetstechnische analyse van het resultaat

#### 3.1.1 Uitgangspunten

In september 2004 is onder leiding van een toetsdeskundige van Cito een constructiegroep opgericht voor de productie van vragen voor het havo-examen geschiedenis nieuwe stijl. Bij de samenstelling van de constructiegroep zijn grotendeels dezelfde criteria gehanteerd als bij een constructiegroep voor het reguliere examen. De constructiegroep bestond uit drie docenten met een lestaak in de tweede fase van het havo, waarvan twee werkzaam op scholen die deelnamen aan het pilotproject en een docent van een school die niet aan het project deelnam. Twee leden combineerden het lidmaatschap van de reguliere constructiegroep met dat van de experimentele constructiegroep. Voor deze opzet is gekozen om met zoveel mogelijk ervaren constructeurs te kunnen werken. Van alle docenten is het contract verlengd voor het tweede experimentjaar.

Uitgangspunt voor de constructieopdracht vormden:

- 1 de kwaliteitscriteria die voor elke toets gelden op het gebied van validiteit, efficiëntie (bruikbaarheid) betrouwbaarheid en aanvaardbaarheid;
- 2 de opdracht die is voortgekomen uit de 'informele werkgroep', waarop de hoofdvraag van het pilotproject is gebaseerd: Welke vorm of vormen van examinering zijn passend en praktisch haalbaar, uitgaand van de examenprogramma's geschiedenis (uitgezonderd het programma voor het combinatievak geschiedenis en maatschappijleer) zoals omschreven in het rapport van de Commissie Historische en Maatschappelijke Vorming?;
- 3 de bestaande examenpraktijk: de vorm van het bestaande reguliere havo-examen diende zoveel mogelijk gevolgd te worden; dat betekende onder andere dat er open vragen geconstrueerd moesten worden en dat het een 'papieren examen' zou worden.

Vooraf was in de 'informele werkgroep' al vastgesteld dat de door de CHMV voorgestelde casustoetsing als niet praktisch en onhaalbaar van de hand gewezen moest worden. In een brief aan de Minister van OC&W d.d. 14 maart 2005 zijn de overwegingen om tot dit besluit te komen samengevat (zie paragraaf 1.2.).

#### 3.1.2 Validiteit

De bepaling van de begripsvaliditeit van centrale toetsvragen bij het vak geschiedenis is door de aard van het vak vaak gecompliceerd. Door het talige karakter van geschiedenis en het gebruik van tekstbronnen bij de vragen is het onderscheid tussen een vraag die historische kennis en inzicht toetst en een vraag die tekstbegrip en taalvaardigheid van de kandidaat toetst, soms moeilijk te maken. Door bespreking van de vragen in de constructiegroep (cg) en in de cevo-vaksectie en de vragen vervolgens voor te leggen aan interne en externe screeners, is geprobeerd de vragen zo 'historisch' mogelijk te houden.

De reguliere examens worden door een universitair hoogleraar met vakspecialistische kennis op het terrein

van het examenonderwerp gescreend op vakinhoudelijk validiteit, dat wil zeggen: er wordt nagegaan of de inhoud van de leerstof in overeenstemming is met de huidige wetenschappelijke inzichten. Dat impliceert dat per examen twee hoogleraren de vragen bekijken om vakinhoudelijke onvolkomenheden te voorkomen. Voor het pilotexamen zouden veel meer deskundigen moeten worden ingeschakeld om vakspecialisten van hetzelfde niveau als bij de wisselende onderwerpen te kunnen vinden. Het vakgebied dat het nieuwe examen bestrijkt is veel groter dan bij het reguliere examen, vakinhoudelijke fouten zijn sneller gemaakt. Het pilotexamen is niet op deze wijze door hoogleraren gescreend. Voor dit probleem is nog geen structurele oplossing gevonden, maar het examen is wel in de projectleiding van de pilot besproken en gescreend op overeenstemming met de eindtermen van het examenprogramma.

Bij de reguliere examens is sprake van de uitwerking van een (sub)domein in een stofomschrijving die de inhoud van de leerstof voor het examen vastlegt. Die stofomschrijving bepaalt de inhoudsvaliditeit van het reguliere examen. Het pilotexamen kent geen stofomschrijving die de inhoud van de leerstof omschrijft, maar de kandidaten moeten 49 kenmerkende aspecten, verdeeld over tien tijdvakken met eigen voorbeelden, kunnen verduidelijken. Om een evenwichtige spreiding te bereiken is, conform de huidige methodiek, een matrijs opgesteld die de relatieve vulling van de tijdvakken met kenmerkende aspecten als uitgangspunt voor de spreiding van de vragen over de stof neemt. De tien tijdvakken zijn samengevat in vier grotere gehelen: Prehistorie/Oudheid, Middeleeuwen, Vroegmoderne tijd en Moderne tijd - in overeenstemming met de indeling in eindterm 3 van domein A. De verdeling van de stof over deze vier eenheden is toegepast om een hanteerbare toetsmatrijs te verkrijgen. Een matrijs met een toedeling van puntenaantallen specifiek aan elk van de tien tijdvakken zou te gedetailleerd en daardoor onwerkbaar zijn. Er zou ook te veel nadruk komen te liggen op een indeling die ontworpen is als een didactisch hulpmiddel. De indeling in vier rubrieken komt in deelopjes terug in het examen. De verdeling van de te behalen punten is gebaseerd op het aantal kenmerkende aspecten voor elke periode, waaruit volgt dat aan de Vroegmoderne en de Moderne Tijd relatief meer aandacht wordt besteed, conform de tekst van het examenprogramma. In tabel 3 wordt dit weergegeven.

**Tabel 3: Toetsmatrijs met het relatieve gewicht van de perioden in het pilotexamen.**

Tijdvak	percentage	aantal punten 80 – 85
Prehistorie en Oudheid	± 15% van het aantal punten	12 – 13 punten
Middeleeuwen	± 15% van het aantal punten	12 – 13 punten
Vroegmoderne tijd	± 25% van het aantal punten	20 – 21 punten
Moderne tijd	± 45% van het aantal punten	36 – 38 punten
Totaal		80 – 85 punten

Behalve de inhoudelijke eindterm van domein B diende het historische denken en redeneren (structuurbegrippen en vaardigheden) van domein A in het examen getoetst te worden. Over het algemeen wijkt dit niet sterk af van het bestaande examenprogramma. Een belangrijk verschil vormen de eindtermen 1 en 2 van domein A, waarin het stelsel van tien tijdvakken wordt beschreven, eindterm 3, waarin de tien tijdvakken worden gekoppeld aan de traditionele periodenindeling, en een deel van eindterm 8, namelijk de oriëntatie op de hedendaagse werkelijkheid ('De kandidaat kan voor elk van de tien tijdvakken die genoemd zijn in eindterm 2 uitleggen hoe kennis van het betreffende tijdvak de oriëntatie op de hedendaagse werkelijkheid beïnvloedt').

Voor toetsing van dit deel van eindterm 8 zijn nog geen goede vragen ontwikkeld. Wel is daartoe een aanzet geleverd door vragen 'door de tijd heen' te stellen, waarbij vraagstukken in verschillende tijdvakken aan de orde worden gesteld.

### 3.1.2 Validiteitsproblemen bij de constructie

Uitgangspunt van de CHMV was, dat het voorgeschreven stelsel van tijdvakken met kenmerkende aspecten een 'lege' canon is, die docenten, scholen en leerboekenschrijvers zelf mogen invullen. De 49 kenmerkende aspecten vormen samen de oriëntatiekennis, die niet wordt ingevuld met concrete gebeurtenissen, personen, jaartallen of feiten. Voor het schoolexamen is dit geen groot probleem, voor een centrale toets wel. Binnen het gesloten systeem van een klas, met een beperkte groep leerlingen en de eigen docent kunnen onduidelijkheden, interpretatieverschillen en misverstanden sneller gesignaleerd en opgelost worden dan bij een centraal examen.

Tijdens het eerste constructiejaar zijn een aantal stofinhoudelijke validiteitproblemen naar voren gekomen:

- Het eerste probleem vormen de tien tijdvakken zelf. De tijdsbalk van de CHMV is geen historische tijdsbalk, maar een as die kalendertijd weergeeft: 3000 v Chr -500 na Chr; 500-1000; 1000-1500; 16e, 17e, 18e, 19e eeuw; 1e helft 20e en 2e helft 20e eeuw. Zo'n as wordt in het onderwijs gebruikt om de (van interpretatie afhankelijke) historische tijd (bijvoorbeeld gebeurtenissen of periodes) te rangschikken in kalendertijd. Dat is een verschil met de op stofomschrijvingen gebaseerde examens, want die gaan over historische gebeurtenissen met een door het onderwerp en de vraagstelling van de stofomschrijving bepaalde tijdsindeling die op een tijdsbalk van kalendertijd kan worden afgezet. Bijvoorbeeld in het examenthema Indonesië dat in 2007 en 2008 geëxamineerd wordt is 1600 (dicht bij de oprichting van de VOC in 1602) een door het onderwerp bepaald beginpunt en 1950 (dicht bij de soevereiniteitsoverdracht) een binnen het thema logisch eindpunt. De globale aanduiding van de grenzen van de tijdvakken is daarom een andere manier om met jaartallen om te gaan dan in een afgebakende stofomschrijving waarbij het gaat om jaartallen met daaraan gekoppelde specifieke (en vaak concrete) gebeurtenissen. Met name in de eerste tijdvakken (prehistorie en oudheid en de vroege middeleeuwen) zijn de begrenzingen in jaartallen erg globaal. Vooral wat de prehistorie betreft is de afbakening ervan in de tijd sterk afhankelijk van het gebied en de cultuur die wordt bestudeerd. Voor de vraagconstructie betekent deze onduidelijkheid een probleem bij het formuleren van eenduidige antwoordkernen; doordat de tijdgrenzen van een juist antwoord minder exact aangegeven zijn, zijn meer juiste antwoorden (door de kandidaat gegeven voorbeelden) mogelijk.
- In de tweede plaats gaat de CHMV ervan uit dat elk van deze kalendertijdvakken specifieke historische kenmerken heeft, waardoor het verschilt van andere tijdvakken. Deze 'kenmerkende aspecten' vormen elk een thema dat in het betreffende tijdvak past, maar waarvan niet duidelijk is in hoeverre dit aspect verder een rol speelt in de volgende of in de vorige tijdvakken. De formulering van de kenmerkende aspecten zelf geeft daarbij, op een enkele na, geen houvast. Dit maakt de constructie van doorsnedevragen moeilijk. In het eerste constructiejaar hebben de toetsconstructeurs zich geconcentreerd op de bevraging van de kenmerkende aspecten en de koppeling van die kenmerkende aspecten aan de tijdvakken waarin zij zijn geplaatst. Pas bij een koppeling van de kenmerkende aspecten aan de eindtermen van domein A, (bijvoorbeeld: 'in historische processen de samenhang tussen verandering en continuïteit beschrijven' - eindterm 4) zou er meer ruimte voor doorsnedevragen kunnen worden gecreëerd. Maar van die mogelijk-

heid heeft de constructiegroep (nog) spaarzaam gebruik gemaakt. De kenmerkende aspecten staan binnen hun tijdvak niet in chronologische volgorde en er worden (vrijwel nergens) causale verbanden tussen de aspecten gelegd. Op grond van eindterm 7 van domein A ('in het kader van een historische vraagstelling verklaringen geven voor historische gebeurtenissen, verschijnselen en ontwikkelingen') zouden er wel vragen naar causale relaties gesteld mogen worden, maar de mogelijkheden die deze eindterm biedt zijn nog maar in beperkte mate benut door de constructiegroep. Mogelijke verschuiving in betekenis van begrippen (bijvoorbeeld de invulling van het begrip democratie door de tijd heen) wordt niet nader aangeduid. Vooral de eerste tijdvakken (prehistorie met name) leverden problemen op door hun grootte en de moeilijk af te bakenen inhoud van de kenmerkende aspecten die daarin horen. Globale aanduidingen van culturen als 'de leefwijze van jager- verzamelaars' of 'het ontstaan van landbouw en landbouwsamenlevingen' hebben niet meer dan oppervlakkige vragen opgeleverd.

