

Uitslag

Examenverslag 2002

vbo/mavo/havo/vwo

R.V.J. Alberts
L. Noordermeer

Cito groep

Voortgezet onderwijs

Examenverslag

vbo/mavo/havo/vwo 2002

R.V.J. Alberts

L. Noordermeer

Citogroep

Arnhem, september 2002

Art. nr. 58266

© Citogroep Arnhem 2002. Auteursrecht voorbehouden.

Niets uit dit werk mag zonder voorafgaande schriftelijke toestemming van de Citogroep worden openbaar gemaakt en/of verveelvoudigd door middel van druk, fotokopie/reprografie, scanning, computersoftware of andere elektronische verveelvoudiging of openbaarmaking, microfilm, geluidskopie, film- of videokopie of op welke wijze dan ook.

Woord vooraf

Het examenverslag geeft u een indruk van de wijze van produceren van de examens en van de resultaten in 2002. In opdracht van de CEVO (Centrale Examencommissie Vaststelling Opgaven) produceert de Citogroep jaarlijks de examens. De evaluatie, waarvan dit verslag een onderdeel is, behoort ook bij deze activiteiten.

Examens krijgen elk jaar in de examenperiode veel aandacht. Dit jaar werd met extra belangstelling naar de examens en examenresultaten voor de vernieuwde tweede fase gekeken. Namen in 2001 al alle havo-leerlingen voor het eerst deel aan de tweede fase examens, dit jaar was dat het geval voor de vwo-leerlingen. Men is benieuwd naar de effecten van de veranderingen in het onderwijs. Maar er is nog een aanleiding voor de toegenomen belangstelling. De centrale examens gaan steeds meer een ijkpunt vormen, niet alleen voor de niveaubepaling en certificering van de leerling, maar ook voor de resultaten van de school. Omdat examenresultaten onderdeel van de kwaliteitskaart uitmaken, wordt ook buiten de examenperiode over examens en examenresultaten van scholen gesproken. Centrale examens maken het mogelijk vergelijkingen met landelijke gemiddelden en vergelijkingen tussen scholen onderling te maken. Daarbij dient echter te worden opgemerkt dat centrale examens zich beperken tot een deel van het examenprogramma. Alleen voor dat deel vormen de examens een ijkpunt.

Bij de samenstelling van de examens is, naast het leerstofinhoudelijke aspect, de moeilijkheidsgraad steeds van groot belang. Gelijkwaardigheid qua moeilijkheidsgraad is een van de billijkheidseisen waaraan centrale examens dienen te voldoen. Procedures van normvergelijking, waarvan in dit verslag ook gegevens opgenomen zijn, dragen ertoe bij dat de centrale examens een echt ijkpunt kunnen zijn en maken het mogelijk om prestaties van een gelijk niveau over verschillende jaren en tijdvakken heen met een zelfde cijfer te waarderen.

Het accent op de inhoudelijke ontwikkeling en de aandacht voor de niveaubepaling heeft een belangrijke rol gespeeld bij de ontwikkeling van de nieuwe examens voor de Tweede Fase havo/vwo.

De examens voor vbo/mavo C- en D-niveau verkeerden dit jaar wat de ontwikkelingen betreft nog in een rustiger vaarwater. De examens voor het vmbo gemengde en theoretische leerweg zullen in 2003 voor het eerst worden afgenomen.

Wij hopen dat dit verslag u een goed beeld geeft van het verloop en de resultaten van de centrale examens 2002.