- Er bestaan grote verschillen in aard en karakter tussen de omschrijving van de verschillende kenmerkende aspecten. Soms wordt wel een oorzaak-gevolgrelatie gelegd ('de protestantse reformatie die splitsing van de christelijke kerk in West-Europa tot gevolg had' - nr 21) soms wordt volstaan met een zeer summiere aanduiding ('het voeren van twee wereldoorlogen' - nr 40), soms wijkt het kenmerkende aspect af van de andere kenmerkende aspecten ('de Grieks-Romeinse vormentaal' - nr 5).

Behalve toetsing van oriëntatiekennis, vereist het examenprogramma meer. 'Uitleggen hoe kennis van het betreffende tijdvak de oriëntatie op de hedendaagse werkelijkheid beïnvloedt' (eindterm 8) heeft tot nu toe geen voor een centrale toets aanvaardbare vragen opgeleverd. Er zijn wel aanzetten gezocht in pogingen om bepaalde verschijnselen door meerdere perioden heen te bevragen, maar of dit de eindterm werkelijk dekt is de vraag. Of deze eindterm op een centrale toets wel kán leiden tot een vraag met een aanvaardbaar, objectief, antwoordmodel blijft onderdeel van het debat in constructiegroep en Cevo.

Het historisch denken en redeneren van domein A is in verschillende vragen aan de orde gesteld. De discussie over de manier waarop deze redeneerwijzen, die voor het grootste deel ook in het oude programma voorkwamen (verandering en continuïteit, bronnen, oorzaak-gevolgrelaties, feit en mening onderscheiden, plaats- en tijdgebondenheid onderkennen, argumenteren) was niet veel anders dan bij de toetsing van vaardigheden in het reguliere examen.

Domein A voert als titel 'Historisch Besef'. In de discussie rondom het rapport van de CHMV is dit begrip voortdurend aan de orde geweest. In de informele werkgroep is afgesproken ervan uit te gaan dat 'historisch denken en redeneren' in veel opzichten het 'historisch besef' omvat. In hoeverre in het pilotexamen toetsing van historisch denken en redeneren heeft plaatsgevonden in een erkende en herkenbare vorm, blijft onderwerp van discussie en wordt nader onderzocht.

### 3.1.3 De toetstechnische gegevens

Na de versnelde correctie zijn door de docenten van de pilotscholen de gegevens van al hun examenkandidaten ingezonden via Wolf, het digitale verwerkingssysteem van Cito. Deze gegevens zijn op dezelfde manier verwerkt als bij het bestaande examen. Daardoor zijn de resultaten van het pilotexamen goed vergelijkbaar met die van het reguliere examen.



In tabel 4 is de uitslag van het pilotexamen weergegeven naast die van de reguliere havo-examens van de afgelopen vijf jaar. Dit is het resultaat bij N=1, dat wil zeggen: voor de aanpassing van de N-term. (Aanpassing van de N-term houdt in dat de examenresultaten door aanpassing van de cijferscores gemiddeld op een bepaald gewenst peil worden gebracht. De pilotexamens zijn op deze wijze gelijk getrokken met de reguliere examens, dit om de pilotleerlingen niet te benadelen. Ook van reguliere examens wordt de N-term aangepast om ongerechtvaardigde schommelingen in de resultaten tegen te gaan. De resultaten vóór aanpassing van de N-term geven dus inzicht in de 'reële resultaten' van een examen). Uit de tabel blijkt dat het resultaat van het pilotexamen onder het gemiddelde van het reguliere havo-examen van dit jaar, en ook onder het gemiddelde van de afgelopen vijf jaar ligt. Voor het percentage onvoldoendes geldt dat dit bij het pilotexamen hoger ligt bij beide vergelijkingen.

**Tabel 4: Resultaten van het reguliere examen en het pilotexamen voor N=1.**

jaar	gemiddeld	cijfer	percentage onvoldoende	
			regulier	pilot
	regulier	pilot	regulier	pilot
2006	6,7	6,1	12%	24%
2005	6,4		19%	
2004	6,7		11%	
2003	5,9		32%	
2002	5,8		41%	
gemiddeld	6,3		21%	

De moeilijkheidsgraad van een examen kan worden bepaald door de p-waarde vast te stellen. De p-waarde van een vraag is het percentage van de leerlingen dat een vraag juist heeft beantwoord. De p-waarde van een hele toets is het gemiddelde van alle p-waarden van vragen. Toetsconstructeurs streven naar een gemiddelde p-waarde tussen 40 en 70 per vraag. Waarden lager dan 40 duiden op een te moeilijke, waarden hoger dan 70 op een te makkelijke vraag. Het gemiddelde van een heel examen moet rond de 55-60 liggen. In tabel 5 is weergegeven welke gemiddelde p-waarde het pilotexamen had en hoe zich dit verhoudt tot gemiddelde p-waarden van reguliere examens. Uit de gegevens blijkt dat het pilotexamen aanzienlijk moeilijker was dan het reguliere examen van 2006, en (met een minder groot verschil) moeilijker was dan het gemiddelde van de afgelopen vijf jaar. De gemiddelde p-waarde van het pilotexamen komt meer overeen met de streefwaarde dan die van het reguliere examen van 2006.

**Tabel 5: Moeilijkheidsgraad, p-waarden regulier en pilotexamen.**

jaar	regulier	pilot
2006	63,3	56,2
2005	59,6	
2004	62,9	
2003	54,6	
2002	52,8	
gemiddeld	58,6	

In het reguliere examen worden vragen met bronnen aanzienlijk beter gemaakt dan vragen zonder bronnen. In het pilotexamen was dat net zo. Uit tabel 6 blijkt dat op vragen zonder bronnen minder is gescoord dan op vragen met bronnen. Bij beide examenvormen ondersteunt een bron blijkbaar de kandidaat in het beantwoorden van de vraag.

**Tabel 6: Vragen met en zonder bronnen, p-waarden.**

	regulier	pilot	regulier	pilot
	met bronnen		zonder bronnen	
2006	65,7	58,7	58,9	51,9
2005	62,6		54,9	
2004	67,8		59,3	
2003	57,0		50,5	
2002	60,1		47,6	
gemiddeld	62,6		54,3	

Uitgesplitst naar de verschillende onderdelen van het pilotexamen was de moeilijkheidsgraad, de rit-waarde (een maat voor de differentiatie, onderscheid tussen goed en minder goed presterende leerlingen), de standaarddeviatie (idem) en de maximale score per onderdeel zoals in tabel 7 is weergegeven. De vragen over prehistorie, middeleeuwen en vroegmoderne tijd zijn beter gemaakt dan de vragen over de moderne tijd. Op de vragen over de middeleeuwen is door de kandidaten het hoogst gescoord.

**Tabel 7: Deelonderwerpen, pilotexamen.**

	p waarde	Rit	st.dev	max score
prehistorie	59,79	0,54	2,06	12
middeleeuwen	60,43	0,54	2,55	13
vroegmoderne tijd	58,47	0,42	3,48	26
moderne tijd	50,43	0,43	3,57	27
gemiddeld	56,22	0,34	8,45	78

Voor het pilotexamen is een aantal sjabloonvragen ontwikkeld, die in de voorbereiding door docenten en leerlingen gebruikt zijn. In het examen zijn die in min of meer herkenbare vorm teruggekomen. Deze vragen zijn voor de analyse van de examenresultaten beschouwd als specifieke 'nieuwe-stijlvragen', die typerend zijn voor het nieuwe examen. Ook de vragen waarbij van kandidaten werd gevraagd iets te illustreren met een zelfgekozen voorbeeld zijn als 'nieuwe-stijlvragen' geormerkt voor de analyse. In tabel 8 worden de toets-technische analysegegevens voor deze nieuwe-stijlvragen weergegeven.

**Tabel 8: Specifieke nieuwe stijl (pilot) vragen.**

vraag	p waarde	Rit	st.dev	max score
2	71	0,47	1,05	4
3	50	0,58	1,26	4
5	53	0,43	0,89	2
11	54	0,58	1,08	4
16	65	0,40	0,71	2
17	69	0,36	0,77	2
25	26	0,39	0,78	2
30	40	0,37	0,65	2
gemiddeld	55	0,45	3,34	22

Uit tabel 8 blijkt dat de moeilijkheidsgraad van deze vragen iets hoger lag dan het gemiddelde van het hele examen, maar dat deze afwijking tamelijk gering is. Vraag 3 was bijvoorbeeld een sjabloonvraag waarin met een eigen voorbeeld de ontwikkelingen in de Griekse stadstaten moest worden toegelicht. Omdat hiermee goed geoefend had kunnen worden, zou dit moeten resulteren in een hoge p-waarde. Dat blijkt niet zo te zijn: de p-waarde van vraag 3 duidt eerder op een wat moeilijkere dan op een erg makkelijke vraag. De rit-waarde van deze vragen is wel aanzienlijk hoger dan gemiddeld. Dat duidt erop dat deze vragen sterker differentiëren.

Uit de toetstechnische gegevens kan een waarde 'alpha' worden afgeleid voor de betrouwbaarheid van een toets. Het gaat hierbij om de mate waarin men staat kan maken op meetresultaten, dat wil zeggen de mate waarin de scores consistent, nauwkeurig en reproduceerbaar zijn, kortom: vrij van meetfouten. Gestreefd wordt naar een betrouwbaarheidsscore van 0,75. Bij die waarde wordt de toets als een redelijk betrouwbaar meetinstrument gezien. In tabel 9 wordt de betrouwbaarheidsscore van het pilotexamen vergeleken met die van reguliere examens. Uit deze tabel blijkt dat de beoogde betrouwbaarheid van de geschiedenisexamens niet gehaald wordt, maar dat het pilotexamen in dit opzicht iets beter voldoet dan het reguliere examen.

**Tabel 9: Betrouwbaarheid, regulier en pilot.**

jaar	regulier	pilot
2006	0,70	0,72
2005	0,67	
2004	0,67	
2003	0,70	
2002	0,76	
gemiddeld	0,70	

Een belangrijk probleem bij de bestaande geschiedenisexamens vormt de geringe differentiatie: de mate waarin goede leerlingen goed scoren en de minder bekwame kandidaten duidelijk mindere resultaten boeken. De standaarddeviatie is daarvoor een maat: hoe groter de differentiatie, hoe groter dit getal wordt. De afgelopen jaren is deze spreiding telkens kleiner geworden. Dat hangt samen met de lengte van de score-schaal. Die is verminderd: het examen van 2003 telde 84 scorepunten, het pilotexamen 78 en het reguliere havo-examen 77. Uit de eerste reacties van de docenten van de pilotscholen viel op te maken dat zij de indruk hadden dat de differentiatie bij het pilotexamen groter was dan die bij het reguliere examen. Die indruk wordt door de gegevens in tabel 10 echter niet bevestigd.

**Tabel 10: Differentiatie (standaarddeviatie) regulier en pilot.**

jaar	regulier	pilot
2006	8,78	8,45
2005	8,98	
2004	8,43	
2003	9,26	
2002	10,46	
gemiddeld	9,18	

#### **3.1.4. Samenvatting**

De constructie van het geschiedenisexamen nieuwe stijl is zoveel mogelijk georganiseerd en verlopen zoals bij het reguliere examen gebruikelijk is. Er is gewerkt met een constructiegroep waarin ervaring met vraagconstructie voor het reguliere examen ruim voorhanden was. Uiteraard had het merendeel van de constructiegroepleden, net als bij het reguliere examen, ervaring met het lesgeven volgens het examenprogramma waarvoor de vragen werden geconstrueerd. Bij de spreiding van de vragen over de tien tijdvakken en de 49 kenmerkende aspecten is uitgegaan van een toetsmatrix zoals die ook voor het reguliere examen wordt opgesteld. De relatieve vulling van de tijdvakken komt terug in het aantal vragen en scorepunten dat per tijdvak 'verdiend' kan worden.

De toetsing van oriëntatiekennis op centraal niveau leverde in de vraagconstructie problemen op met betrekking tot de validiteit op meerdere terreinen. Reproductieve vragen konden niet worden gesteld, omdat

de oriëntatiekennis geen concrete gebeurtenissen, personen of jaartallen bevat. Doorsnedevragen bleken moeilijk te stellen; vragen waarbij de historische benaderingswijzen van domein A worden gekoppeld aan de kenmerkende aspecten zijn nog maar in beperkte mate ontwikkeld.

Door het ontwikkelen en verspreiden van sjabloonvragen werden docenten, kandidaten en vragenmakers in staat gesteld te oefenen met het opstellen van andersoortige vragen.

De toetsing van historisch denken en redeneren leverde niet veel minder of andere problemen op dan bij de bestaande examens. Over de manier waarop dat op een herkenbare en erkende manier gebeurt, blijven grote verschillen van mening bestaan.

De toetstechnische gegevens die na afloop van het examen verzameld zijn, laten zien dat het nieuwe stijl-examen op sommige onderdelen afwijkt van de reguliere havo-examens, maar dat de verschillen in het algemeen minder groot zijn dan wellicht verwacht werd. De moeilijkheidsgraad was hoger, de betrouwbaarheid was groter, en de differentiatie iets kleiner dan bij het reguliere examen.

### **3.2. Evaluatie van het examen door docenten en leerlingen**

In hoeverre is het centraal schriftelijk eindexamen geschiedenis nieuwe stijl een haalbaar en aanvaardbaar examen? Om deze vraag te beantwoorden is zowel docenten als leerlingen gevraagd of het examen voldeed aan hun verwachtingen en of ze het een haalbaar en eerlijk examen vonden. Verder is de docenten gevraagd of het examen in hun ogen voldeed aan de eindtermen en in hoeverre het voldoende betrouwbaar te corrigeren was. Tot slot vroegen we docenten of zij landelijke implementatie van het nieuwe examen wenselijk en haalbaar vinden voor havo en vwo.

#### **3.2.1 Was het examen zoals verwacht?**

##### **Docenten**

Vooraf leefde bij de meeste docenten van de pilotscholen onzekerheid over hoe de leerstof in het eindexamen bevraagd zou gaan worden. Het proefexamen maakte hierin al wel veel duidelijk, maar kwam vrij laat in het examenjaar. Het kwam voor de meeste docenten wel overeen met wat ze verwacht hadden. Deze herkenning lag er met name in dat de docenten de vraagtypen herkenden en ook geprobeerd hadden om dit soort vragen zelf te maken voor de schoolexamens. Sommigen hadden de vragen wat dichter bij de kenmerkende aspecten verwacht, of vonden sommige vragen wat onduidelijk. Naar aanleiding van het proefexamen wilden de meeste docenten nog extra uitleg geven over een aantal kenmerkende aspecten en vraagvormen uitleggen en oefenen.

Het CE zelf was voor de meeste docenten zoals verwacht. Docenten van drie scholen hadden wel verwacht dat de laatste twee tijdvakken meer bevraagd zouden worden in het examen. Op een andere school vond men de tijdvakken 5, 6 en 7 mager terugkomen in het examen. Ook was onduidelijk of leerlingen nu precies moesten weten wat 'representatief' is. Een docent merkte op dat er helaas geen vragen werden gesteld door de tijdvakken heen, terwijl dit in de ogen van deze docent wel belangrijk is.

De leerlingen scoorden over het algemeen zoals de docenten verwacht hadden. Wel scoorden leerlingen op een afzonderlijke school (of klas) soms wat beter of slechter dan verwacht. Op één school vielen de scores van twee klassen erg tegen. Ook ten aanzien van individuele leerlingen kwamen de verwachtingen in enkele gevallen niet helemaal overeen met de behaalde score.

### Leerlingen

Uit de vragenlijst die de leerlingen na afloop van het examen hebben ingevuld<sup>15</sup>, blijkt dat de meeste leerlingen uit de pilotgroep het examen 'goed te maken' (45%) of 'moeilijk' (45%) vonden (zie tabel 11). De vragen waren zoals ze verwacht hadden en zowel de vragen als het bronnenmateriaal werden over het algemeen duidelijk gevonden. Toch geldt dit algemene beeld niet voor alle leerlingen. Ruim een kwart van de leerlingen (26%) gaf aan moeite te hebben met het soort vragen en deze niet duidelijk te vinden. Als we deze uitkomsten vergelijken met de antwoorden van de leerlingen die het reguliere examen gemaakt hebben, blijkt dat de leerlingen van het reguliere examen de examenvragen duidelijker vonden en het meer eens waren met de stelling dat de belangrijkste onderdelen van de te bestuderen stof bevroegd werden. Dit kan te maken hebben met het feit dat zowel leerlingen als docenten in de pilotgroep lange tijd weinig zicht hadden op wat ze op het examen konden verwachten. De leerlingen uit de pilotscholen vonden in vergelijking met de leerlingen die het reguliere examen maakten de examenvragen meer op elkaar lijken. De leerlingen uit beide groepen verschilden niet in hun antwoorden op de vraag of ze op dit examen konden laten zien wat ze aan historische kennis, inzicht en vaardigheden hebben.

Tabel 11. Mening van leerlingen die het pilotexamen en leerlingen die het reguliere examen hebben gemaakt over het examen.

Items uit de vragenlijst	Pilot examen			Regulier examen		
	N	Gem.	St.dev.	N	Gem.	St.dev.
Ik vond dit examen [1 = te moeilijk – 5 = te makkelijk]	547	2.43	.62	293	2.48	.60
De vragen waren zoals ik ze had verwacht [1 = zeer meeneens – 5 zeer meeeens]	538	3.01	.88	286	3.11	.89
De belangrijkste onderdelen van de te bestuderen stof werden in dit examen bevroegd	538	3.08	.89	286	3.29*	.90
De vragen van dit examen waren duidelijk	538	2.96	.94	286	3.31*	.90
De vragen van dit examen leken erg op elkaar	538	2.70*	.88	286	2.57	.91
Het bronnenmateriaal bij dit examen was duidelijk	538	3.79	.75	286	3.73	.88
Ik heb op dit examen goed kunnen laten zien wat ik aan historische kennis heb	538	3.03	.94	286	3.05	.89
Ik heb op dit examen goed kunnen laten zien wat ik aan historisch inzicht heb	538	3.33	.78	286	3.24	.80

N = aantal leerlingen dat de vraag heeft ingevuld

\* significant verschil tussen pilot en regulier ( $p < .05$ ).

<sup>15</sup> Deze vragenlijst is opgenomen in bijlage 8.

Het beeld dat uit de vragenlijst naar voren komt, werd bevestigd in de leerlingenbijeenkomst. Van de 21 aanwezige leerlingen vonden 7 leerlingen het examen moeilijk en 14 leerlingen vonden het wel goed te doen, maar ook niet makkelijk. Het examen was volgens de leerlingen erg op inzicht gericht. Sommigen vonden dit prettig en niet zo moeilijk ('ik heb een goed overzicht gekregen en kan een antwoord beredeneren als ik het niet weet'). Anderen hadden hier wel moeite mee en zouden graag wat meer leervragen willen. Een aantal leerlingen gaf aan dat het soort vragen voor hen nieuw was. Dit punt werd niet door iedereen gedeeld en bleek erg te verschillen per school. Vragen waarin een voorbeeld genoemd moest worden vond men lastig en zorgden voor twijfel: is het voorbeeld dat ik noem een goed voorbeeld? Hoe ver, hoe diep moet je gaan met het beantwoorden van de vraag? De leerlingen die het examen goed te doen vonden, gaven ook aan dat je met goed lezen en logisch nadenken een heel eind komt. Ook vonden ze dat het examen 'niet heel diep' ging. De vragen waren volgens deze leerlingen vrij breed en oppervlakkig en daardoor de antwoorden ook.

Ten aanzien van de voorbereiding op het examen vonden de leerlingen uit de pilotscholen dat de lessen over het algemeen goed aansloten op het examen (50% eens, 29% neutraal, en 18% oneens). De groep leerlingen die het reguliere examen had gemaakt scoorde hoger op deze vraag en vond de aansluiting dus beter (zie tabel 12). Zowel de pilotleerlingen als de leerlingen van scholen met het reguliere examenprogramma vonden dat ze 'veel' stof te bestuderen hadden. De leerlingen die het reguliere examen maakten, hebben in de week voorafgaand aan het examen gemiddeld meer tijd besteed aan de voorbereiding dan de pilotgroep: 12 uur ten opzichte van 9 uur. Dit verschil sluit aan bij de opmerking van de pilotdocenten dat de score op het nieuwe examen niet zozeer bepaald wordt door het leren van de leerling vlak voor het examen, maar vooral door de inspanning die de leerling gedurende de twee jaren van examenvoorbereiding heeft geleverd. De voorbereiding verliep ook anders. Waar de leerlingen van de pilotscholen zich met name richtten op de aantekeningen van de uitleg van de docent, richtten de leerlingen die het reguliere examen maakten zich met name op de teksten uit het examenkatern en de proefexamens. De leerlingen van de pilotscholen gaven vaker aan zich op een andere manier voor te bereiden. In de leerlingenbijeenkomst werden onder andere genoemd: het boekje 'Geschiedenis in je pocket', de strip 'Van nul tot nu', de CD-rom 'Tijdvakken concreet' (van het IVGD) en de lijst met kenmerkende aspecten en begrippen.