De gegevens waarop dit verslag gebaseerd is, ontvangen wij van de scholen. Zonder de medewerking van vele docenten en schoolleiders was dit verslag niet mogelijk. Wij willen dit woord vooraf besluiten met een woord van dank aan allen die aan de totstandkoming van dit verslag meegewerkt hebben.

drs. E.J.J. Kremers
accountmanager centrale
examens havo/vwo

drs. H.P.I. Tax
accountmanager centrale
examens vbo/mavo

Inhoud

Woord vooraf

Inleiding 7

1 Algemeen 9

- 1.1 De eindexamenkandidaten 9
- 1.2 Eindexamenvakken en vraagvormen 9
- 1.3 De constructie van opgaven en correctievoorschriften 9
- 1.4 De afname van de examens 12
- 1.5 De verzameling van afnamegegevens 12
- 1.6 Fouten in examenopgaven 13
- 1.7 Normering 13
- 1.8 Normvergelijking 19
- 1.9 Vergelijking van deelvak en totaalvak 24
- 1.10 Vergelijking van reguliere en experimentele examens 24
- 1.11 Niveauverschillen C- en D-examens 25
- 1.12 Verschillen tussen de resultaten van jongens en meisjes 26
- 1.13 De examens tweede tijdvak 28

2 Nieuwe examenvormen 33

- 2.1 Nieuwe examenprogramma's 33
- 2.2 Computers en examens 33
- 2.3 Vernieuwingen in de examens 2002 35
 - 2.3.1 Introductie open vraagvorm en woordenboekgebruik bij de MVT 35
 - 2.3.2 Nieuwe en vernieuwde examens voor het vwo 36
 - 2.3.2.1 Culturele en kunstzinnige vorming 2 36
 - 2.3.2.2 Talen 36
 - 2.3.2.3 Maatschappijvakken 37
 - 2.3.2.4 Exacte vakken 39
 - 2.4 Vakken met wisselende onderwerpen 42
 - 2.5 Vooruitblik 2003 44

3 Examengegevens per vak 47

- 3.1 Cito-vakmedewerkers 47
- 3.2 Tabellen 48

Bijlage 59

- 1 Enkele begrippen 60

Inleiding

De examens 2002 stonden evenals andere jaren volop in de publiciteit. Toch kenden ze over het algemeen een rustig verloop. De reacties op de examens waren zoals dat gebruikelijk is: wisselend kritisch en positief.

Het LAKS (landelijk actie comité scholieren) verzamelt en ordent al sinds enkele jaren op systematische wijze de klachten van examenkandidaten. Het aantal klachten stijgt van jaar tot jaar. Dat wordt vermoedelijk mede veroorzaakt door het gemak waarmee steeds meer kandidaten gebruik maken van internet. Opmerkelijk is dat er, vooral in het havo, toch een zekere mate van overeenstemming bestaat tussen de inhoudelijke reacties die het LAKS mocht ontvangen en de analyses van de examenresultaten. De examens waarover in relatie tot het aantal kandidaten de meeste klachten binnen kwamen, bleken vaak bij nadere analyse van de resultaten ook de examens te zijn die een hogere moeilijkheidsgraad hadden dan was beoogd. Het door het Laks aangelegde bestand zal in het najaar door de Citogroep aan een nadere inhoudelijke analyse worden onderworpen.

Bij de normering van de examens kon dit jaar bij veel examens gebruik worden gemaakt van normvergelijkingsgegevens. Het systeem van normvergelijkingsprocedures waarmee in 1993 gestart is, wordt nog steeds verder uitgebreid. De talen Engels, Frans en Duits zijn nu voor alle niveaus in het systeem opgenomen. Bij havo en vwo konden dit jaar de resultaten voor enkele vakken voor de zesde keer in successie gebruikt worden. Daarnaast werd voor enkele examens deze methodiek dit jaar voor het eerst toegepast. De veranderingen in het onderwijs als gevolg van de invoering van de tweede fase leiden er in een aantal gevallen toe dat de methodiek niet zonder meer kan worden toegepast. (In paragraaf 1.8 wordt nader ingegaan op de resultaten van de normvergelijking.)