Tabel 12. Mening van leerlingen die het pilotexamen en leerlingen die het reguliere examen hebben gemaakt over de voorbereiding op het examen.

Items uit de vragenlijst	Pilot examen			Regulier examen		
	N	Gem.	St.dev.	N	Gem.	St.dev.
Het examen sloot goed aan bij wat we in de lessen behandeld hebben [1 = zeer mee oneens – 5 zeer mee eens]	532	3.35	.89	277	3.59*	.90
Ik vond de hoeveelheid stof die we voor dit examen moesten bestuderen [1 = te veel – 5 = te weinig]	547	2.03	.76	292	1.96	.66
Waar heb je je bij het leren voor dit examen vooral op gericht? Kruis de twee belangrijkste aan:						
-Mijn aantekeningen van de uitleg van de docent	561	.51*	.50	299	.38	.49
-De opdrachten die ik over de stof gemaakt heb	561	.09	.28	299	.08	.27
-De samenvattingen die ik van de stof gemaakt heb	561	.48	.50	299	.42	.50
-De teksten (uit een schoolboek/katern) die de stof uitleggen	561	.31	.46	299	.59*	.49
-Een of meer proefexamens die ik heb gemaakt	561	.14	.35	299	.19*	.39
-Anders..	561	.20*	.40	299	.13	.33

N = aantal leerlingen dat de vraag heeft ingevuld

\* significant verschil tussen pilot en regulier ( $p < .05$ ).

### 3.2.2 Voldeed het examen aan de eindtermen CHMV?

Op de vraag wat dit examen vooral toetst, gaven de geïnterviewde docenten van de pilotscholen verschillende antwoorden. Enkelen noemden 'de kenmerkende aspecten in de tijdvakken', anderen noemden inzicht, of de vaardigheid van redeneren en in de context van de tijd plaatsen. De docenten verschilden van mening over de vraag of het examen meer inzicht of kennis toetst. Waar de ene docent vond dat er te veel inzichtvragen opgenomen waren en dat er wat meer leervragen in mochten, vond de andere docent dat de vragen voornamelijk op een meer feitelijk niveau gesteld werden en dat er meer inzichtvragen gesteld mochten worden.

De docenten vonden het prettig om van tevoren de verdeling van de vragen over de tijdvakken te weten. Wel bestond de indruk dat sommige tijdvakken, met name de laatste twee, weinig bevroegd werden. Bij sommige vragen twijfelde men of iets, gezien de omschrijving van het kenmerkend aspect, bekend kon worden verondersteld, zoals de scheiding tussen kerk en staat (vraag 10), voorbeelden van politieke emancipatie van vrouwen naast kiesrecht (vraag 25) en of een bron representatief is (vraag 29). De variatie in vragen vond men ten opzichte van het reguliere examen niet zo groot, maar dit vond men geen probleem. Ook merkte een



docent op dat de vragen op zichzelf niet zo uitdagend waren voor de leerlingen.

Volgens de meeste docenten kwam het examen overeen met de bedoelingen zoals beschreven in het rapport van de CHMV. Het examen toetste overzichtskennis. Bij enkele docenten bestond er twijfel ten aanzien van de mate waarin het examen nu echt historisch besef toetste en in hoeverre dit door het nieuwe examenprogramma ontwikkeld wordt.

### 3.2.3 Was het examen deugdelijk, haalbaar en eerlijk?

#### Docenten

Over het algemeen zijn de docenten van de pilotscholen tevreden over het examen. Ze vonden het examen goed te maken en het was volgens een docent 'echt geschiedenis en niet alleen maar feiten leren'. Volgens de docenten lijkt het examen op het eerste gezicht eenvoudig voor docenten en experts, maar is het dat niet voor zestienjarigen. Het is voor hen moeilijk om tot overzichtskennis te komen. Ook is het lastig voor leerlingen dat historisch begrip steeds gekoppeld wordt aan kennis. Een docent merkt op: 'Leerlingen hebben veel meegekregen aan overzichtskennis. Bij veel antwoorden van leerlingen voelde ik toch wel enige trots.' Op één school vond men het examen eigenlijk te makkelijk. Het niveau dat bevroegd werd is volgens deze docenten het niveau dat ze op het einde van de onderbouw zouden moeten beheersen. Het examen mag voor deze docenten dan ook moeilijker, met name ten aanzien van het schoolse historisch redeneren.

Over het algemeen vinden de docenten het een eerlijk examen. De goede leerlingen halen een goed cijfer, de slechte een minder goed cijfer. Als je twee jaar lang de zaken goed bijhoudt, kom je volgens de docenten een heel eind. Maar ook met een goede algemene ontwikkeling kom je volgens een aantal docenten ver. Eén docent merkte op dat het examen voor leerlingen die van de mavo/vmbo-t komen en over minder algemene ontwikkeling beschikken wel erg lastig is. Op twee scholen vond men dat het examen te weinig differentieerde.

We vroegen de docenten van de pilotscholen of ze denken dat het nieuwe programma geschikt is om jaar in jaar uit centrale examens over te maken. De docenten vinden over het algemeen dat het examenprogramma voldoende ruimte biedt om er langere tijd les over te geven.

Er is stof genoeg. Wel moeten inhoud en werkvormen steeds veranderen. De valkuil is dat het een trucje wordt en dat je ieder jaar hetzelfde doet. De leerlingen kunnen dan veel van elkaar kopiëren.

Je moet volgens een aantal docenten accenten verleggen en het programma creatief invullen. De kenmerkende aspecten bieden hiervoor voldoende ruimte. Ten aanzien van het examen zijn er meer twijfels.

Eenzijds merken enkele docenten op dat er stof en bronnen genoeg zijn om examens mee te maken. Ook kun je dezelfde bronnen op verschillende manieren gebruiken. Op een aantal scholen heeft men zelf vragen gemaakt en dat ging goed. Andere docenten zien problemen ten aanzien van de verscheidenheid van vragen die gemaakt kunnen worden. Voor sommige perioden zijn bronnen schaars (een docent merkt op dat voor hem niet per se alle tijdvakken in alle examens terug hoeven te komen). Bij sommige docenten leeft de angst dat het examen een trucje zal worden. Er zou volgens een aantal docenten meer variatie in het type vragen mogen komen. De examens zouden zich dan meer moeten richten op het toetsen van historisch redeneren. De meeste docenten zouden het jammer vinden als de stof voor het examenprogramma toch meer gespecificeerd zou worden. Daarmee beperk je volgens hen de vrijheid die het programma nu heeft.

### Leerlingen

Tijdens de leerlingenbijeenkomst werd de volgende stelling aan de leerlingen voorgelegd: Als je goed bent in geschiedenis haal je voor dit examen een goed cijfer, als je slecht bent een slecht cijfer. Tien leerlingen waren het met de stelling eens, zeven oneens en vier twijfelden. Discussie over de stelling leidde tot de vraag wanneer je goed bent in geschiedenis en wanneer je voor dit examen een goed cijfer haalt. Volgens de meesten moet je om een goed cijfer te halen, twee jaar lang alles goed bijhouden en inzet tonen. 'De stof is breed en algemeen en op een gegeven moment ben je er zo mee bezig geweest dat je het wel weet en herkent.' Inzet tonen en alles bijhouden hangt volgens leerlingen ook samen met motivatie. Als je goed bent in geschiedenis, vind je het vak leuker en leer je het daardoor makkelijker. Een tweede belangrijk punt dat genoemd werd is inzicht hebben. Met feiten leren alleen kom je er niet. 'Het examen ging om inzicht en herkenning, als je goed bent in geschiedenis heb je dat wel.' Als derde punt werd genoemd het goed kunnen lezen van vragen en bronnen.

Er werden ook argumenten tegen de stelling gegeven. Een leerling die tegen de stelling was drukte het als volgt uit: 'Als je slecht bent in geschiedenis kun je dankzij de oppervlakkigheid van het examen toch een voldoende halen. Het ligt dus aan je voorbereiding.' Als argument tegen werd ook genoemd dat de vragen op het examen zo breed waren, dat het moeilijk gevonden werd om er gericht op te antwoorden. 'Ik heb mijn kennis niet goed kunnen toepassen terwijl ik het wel allemaal begreep.'

#### 3.2.4 Correctie van het examen

In hoeverre was het examen voldoende betrouwbaar te corrigeren? Volgens een aantal docenten was dit examen prettiger en sneller na te kijken dan het reguliere examen waarin twee thema's worden bevroegd. Over het algemeen vonden de pilotdocenten dat de kenmerkende aspecten voldoende houvast bieden bij het corrigeren. Aan de andere kant werd opgemerkt dat het examen veel meer vage antwoorden oplevert, die het nakijkwerk juist lastiger maken. Een docent merkte op dat het examen voor het tweede tijdvak lastiger na te kijken was, mede door het ontbreken van een normbespreking.

Docenten hebben bij het corrigeren van de examens in een tabel bijgehouden welke vragen zij lastig te corrigeren vonden. Driekwart van de vragen leverde bij geen of slechts een of twee docenten problemen op bij de correctie. Er waren een paar vragen (een kwart van het totaal) die veel docenten lastig te corrigeren vonden. Ten eerste vraag 11, een zogenoemde 'stapelvraag' over de uitspraak van Willem van Oranje. Er was een grote variatie aan antwoorden en veel leerlingen gaven vage antwoorden. Ten tweede vraag 15, de vraag over de betrouwbaarheid van de Sonttolregisters. Het was voor docenten niet duidelijk hoe precies het antwoord moest zijn. Veel leerlingen wisten niet dat Sonttolregisters een overheidsdocument zijn. Ten derde vraag 29 over de representativiteit van de mening van Stalin en Churchill. Veel leerlingen gaven vage antwoorden of refereerden aan eigen kennis met betrekking tot de periode. Over het algemeen waren er geen problemen met de tweede corrector. De normbespreking heeft hier waarschijnlijk een belangrijk aandeel in gehad.

De meeste docenten zijn van mening dat de correctie geen probleem zal vormen als het programma landelijk ingevoerd wordt. Soms zal er wat discussie zijn, maar dat is volgens de meeste docenten van de pilotscholen niet anders dan bij het reguliere examen. Er is een leidraad van wat goede antwoorden zijn en de verwachting is dat leerlingen met standaardantwoorden zullen komen. Daarnaast speelt de argumentatie een belangrijke rol. Docenten van één school zien de correctie echter wel als problematisch.

### 3.2.5 Landelijke implementatie

De docenten van de pilotscholen zijn van mening dat het examenprogramma zoals zij dat hebben uitgetest landelijk kan worden ingevoerd. Een aantal docenten verbindt daar wel de voorwaarde aan dat de lijst met kenmerkende aspecten nog eens kritisch wordt bekeken en daar waar mogelijk verbeterd (in de zin van specifiek, duidelijker en/of meer consistent).

Onder andere omdat de docenten met name in de beginfase van de pilot onzeker waren en een aantal docenten zijn werkwijze gedurende de pilot heeft aangepast, benadrukken de deelnemers van de pilot dat er voldoende geïnvesteerd moet worden in een zorgvuldige implementatie. Docenten moeten ondersteund worden door middel van voorlichtingsbijeenkomsten waarbij de pilotscholen een rol hebben, maar ook brochures, conferenties en nascholingsbijeenkomsten.

De docenten van de pilotscholen is naar hun mening gevraagd over de geschiktheid van het examenprogramma voor vwo. De docenten zijn van mening dat het examenprogramma en het examen voor vwo meer diepgang moeten hebben. Er worden verschillende voorbeelden gegeven van manieren waarop die diepgang bereikt kan worden: meer vergelijkingen maken tussen bronnen, meer vergelijkingen maken tussen perioden/tijdvakken, meer aandacht voor diachronische ontwikkelingen (ontwikkelingen over tijdvakken heen), meer causale verbanden leggen en meer werken met vakoverstijgende opdrachten. Leden van de constructiegroep twijfelen echter of ze op basis van het examenprogramma examens kunnen maken die het voor vwo verwachte niveau zullen hebben (zie 3.1.5).

### 3.3. Evaluatie van het examen door experts

In hoeverre wordt het curriculum en het examen zinvol en uitvoerbaar geacht door 'de buitenwacht'? Er zijn drie groepen die hier onder 'de buitenwacht' worden verstaan: geschiedenisdocenten die niet deelnemen aan de pilot, vakdidactici geschiedenis van de eerstegraads lerarenopleidingen en experts in binnen- en buitenland die als historicus werkzaam zijn en betrokken zijn of zijn geweest bij de ontwikkeling van geschiedenisonderwijs in het voortgezet onderwijs.

#### 3.3.1 Geschiedenisdocenten

In samenwerking met het bestuur van de VGN is een vragenlijst opgesteld waarin geschiedenisdocenten in Nederland hun mening over het nieuwe examenprogramma en het pilot-examen konden geven. De enquête is geplaatst op de website van de VGN en via het blad Kleio en e-mail onder de aandacht gebracht van de leden. De enquête is ingevuld door 72 docenten. Het VGN-bestuur heeft in september 2006 een bijeenkomst belegd om geschiedenisdocenten te informeren over de pilot. Deze bijeenkomst was ook bedoeld om leden te raadplegen ten behoeve van het advies dat het VGN-bestuur over de invoering van een nieuw examen zou gaan geven. Tijdens de bijeenkomst werden docenten geïnformeerd over de procedure van de pilot, de opzet van de evaluatie en de ervaringen van een aantal docenten die het nieuwe examenprogramma hebben geïmplementeerd.<sup>16</sup>

Uit de enquête en de bijeenkomst kwam een aantal vragen en kritiekpunten naar voren. Er was bij veel docenten onduidelijkheid over het onderscheid tussen de invoering van het nieuwe examenprogramma (die

<sup>16</sup> Het verslag van de docentenraadpleging en de uitkomsten van de enquête zijn te vinden op de website van de VGN ([www.vgnkleio.nl](http://www.vgnkleio.nl)).

per 2007 plaatsvindt) en de mogelijke alternatieven voor een centraal examen. Ruim de helft van de docenten die de enquête hebben ingevuld geeft aan het niet eens te zijn met de stelling dat het toetsen van oriëntatiekennis in het CE een verstandige keuze is. Toch scoort bij de vraag wat het examen nieuwe stijl vooral zou moeten meten 'het in de tijd kunnen plaatsen van historische gebeurtenissen, verschijnselen en ontwikkelingen' hoog, naast 'het kunnen geven van verklaringen voor historische gebeurtenissen, verschijnselen en ontwikkelingen'. Tijdens de bijeenkomst kwam ook een verschil van mening naar voren ten aanzien van de stof die bevestigd moet worden op het CE. Kies je voor een globaal overzicht ten dienste van de algemene ontwikkeling van leerlingen of voor nadruk op 'grote thema's' in de geschiedenis van de laatste honderd jaar om leerlingen meer zicht te geven op de wereld van nu? Veel respondenten van de enquête denken dat de kenmerkende aspecten onvoldoende houvast bieden voor een goede voorbereiding op het examen. Deze twijfel bestond aanvankelijk ook bij de docenten van de pilotscholen, maar is na twee jaar grotendeels verdwenen. Er is twijfel ten aanzien van de geschiktheid van het nieuwe examenprogramma voor vwo. Ruim een derde van de docenten die de enquête invulden blijkt vóór invoering van het examen nieuwe stijl. Ruim een derde is tegen en verkiest diachrone thema's. Een kwart van de docenten is van mening dat er een langere pilotperiode moet zijn. Omdat een a-selecte steekproef uit de groep geschiedenisdocenten met examenklassen ontbreekt, kan moeilijk vastgesteld worden in hoeverre de uitkomsten van de enquête representatief zijn.

### 3.3.2. Vakdidactici

De mening van een zevental vakdidactici van eerstegraads lerarenopleidingen is weergegeven in een artikel van dr. S.R.E. Klein in *Kleio*.<sup>17</sup> Dit artikel behandelt de vraag of het pilotexamen inhoudelijk aanvaardbaar is. Klein komt tot de conclusie dat het examen niet aanvaardbaar is. Daarvoor worden drie argumenten aangevoerd. In de eerste plaats ontbreekt in het pilotexamen een vraag die diachronische ontwikkelingen - ontwikkelingen en verbanden over tijdvakken heen - toetst. Vanuit de vakdiscipline en het examenprogramma wordt dit inzicht echter wel verwacht. Leerlingen moeten in historische processen de samenhang tussen veranderingen en continuïteit kunnen beschrijven. In de tweede plaats impliceert het toetsen van oriëntatiekennis in het centraal examen volgens Klein dat er concessies worden gedaan aan de diepgang. Het stelsel van tien tijdvakken, bedoeld als een didactisch hulpmiddel, is ten onrechte verplichte examenstof geworden. Diverse vragen worden getypeerd door Klein als onderbouwvragen, bij bronnen worden volgens hem lange toelichtingen gegeven omdat examenmakers niet van contextkennis kunnen uitgaan en bij de toetsing van een belangrijk vakconcept als 'betrouwbaarheid' lijkt er door het ontbreken van specifieke contextkennis bij leerlingen van historisch denken geen sprake meer te zijn. In de derde plaats dreigen de contacten tussen wetenschap en schoolpraktijk te verwateren. Bij de ontwikkeling van de stofomschrijving en de screening van het concept-examen is geen wetenschappelijke specialist betrokken. Examenmakers bij het CITO kunnen zich niet op één thema concentreren. De examentraditie met wisselende thema's impliceerde vakinhoudelijke nascholing van vele docenten, waardoor nieuwe inzichten gemakkelijker hun weg konden vinden naar het klaslokaal. Klein opteert voor de toevoeging van een diachronisch thema, zodat ook het historisch proces toetsbaar is, vervlakking van het niveau kan worden voorkomen en vakspecifieke nascholing kan worden gecontinueerd.

### 3.3.3. Hoogleraren uit binnen- en buitenland

Het nieuwe examenprogramma en het eerste pilotexamen zijn voorgelegd aan een aantal historici: professor

<sup>17</sup> S. Klein, 'Het pilotexamen geschiedenis havo', in: *Kleio* 47 (2006) nr 6, p. 40-45.

W. Goegebeur (Vrije Universiteit, Brussel), professor P. de Rooij (Universiteit van Amsterdam) en professor M. Grever (Erasmus Universiteit Rotterdam). In een semi-gestructureerd interview hebben deze historici aan de hand van een aantal kernvragen hun mening over het nieuwe examenprogramma en het pilotexamen gegeven. Daarnaast hebben professor K. Barton (University of Cincinnati, Verenigde Staten) en professor S. Wineburg (Stanford University, Verenigde Staten) schriftelijk hun reactie gegeven op het pilotexamen.

Professor De Rooij was voorzitter van de Commissie Historische en Maatschappelijke Vorming. Het rapport van deze commissie was de basis voor het nieuwe examenprogramma. De Rooij is tevreden over het nieuwe examen. Het examen toetst, zoals voorgesteld in het rapport van de Commissie Historische en Maatschappelijke Vorming, een algemeen overzicht van de geschiedenis. Het gaat niet om zogenoemde schoolkennis van twee specialistische onderwerpen maar om het kunnen toepassen van overzichts-kennis. Volgens De Rooij oogt het examen hierdoor als van meer belang dan het huidige geschiedenisexamen. Het nieuwe examen biedt daarnaast ook continuïteit in de toetsing van historische benaderingswijzen zoals die in het oude programma bekend waren. Het examen heeft volgens hem voldoende niveau. Het examen vergt niet alleen kennis, maar ook beheersing, dat je er iets mee kan doen. Je kunt niet terugvallen op geleerde rijtjes en vaardigheden. Je moet inzicht hebben, dat is moeilijk voor veel kinderen. Het niveau is ook voldoende voor een goede aansluiting voor vervolgoopleidingen. Het algemene overzicht dat geschiedenisstudenten in de propedeuse krijgen zal volgens De Rooij een stuk makkelijker zijn met deze achtergrond. De Rooij acht het nieuwe examenprogramma ook geschikt voor vwo. Van een vwo-leerling verwacht hij meer uitgewerkte en subtielere antwoorden. Het examen zou ter bewaking van de inhoudelijke kwaliteit gescreend kunnen worden door universitaire docenten die in de bacheloropleiding geschiedenis een algemeen overzicht van de geschiedenis doceren.

Professor Goegebeur is kritisch over het nieuwe examenprogramma en het daarbij ontwikkelde centraal examen. Zijn kritiek betreft met name het kader van tien tijdvakken met kenmerkende aspecten. Dat kader doet volgens hem geen recht aan wat geschiedenis eigenlijk is. Geschiedenis bestaat uit continuïteit en breuken, terwijl in het examenprogramma verschijnselen en ontwikkelingen worden 'opgesloten' in tijdvakken die door middel van kunstmatige cesuren onderscheiden zijn. De kenmerkende aspecten besteden weinig aandacht aan specifieke historische ruimtes, aan wereldgeschiedenis en aan sociaal-economische ontwikkelingen. Professor Goegebeur is voor een beter evenwicht tussen een verdiepende thematische en een overzichtsbenadering. Dit zou kunnen door het toevoegen van een aantal tijdvakoverschrijdende probleemstellingen. Hij vraagt zich af welke criteria de keuze van de in het examenprogramma opgenomen vaardigheden hebben bepaald. Hoewel professor Goegebeur zich kan vinden in een deel van de examenvragen, heeft hij kritiek op de open formulering van de kenmerkende aspecten die volgens hem een centrale toetsing bemoeilijken.