De reeds eerder ingezette vernieuwingen in de examens voor de vernieuwde tweede fase werden dit jaar ook voor het vwo in de volle breedte doorgevoerd. Voor het havo/vwo nieuwe stijl bestonden de examens voor een deel uit vernieuwende opgaven. Deze opgaven doen expliciet een beroep op vaardigheden. De programma's voor de vernieuwde tweede fase leggen immers naast kennis en inzicht een sterke nadruk op vaardigheden. Voor de 'nieuwe' vakken Management en organisatie, Culturele en kunstzinnige vorming 2, Turks en Arabisch werden dit jaar voor de tweede maal op vwo-niveau examens afgenomen. De vbo/mavo examens werden dit jaar voor het laatst over de volle breedte aangeboden. In 2003 zullen de eerste examens voor de leerwegen in het nieuwe vmbo worden afgenomen.

Dit examenverslag geeft inzicht in de landelijke resultaten van de centrale examens 2002. In het eerste hoofdstuk wordt een aantal globale cijfermatige gegevens verstrekt en toegelicht. Tevens wordt nader ingegaan op de werkwijze bij de examenproductie, de verwerking van gegevens en de normering. Speciale aandachtspunten zijn normvergelijking, niveauverschillen, verschillen tussen de resultaten van jongens en meisjes en de resultaten van het tweede tijdvak.

In het tweede hoofdstuk krijgen enkele nieuwe of bijzondere examenvormen extra aandacht. Er wordt een overzicht gegeven van examenprogramma's in ontwikkeling en de uitwerking daarvan in de (vwo-)examens. Ook worden de onderwerpen vermeld die aan de orde kwamen bij vakken die jaarlijks wisselende leerstof bevragen.

Het derde en laatste hoofdstuk bevat de cijfermatige gegevens van alle examens afzonderlijk.

De gegevens over de behaalde resultaten zijn afkomstig van de informatie die de scholen verstrekt hebben. Het betreft hier bij het merendeel van de vakken de scores van de (alfabetisch) eerste vijf kandidaten per vak, na eerste correctie.

Sinds 1998 is er uit oogpunt van vergelijkbaarheid voor gekozen om bij de normering expliciet te kijken naar de afnameresultaten van de kandidaten uit het reguliere VO. Bij de presentatie van de afnameresultaten in hoofdstuk 3 zijn in de steekproef dan ook geen kandidaten uit het VAVO opgenomen.

1 Algemeen

1.1 De eindexamenkandidaten

In 2002 namen er in het gehele voortgezet onderwijs (vbo, mavo, havo oude en nieuwe stijl, vwo oude en nieuwe stijl) ongeveer 160.000 kandidaten deel aan het Centrale Examen, eerste tijdvak, dat in mei afgenomen werd. De kandidaten die examen doen op A en/of B-niveau zijn niet in deze aantallen opgenomen.

Tabel 1 geeft een overzicht van de aantallen eindexamenkandidaten die stonden ingeschreven voor het eerste tijdvak van de Centrale Examens 2002.

Ter voorkoming van misverstanden het volgende: de vermelde aantallen zijn de aantallen ingeschreven kandidaten bij het begin van het cursusjaar, zoals verstrekt door de Informatie Beheer Groep te Groningen. Deze aantallen zijn in de regel een overschatting van de aantallen kandidaten die daadwerkelijk aan het examen deelnamen.

1.2 Eindexamenvakken en vraagvormen

In tabel 1 zijn alle algemene eindexamenvakken genoemd die in 2002 in het vbo/mavo en havo en vwo oude en nieuwe stijl centraal geëxamineerd zijn en waarvoor de Citogroep de productie en verwerking verricht.

Sinds 2001 kennen de algemene vakken geen examens meer die volledig uit machinaal scoorbare vragen bestaan. De antwoorden op de open vragen of opdrachten worden beoordeeld door docenten aan de hand van door de CEVO vastgestelde correctievoorschriften. De examens die zowel open als meerkeuzevragen bevatten, worden eveneens geheel door docenten gescoord. De meerkeuzevragen in deze examens worden gescoord aan de hand van de in het correctievoorschrift opgenomen antwoordsleutels. De scoring van examens met alleen open vragen en examens met zowel open vragen als gesloten vragen wordt tweemaal uitgevoerd: eerst door de eigen docent en vervolgens door een docent van een andere school, de tweede corrector.