Professor Grever was in 1994 en 1995 voorzitter van de werkgroep 'Geschiedenis in de Tweede Fase'. Zij is kritisch over zowel het nieuwe examenprogramma als het examen dat op grond van dat programma gemaakt is. Grever brengt twee fundamentele bezwaren naar voren. In de eerste plaats is ze van mening dat leerlingen van 5 havo en 6 vwo voor hun verdere studie en leven meer gebaat zijn bij een examen dat zich vooral richt op de Vroegmoderne en Moderne Tijd. Prehistorie, Oudheid en Middeleeuwen kunnen volgens haar wel in schoolexamens aan de orde komen. In de tweede plaats sluit het examen volgens haar niet aan bij wat het geschiedenisonderwijs leerlingen te bieden heeft en bij de complexiteit van het vak. Het examen komt op haar

'antiquarisch' over door het gebrek aan samenhang. Ze vindt dat het examen blijft steken bij het toetsen van het 'gereedschap' (het referentiekader van tien tijdvakken met kenmerkende aspecten) en niet het gebruiken van dat gereedschap centraal stelt. Het examen toetst naar haar mening te weinig of leerlingen kennis van de kenmerkende aspecten kunnen gebruiken om een redenering op te bouwen in relatie tot een relevante probleemstelling. Ze constateert dat het examen naast oppervlakkige en enigszins onnozele vragen ook vragen bevat die zich richten op kennis van verbanden en langetermijnontwikkelingen. Grever vindt het examen te weinig uitdagend voor havo-leerlingen en zeker voor vwo-leerlingen. Grever ziet meer in een examen dat uit twee delen bestaat. Een deel toetst oriëntatiekennis, begrip en historisch kunnen redeneren in een diachronisch en probleemgericht thema met een beknopte stofomschrijving. Het andere deel richt zich op kennis van kenmerkende aspecten, met de nadruk op besef van continuïteit en discontinuïteit en het beschrijven van verschillen en overeenkomsten.

Professor Barton zegt onder de indruk te zijn van 'the structure and focus of the exam', vooral omdat het zich ten doel stelt leerlingen te bevragen op de betekenis van historische inhoud, zonder te vervallen in simpele kennisvragen over feitelijke details. Het is een probleem als een toets alleen vaststelt in hoeverre leerlingen de details van bepaalde voorbeelden van historische inhoud hebben onthouden, aldus Barton. Een betekenisvol geschiedeniscurriculum concentreert zich op de oorzaken en gevolgen van historische ontwikkelingen en gebeurtenissen, de maatschappelijke context waarin samenlevingen in het verleden leefden en het proces van constructie van historische beelden. Binnen een dergelijk raamwerk illustreren specifieke personen en gebeurtenissen de 'big ideas' van de geschiedenis; het is minder belangrijk voor leerlingen om de voorbeelden te leren dan datgene te begrijpen waarvan zij voorbeelden zijn. De taak waarvoor de toetsontwerpers staan, zo gaat Barton verder, is dan een examen te ontwikkelen dat op een geschikte manier het begrip van leerlingen voor kritische onderdelen van historische inhoud in kaart brengt, zonder hen te beperken in het gebruik van bepaalde voorbeelden. Het experimentele examen doet dat heel goed, zowel door leerlingen te vragen eigen voorbeelden te geven om te illustreren wat ze geleerd hebben, als ook door voorbeelden aan hen voor te leggen en ze vervolgens niet te vragen nadere details te geven, maar te vragen hoe ze grotere historische thema's kunnen verduidelijken.

Professor Wineburg reageert op gelijksoortige wijze. Ook hij zegt onder de indruk te zijn van het materiaal: 'What impresses me most is that they [de examenopgaven] ask young people to think about history in context, rather than merely memorizing details that will be soon forgotten. Indeed, this type of examination could serve as a model for other national systems seeking to reform history instruction. The questions being asked prepare young people for the kinds of challenges that await them should they pursue history in the university, and this is a very positive development in high school assessment.'

## 4 | Vooruitzichten

Dit hoofdstuk beschrijft vooruitzichten in tweeledig opzicht: aan de ene kant wordt een blik vooruit geworpen op ontwikkelingen binnen de acht pilotscholen. De proefnemingen met een examen nieuwe stijl zullen op deze scholen in de jaren 2007 en 2008 worden voortgezet. Andere vooruitzichten betreffen beschouwingen over mogelijke ontwikkelingen in het geschiedenisonderwijs na invoering van examens nieuwe stijl. Deze beschouwingen kunnen uiteraard niet meer zijn dan veronderstellingen, zeker waar het gaat om de toekomst op langere termijn.

### 4.1. Ontwikkelingen binnen de pilotscholen

#### 4.1.1. De pilotdocenten en de scholen

Het laat zich aanzien dat de jaren van voortgezette proefneming geen exacte kopie zullen worden van het onderwijs aan de eerste pilotgeneratie. Van beslissende invloed hierop lijkt te zijn de omstandigheid dat de pilotdocenten een volledige cyclus hebben afgerond, inclusief afname en correctie van een examen. Nu de resultaten hiervan compleet zichtbaar geworden zijn, is de zekerheid onder de docenten toegenomen, zo gaf een aantal van hen te kennen aan het begin van de cursus 2006-2007. Deze zekerheid straalt ook weer af op de leerlingen. Docenten die weten wat ze kunnen verwachten en weten wat hen te doen staat, kunnen hun leerlingen op een andere manier motiveren dan docenten die zelf nog zoekend zijn. Dat is althans de mening van een belangrijk deel van de pilotdocenten.

Behalve deze houding is er ook de vergrote expertise die de docenten door hun ervaring hebben opgedaan. Op een aantal punten wordt het curriculum door hen aangepast. In hoofdstuk 2.2.1. werd reeds beschreven wat voornemens op dit punt zijn. Een aantal docenten geeft aan na te denken over de uitstralingseffecten van de pilot in de overige sectoren van het onderwijs: het vwo, en de onderbouw. Op scholen waar de onderbouw (grotendeels) door andere docenten bediend wordt dan de bovenbouw komt overleg tot stand over de wijze waarop leerlingen in de lagere leerjaren het best kunnen worden onderwezen in het perspectief van het nieuwe examen.

#### 4.1.2. Voornemens voor het vervolgproject

Het ligt voor de hand bij de uitvoering van het vervolgproject aandacht te schenken aan tijdens de eerste pilotcyclus gerezen problemen en geconstateerde tekortkomingen. Deze liggen vooral op het terrein van de constructie van (inhouds-)valide toetsen in relatie tot een globaal geformuleerd examenprogramma. De uitvoering van een onderwijscurriculum zal - gezien de opgedane ervaringen van de docenten en het beschikbaar komen van meer onderwijsmaterialen - vermoedelijk op steeds minder problemen stuiten.

De mogelijkheden voor het construeren van valide toetsen kunnen naar het zich laat aanzien worden verruimd op een aantal manieren:

- Door het formuleren van nadere toelichtingen op kenmerkende aspecten van tijdvakken, waardoor duidelijker wordt wat onder een globaal aangeduide inhoud kan worden verstaan, zonder deze op het niveau van feitelijke en concrete historische gegevens in te vullen of vast te leggen;
- Door het formuleren van nadere toelichtingen op de eindtermen 4 t/m 7 betreffende historisch denken en redeneren, waardoor duidelijker wordt uit welke componenten een historische redenering kan worden opgebouwd en de centrale toetsbaarheid ervan in het algemeen - ongeacht de relatie met een bepaalde historische inhoud - kan worden vergroot;

- Door voorbeeldmatig aan te duiden hoe ontwikkelingen door de tijd heen kunnen worden bestudeerd door het combineren van kenmerkende aspecten uit diverse tijdvakken, en het toepassen van algemene begrippen (uit de lijst van 91, zie bijlage 6) op meer dan één tijdvak;
- Door voorbeeldmatig aan te duiden hoe causale relaties rond de in kenmerkende aspecten aangeduide verschijnselen kunnen worden bestudeerd.

De hier aangeduide toelichtingen - en wellicht nog andere - kunnen de vorm krijgen van een 'syllabus' bij het eindexamenprogramma. Gezien de door de pilotdocenten duidelijk onderkende voordelen van een tot op grote hoogte vrij invulbaar programma, zowel voor hun eigen functioneren als voor het leren van de leerlingen, zou de syllabus niet het karakter moeten krijgen van een traditionele stofomschrijving.

Indien in het vervolgproject aan de nadere toelichtingen gewerkt wordt, zal het effect hiervan pas bij het examen van 2008 goed merkbaar zijn. De leerlingen die zich voorbereiden op het examen van 2007 zijn immers reeds in de cursus 2005-2006 in havo-4 begonnen en zitten op het moment van schrijven van dit rapport in havo-5. Voor hen is nog maar een beperkte mate van bijsturing mogelijk.

Een ander aandachtspunt voor het vervolg is het vwo-niveau. In de opzet van het vervolgproject is uitvoering van een vwo-pilot weliswaar niet opgenomen. Dat neemt echter niet weg dat zowel van binnen als buiten de pilot zorg is geuit over de wijze waarop een examen op vwo-niveau volgens het nieuwe model tot stand zou kunnen komen. Bij de ontwikkeling van de hierboven aangeduide nadere toelichtingen zou het onderscheid tussen een havo- en een vwo-niveau kunnen worden betrokken.

#### **4.2. Te verwachten effecten bij implementatie**

##### **4.2.1. De vraag naar scholing**

Uit het pilotproject is gebleken dat de overgang naar een nieuw examenregime grote aanpassingen vergt van docenten en leerlingen. Een probleem hierbij vormden de onzekerheid en onduidelijkheid. Deze problemen zullen bij landelijke implementatie minder groot zijn dan ze tijdens de pilot waren, omdat men kan beschikken over de ervaringen van de pilot. Na drie examenjaren zijn zes gebruikte examens beschikbaar, die - omdat zij steeds over een gelijklopend examenprogramma gaan - zeer goed als voorbeeld kunnen dienen. Leraren kunnen zich hierop oriënteren.

Dat neemt niet weg dat vooral duidelijkheid zou moeten worden gecreëerd om onzekerheid en daarmee gepaard gaande gevoelens van onvrede zoveel mogelijk voor te zijn. Duidelijkheid kan alleen ontstaan door het geven van informatie. De vraag naar scholing zal dan ook groot zijn. Het is op het moment van schrijven van dit rapport nog niet duidelijk hoe opleidingen, nascholingsinstituten en onderwijsondersteunende instellingen hierop zullen inspelen. Uitgaand van de gedachte dat de vraag sturend zal zijn voor het aanbod, moet verondersteld worden dat de nodige scholingsinitiatieven zullen worden ontwikkeld. Hoeveel extra ondersteuning en centrale regie daarbij wenselijk zal zijn, blijft een belangrijke vraag.

##### **4.2.2. De curricula van de scholen**

Bij implementatie van het nieuwe examenprogramma wordt voor het eerst een situatie gecreëerd waarin in de bovenbouw van het havo en vwo logisch wordt voortgebouwd op in de onderbouw aangebrachte historische kennis. Het gaat immers om dezelfde tien tijdvakken met grotendeels dezelfde kenmerken. Leerlingen die



hiermee vroegtijdig vertrouwd worden gemaakt, zullen een basis ontwikkelen voor de bovenbouw. Het ligt dan ook voor de hand te veronderstellen dat scholen bij hun inrichting van het onderbouwonderwijs met de nieuwe examensituatie rekening zullen houden. Enigszins tegenstrijdig hiermee is de ontwikkeling in de richting van globalisering van geschiedenisdoelen en integratie van geschiedenis in leergebieden, zoals die zich tegelijkertijd voordoet. Deze zijn vaak moeilijk te combineren met een meer op chronologie gebaseerde tijdvakken-aanpak. Het is moeilijk te voorspellen hoe deze twee tendenties in de praktijk in de onderbouw van de verschillende scholen zullen uitwerken.

Het onderwijs in de bovenbouw van havo en vwo zal ingrijpend worden aangepast. De pilot heeft geleerd dat daarin verschillende maten van aanpassing voorstelbaar zijn. Sommige docenten zullen geneigd zijn zich geheel te concentreren op 'het nieuwe', d.w.z. de tijdvakken en de overzichtskennis daarvan. Dat kan vertekeningen opleveren die geen recht doen aan het eindexamenprogramma. In dat programma is immers ook aandacht voor thema's en voor historisch denken en redeneren. Een mogelijkheid is dat docenten zullen ervaren dat - bij een wat andere invulling van de onderbouw - een raamwerk van tijdvakken reeds aanwezig is, en dat juist door toepassing van het raamwerk op thema's en andere voorbeelden een nadere invulling en verdieping kan worden bereikt. Waar mogelijk zou een dergelijke ontwikkeling moeten worden gestimuleerd.

Een belangrijk gegeven is het beschikbaar komen van nieuwe leerboeken. Op het moment van schrijven van dit rapport is duidelijk dat omstreeks de jaarwisseling van 2006-2007 tenminste drie educatieve uitgeverijen methoden publiceren die geënt zijn op het nieuwe eindexamenprogramma havo en vwo. De ervaring leert dat de beschikbaarheid van boeken een grote rol speelt bij de voorstellingen die docenten zich van een nieuw programma (kunnen) maken. Daarmee neemt het vertrouwen toe dat een nieuw curriculum uitvoerbaar zal zijn.

Ten slotte de ontwikkelingen in het vmbo. Wijzigingen in het vmbo-examen zijn pas op langere termijn aan de orde. Op korte termijn is reeds voorzien in de opname in het programma van een overzicht van de (lange) twintigste eeuw. Daarbij is gepoogd voor die eeuw een invulling te geven die in de richting gaat van de algemene oriënterende kennis van de laatste twee tijdvakken. Op welke wijze verdere aanpassingen in het vmbo-programma kunnen plaatsvinden, is op dit moment moeilijk te voorspellen.

#### **4.2.3. De relatie tot vakinhoudelijke wetenschappelijke expertise**

Bij de bestaande examens wordt contact met de ontwikkelingen in de wetenschap in stand gehouden middels scholingen door experts op de jaarlijks wisselende onderwerpen. De examens worden door inhoudelijke experts gescreend, wat niet zo moeilijk te organiseren is omdat ze over twee specialistische onderwerpen gaan. Voor deze contacten met de wetenschap zijn nog geen nieuwe vormen gevonden voor het examen nieuwe stijl. Het is echter wel van belang dat dit contact in stand blijft, zoals door een aantal critici reeds is opgemerkt.

Hoe een en ander zich zal ontwikkelen, is nog niet duidelijk. Voorstelbaar is dat nascholingen zich succesvoller zullen richten op tijdvakken, dan wel kenmerkende aspecten van tijdvakken, die elk achtereenvolgens door inhoudelijke experts van toelichting worden voorzien. Nascholingsinstituten zouden dergelijke cycli kunnen opzetten, waardoor het 'gehele programma' in een bepaald schema telkens opnieuw wordt doorgenomen. Een bijkomend voordeel van dergelijke scholingen zou zijn, dat de resultaten ervan langduriger toepasbaar blijven dan die voor de wisselende examenthema's, die immers slechts twee jaar in gebruik blijven.

Wat betreft de inhoudelijke screening van de examens zou de indeling in perioden een aanknopingspunt kunnen bieden. Specialisten in respectievelijk Prehistorie/Oudheid, Middeleeuwen, Vroegmoderne Tijd en Moderne Tijd zouden de desbetreffende vragen uit een examen kunnen controleren. Daarbij zou een beroep kunnen worden gedaan op universitaire docenten die in de bachelor-opleidingen de algemene introducties van perioden verzorgen.

# 5 | Conclusies

## 5.1. Beantwoording van de deelvragen

In hoofdstuk 1 werd een aantal deelvragen gepresenteerd die uit de hoofdvraag van de examenpilot werden afgeleid. In de loop van de voorgaande hoofdstukken zijn deze deelvragen successievelijk beantwoord. Hier worden de hoofdzaken van deze beantwoording kort samengevat.

1 Kan een valide examen worden ontwikkeld dat voldoet aan de eindtermen?

Kort samengevat luidt het antwoord: dat is mogelijk gebleken. De globale formulering van de eindtermen heeft echter wel tot moeilijkheden geleid. Deze zouden bij een nadere toelichting geheel of gedeeltelijk kunnen worden opgelost. Een belangrijk deel van de kritiek van binnen en buiten het pilotproject is op de beperking van de bewegingsvrijheid van de toetsconstructeurs, veroorzaakt door een zeer globale tekst, terug te voeren. Nadere toelichting op de eindtermen in een syllabus kan deze bewegingsvrijheid vergroten.

2 Kan een examen worden ontwikkeld dat voldoet aan de eisen van deugdelijke toetsing?

De toetstechnische analyse heeft aangetoond dat het pilotexamen in dit opzicht aan dezelfde eisen voldeed als de gebruikelijke geschiedenisexamens. Met inachtneming van de bovenstaande opmerkingen over de beperkingen van de inhoudsvaliditeit is een examen geconstrueerd dat valide, betrouwbaar en efficiënt was.

3 Kan een examen worden ontwikkeld dat voldoet aan de verwachtingen van docenten en leerlingen?

De gegevens uit de pilot duiden erop dat dit het geval was. Docenten en leerlingen herkenden in het examen in voldoende mate wat zij ervan hadden verwacht. Sommige docenten tekenen hierbij aan dat naar hun inzicht op grond van het examenprogramma méér mogelijk zou zijn geweest. De toetsdeskundigen zien hier echter grotere problemen.

4 Kan een examen worden ontwikkeld dat voldoende betrouwbaar te corrigeren is?

De ervaringen van de pilotdocenten wijzen erop dat het examen niet moeilijker te corrigeren was dan reguliere geschiedenisexamens. Bij geschiedenisexamens is verschillende interpretatie van het correctievoorschrift en verschil van mening tussen eerste en tweede corrector een vaak voorkomend verschijnsel. Bij het pilotexamen was dat niet anders. Sommige pilotdocenten vonden het examen gemakkelijker te corrigeren dan reguliere examens. De resultaten weerspiegelden in het algemeen hun verwachtingen op grond van de indrukken die ze over leerlingen hadden verzameld.

5 Kan een examen worden ontwikkeld dat in de ogen van docenten en leerlingen haalbaar en eerlijk is?

De uitgevoerde onderwijsprogramma's hebben naar het oordeel van de grote meerderheid van de docenten de leerlingen voldoende op het examen kunnen voorbereiden. In die zin ging het dus om een haalbaar programma, en - ook naar het oordeel van de leerlingen - een examen dat 'goed te doen' was. Docenten herkenden in de examenresultaten de verschillen tussen leerlingen die 'zich twee jaar lang hadden ingespannen' en leerlingen die dat niet hadden gedaan en op de valreep nog probeerden hun achterstand goed te maken. Deze laatste leerlingen werden door het pilotexamen niet beloond, en dat correspondeerde met het gevoel van eerlijkheid onder de docenten. Gegevens van de leerlingen wijzen er niet op dat ze het gevoel hadden door het examen onjuist of oneerlijk te zijn behandeld. De uitvoering van de pilot heeft in dit verband niet tot klachten geleid, ook niet van ouders.

6 Wat zijn de verschillen tussen een examen nieuwe stijl en de gebruikelijke examens?

In de vorm van het examen is gestreefd naar een zo klein mogelijk verschil. Er waren een bronnenboekje, een vragenboekje en een correctievoorschrift die er net zo uitzagen als bij reguliere examens. De formulering (taal) van de vragen sloot aan op wat bij reguliere examens gebruikelijk is. Toepassing van historisch redeneren (structuurbegrippen en vaardigheden) maakte net als bij het reguliere examen deel uit van de toetsing. Het verschil kan dan ook aldus worden samengevat: de inhoudelijke leerstof bestond niet uit twee gedetailleerd in een stofomschrijving beschreven thema's, maar uit globaal omschreven kennis betreffende tien tijdvakken.

7 Slagen scholen en docenten erin om een aanvaardbaar curriculum en een didactiek te realiseren waarmee leerlingen goed voorbereid worden op het beoogde examen?

Het algemene antwoord moet luiden dat zij daarin geslaagd zijn, waarbij vanwege de pilotsituatie gesproken moet worden van voortschrijdend inzicht en realiseren van curricula met vallen en opstaan. Het resultaat is op de overgrote meerderheid van de scholen uiteindelijk bevredigend geweest.

8 Hoe beoordelen experts het pilot-examen?

Onder de binnenlandse experts zijn er kritische geluiden en geluiden die tevredenheid tot uitdrukking brengen. Tevredenheid is er over de continuïteit in de toetsing van historische benaderingswijzen zoals die in het oude programma bekend waren. Over het niveau van het examen verschillen de meningen zeer, tussen 'voldoende' (niet alleen kennis, maar ook beheersing, je moet inzicht hebben) en 'beneden peil' (het zijn eigenlijk brugklas- of basisvormingsvragen). Een positief punt is de aansluiting op het vervolgonderwijs, waar studenten eerder zullen worden aangesproken op hun algemene referentiekader dan op specialistische themakennis. Zorg is er bij binnenlandse specialisten over de geringe aandacht voor ontwikkelingen door de tijdvakken heen en causale verbanden. Een opvatting is ook dat de grote aandacht voor het tijdvakkenkader - dat eigenlijk als een stuk gereedschap bedoeld is - een instrument verheft tot een doel in zichzelf.

Onder de buitenlandse experts zijn er positieve geluiden naast kritische. Kritiek is er op het tijdvakkenkader dat te weinig ruimte zou geven voor ontwikkelingen door de tijden heen en kenmerken teveel in hokjes zou opsluiten. Waardering is er daartegenover voor de wijze waarop van leerlingen gevraagd wordt betekenis te geven aan wat zij geleerd hebben, en niet slechts kennis te tonen van uit het hoofd geleerde voorbeelden. Een enkele buitenlandse expert meent dat dit examen een voorbeeld zou kunnen zijn voor andere landen waar nationale examens worden ontwikkeld.

## 5.2. Conclusies

Op grond van bovenstaande beantwoording van de deelvragen en de door de pilot bereikte resultaten in het algemeen is de volgende algemene eindconclusie te trekken:

Uit de pilot is gebleken dat een examen-havo op basis van het eindexamenprogramma CHMV mogelijk is. Landelijke invoering zal mogelijk zijn, mits goed begeleid en ondersteund.

Deze algemene eindconclusie kan nader worden toegelicht en verduidelijkt in een aantal deelconclusies:

1 Uit het onderzoek van de beginsituatie van de leerlingen is gebleken dat ze de algemene oriënterende kennis van tien tijdvakken in het najaar 2004 slechts in geringe mate beheersten. Naar de mening van de

pilotdocenten is het mogelijk gebleken deze kennis in de loop van twee jaar bij een grote meerderheid van de leerlingen op een aanvaardbaar peil te brengen.

Toelichting: De uitspraak over de beheersing is gebaseerd op enkele toetsen (meerkeuze- en associatief) die bij de aanvang van de pilot aan de leerlingen zijn voorgelegd. Dit houdt tevens een voorbehoud in. Deze toetsen konden maar in beperkte mate op betrouwbare wijze meten wat de leerlingen beheersten. De uitspraak is dan ook eerder een indicatie dan een echt hard gegeven. De uitspraak over het 'aanvaardbare' peil is gebaseerd op de meningen van de pilotdocenten en op de resultaten van het examen.

2 Een adequate voorbereiding van de leerlingen op het examen heeft kunnen plaatsvinden via verschillende didactische werkwijzen en invullingen van het curriculum.

Toelichting: Docenten hebben zelf hun aanpakken en curricula bepaald, en die waren, mede door het niet voorhanden zijn van leerboeken en het niet verstrekken van nadere uitleg van het examenprogramma, zeer verschillend. Uit de resultaten van de pilot is echter niet gebleken dat de ene aanpak significant betere resultaten opleverde dan de andere. Er waren wel niveauverschillen tussen de scholen, maar deze hingen niet samen met de toepassing van een bepaalde didactische strategie. Een voorzichtige conclusie zou dus kunnen zijn dat er bij implementatie van het nieuwe programma meerdere wegen zijn die naar Rome leiden.

3 Over de mate waarin verdere specificatie van de tekst van het examenprogramma gewenst is, bestaat verschil van mening onder de pilotdocenten. Er is kritiek op bepaalde zeer open geformuleerde 'kenmerkende aspecten' van tijdvakken. Aan de andere kant worden vrijheid en eigen verantwoordelijkheid op prijs gesteld.

Toelichting: De vrijheid om zelf concrete invullingen te kiezen is niet alleen uitdagend en aantrekkelijk voor de docenten, zij levert voor de leerlingen ook een andere wijze van leren op die volgens de pilotdocenten waardevolle kanten heeft. Er wordt meer gevraagd van inzicht en toepasbaarheid van kennis dan van de kennis op zichzelf. Daar tegenover staat het probleem van de toetsconstructeurs die moeite hebben met globale formuleringen.

4 De verschillende aard van formulering en detaillering in de tekst van de 'kenmerkende aspecten' levert problemen op voor een valide toetsconstructie.

Toelichting: Met name zeer globaal geformuleerde kenmerken als 'het voeren van twee wereldoorlogen' geven te weinig duidelijkheid over wat moet worden bestudeerd en wat kan worden gevraagd. Bij enkele andere kenmerken staan specifieke elementen er wel (bijvoorbeeld 'grondrechten en burgerrechten in Franse en Bataafse revolutie), waardoor ze gemakkelijker bevroegd kunnen worden. Deze ongelijke bedeling kan tot onevenwichtigheden leiden in de examens op oneigenlijke gronden.

5 De formulering van stof in alleen 'kenmerkende aspecten' van tijdvakken geeft weinig houvast voor toetsopgaven die doorsneden door de tijd, ontwikkelingen en causale relaties beschrijven.

Toelichting: Dit geldt met name als de formulering van kenmerkende aspecten als strikt behorend tot telkens één tijdvak, en geheel op zichzelf wordt beschouwd. Sommige pilotdocenten zijn echter van mening dat de aangeduide toetsopgaven wel degelijk mogelijk zouden moeten zijn, op grond van de toepassing van eindtermen 4 t/m 7 (historisch redeneren) op het inhoudelijke deel van het programma.

6 Pilotdocenten zijn in het algemeen van mening dat het vak geschiedenis met de nieuwe aanpak van het curriculum beter tot zijn recht komt dan met de gebruikelijke aanpak.

Toelichting: Deze conclusie is gebaseerd op veelvuldig voorkomende uitspraken van pilotdocenten in de zin van 'ik heb het gevoel dat ik weer echt geschiedenis geef'. Pilotdocenten gaven daarbij vaak voorbeelden van opvallend inzichtelijke opmerkingen en vragen van leerlingen, die voortkwamen uit het gebruik van het tijdvakkenkader als historisch referentiekader.

7 Leerlingen van de pilotscholen vinden net als leerlingen van niet-pilotscholen het geschiedenisonderwijs in havo-4 en havo-5 belangrijk en interessant. Wat dit laatste betreft scoren ze lager dan leerlingen van niet-pilotscholen.

8 Leerlingen van de pilotscholen zijn van mening dat ze met de nieuwe aanpak nuttige overzichtskennis hebben opgebouwd, die ze kunnen gebruiken en niet snel zullen vergeten. In dit opzicht scoren ze hoger dan leerlingen van niet-pilotscholen.

9 Leerlingen van pilotscholen vinden in mindere mate dan leerlingen van andere scholen dat ze kritisch hebben leren denken en de wereld van vandaag beter hebben leren begrijpen.

Toelichting: De conclusies 7, 8 en 9 zijn gebaseerd op het onderzoek middels de vragenlijst onder pilotleerlingen en niet-pilotleerlingen.

10 Uit toetstechnische gegevens blijkt dat de moeilijkheidsgraad van het pilotexamen hoger ligt dan die van de gebruikelijke examens. Deze moeilijkheidsgraad komt dichterbij de buurt van het streefniveau.

11 De toetstechnische betrouwbaarheid van het pilotexamen is ongeveer even groot als die van gebruikelijke examens.

Toelichting: De conclusies 10 en 11 zijn gebaseerd op de toetstechnische analyses door Cito.

12 Bij de correctie van het pilotexamen zijn geen grotere problemen ontstaan dan bij gebruikelijke examens.

Toelichting: Dit is gebleken tijdens de plenaire besprekingen onder de pilotdocenten na de eerste en tweede correctie van het examen. De conclusie is bovendien gebaseerd op de vragenlijsten die aan de pilotdocenten werden voorgelegd na de correctie van het examen, waarop zij per vraag moesten aangeven

welke moeilijkheden zich voordeden bij de correctie.

13 De docenten van de pilotscholen zijn van mening dat het examenprogramma zoals zij dat hebben uitgetest landelijk kan worden ingevoerd.

Toelichting: Deze opvatting is naar voren gekomen uit de gesprekken met de pilotdocenten na afloop van het examen, alsmede op een plenaire bijeenkomst met pilotdocenten waar deze conclusie werd bekrachtigd.

14 Van de kant van experts bestaat naast tevredenheid ook kritiek op het examen. Deze laatste richt zich met name op de geringe aandacht voor toetsopgaven die doorsneden door de tijd, ontwikkelingen en causale relaties beschrijven - of in het algemeen toetsopgaven die gaan over historisch denken en redeneren ('vaardigheden').

Toelichting: Deze conclusie is gebaseerd op gesprekken met experts en op door experts schriftelijk gegeven reacties.

15 Over de vraag hoe een nieuw examen op vwo-niveau kan worden geïmplementeerd bestaat nog onduidelijkheid. Onzekerheid hieromtrent wordt gevoed door in conclusie 14 genoemde redenen.

Toelichting: Omdat tijdens de pilot geen ervaring is opgedaan in het vwo, bestaat onzekerheid over de vraag in hoeverre de conclusies ook voor het vwo geldigheid hebben. Sommige experts zijn van mening dat de gesignaleerde moeilijkheden met toetsopgaven die doorsneden door de tijd, ontwikkelingen en causale relaties beschrijven - of in het algemeen toetsopgaven die gaan over historisch denken en redeneren - zich op vwo-examens met meer nadruk zullen voordoen.

### 5.3. Aanbevelingen

De conclusie dat een examen mogelijk is en dat landelijke implementatie niet op problemen zal stuiten houdt niet in dat een eindstadium in de ontwikkeling van een examen nieuwe stijl is bereikt. De gesignaleerde problemen zijn daarvoor nog te ernstig. Ook zal aandacht moeten worden besteed aan de wijze waarop het nieuwe programma op het vwo kan worden geïmplementeerd. Op grond van de kritiek van buiten en de ervaringen binnen het pilotproject doet de projectleiding - na overleg met de pilotdocenten - dan ook de volgende aanbevelingen:

1 Het verdient aanbeveling te bezien welke nadere uitwerking van of toelichting op de kenmerkende aspecten van tijdvakken gewenst is. Dit met behoud van de vrijheid van docenten om eigen concrete invullingen te kiezen.

2 Bezien moet worden hoe met een nadere toelichting op de examenstof diachronische ontwikkelingen en causale relaties meer accent kunnen krijgen.

3 Onderzocht moet worden in hoeverre op de eindtermen betreffende 'historisch denken en redeneren' (4 t/m 7) een nadere toelichting gewenst is.

Toelichting: De nadere toelichtingen bedoeld in aanbeveling 1, 2 en 3 hoeven niet te resulteren in een nadere specificering van de tekst van het examenprogramma zelf, maar kunnen de vorm krijgen van syllabi te ontwikkelen door onderwijs ondersteunende instituten.

- 4 Bij de verdere toetsontwikkeling ten behoeve van experimentele examens verdienen de aspecten van 'betekenis voor het heden' en 'historisch denken en redeneren' speciale aandacht.
- 5 Bij de verdere toetsontwikkeling moet worden gezien hoe recht kan worden gedaan aan het niveauverschil tussen havo en vwo.
- 6 De landelijke invoering van het nieuwe examen dient ondersteund te worden door een uitvoerig programma van voorlichting en nascholing.
- 7 De tijdens de uitvoering van de pilot opgedane ervaringskennis van docenten uit pilotscholen moet bij de implementatie worden benut voor voorlichting en scholing van andere docenten.
- 8 Landelijke examens gebaseerd op domein A en B van het nieuwe examenprogramma kunnen in het havo vanaf 2009 en in het vwo vanaf 2010 plaatsvinden, mits bovenstaande aanbevelingen worden opgevolgd en tijdig oplossingen voor gesignaleerde problemen worden gevonden.