De examens voor de moderne vreemde talen bestaan voor ongeveer 2/3 uit meerkeuzevragen en voor 1/3 uit open vragen. De meeste overige examens voor het vbo/mavo met zowel open als meerkeuzevragen bestaan voor ongeveer 50% tot 70% uit gesloten vragen en voor 50% tot 30% uit open vragen. Ook de examens biologie en Nederlands havo en vwo bestaan voor een gedeelte uit open vragen en voor een gedeelte uit meerkeuzevragen. De examens beeldende vakken, Latijn en muziek bevatten een beperkt aantal meerkeuzevragen naast meest open vragen. Tabel 2 geeft een overzicht van de gebruikte vraagvormen in de examens.

1.3 De constructie van opgaven en correctievoorschriften

Bij de constructie van opgaven en correctievoorschriften is een groot aantal docenten uit het voortgezet onderwijs betrokken. De verantwoordelijkheid voor de vaststelling van opgaven en correctievoorschriften ligt bij de CEVO (Centrale Examencommissie Vaststelling Opgaven). In het algemeen bestuur van de CEVO zijn de algemene onderwijsorganisaties en het hoger onderwijs vertegenwoordigd. De CEVO kent voor de verschillende vakken vaksecties. Het voorzitterschap van de CEVO-vaksecties wordt vervuld door een vakdeskundige uit het vervolgonderwijs. Naast de voorzitter hebben twee of drie docenten zitting in de vaksectie. Deze docenten worden voorgedragen door de diverse algemene en vakinhoudelijke onderwijsorganisaties.

Het eigenlijke constructiewerk wordt verricht door constructiegroepen waarvan docenten met ervaring in examenklassen deel uitmaken. Per examen worden door een drietal vakdocenten onder leiding van een Cito-medewerker conceptopgaven geschreven en correctievoorschriften opgesteld.

Tabel 1 Aantal examenkandidaten¹⁾

	vbo/mavo	havo oude stijl ²⁾	havo nieuwe stijl	vwo oude stijl ²⁾	vwo nieuwe stijl
Nederlands (leesvaardigheid)	84344	-	42101	-	29085
Nederlands (schrijfv. / opstel)	83360	3494	-	3312	-
Fries leesvaardigheid	18	-	69	21	-
Fries (opstel)	-	3	-	5	-
Latijn	-	-	-	-	4912
Grieks	-	-	-	-	1846
Frans	11508	-	8691	-	7718
Duits	34699	-	10089	-	7428
Engels	82643	-	41989	-	29359
Spaans	722	-	1062	-	700
Russisch	-	-	1	-	14
Arabisch	87	-	13	-	5
Turks	322	-	18	-	6
Wiskunde	65036	-	-	-	-
Wiskunde A	-	3278	-	3241	-
Wiskunde A 1	-	-	-	-	4703
Wiskunde A 1, 2	-	-	16051	-	9292
Wiskunde B	-	1228	-	1533	-
Wiskunde B 1	-	-	6012	-	6669
Wiskunde B 1, 2	-	-	5601	-	6039
Natuurkunde	27181	-	-	-	-
Natuurkunde 1	-	-	6055	-	6919
Natuurkunde 1, 2	-	-	6735	-	7199
Scheikunde	16504	1211	10882	1650	-
Scheikunde 1	-	-	-	-	6225
Scheikunde 1,2	-	-	-	-	5842
Biologie (1,2)	37694	-	14501	-	10865
Economie	38432	-	-	-	-
Economie I	-	-	-	-	-
Economie II	-	-	-	1281	-
Economie 1	-	-	14667	-	5913
Economie 1, 2	-	-	18526	-	11946
Handelwetenschappen	-	1142	-	-	-
Management en Organisatie	-	-	11650	-	7746
Aardrijkskunde	22902	1629	18794	1672	11505
Geschiedenis en Staatsinrichting	17027	-	29187	-	16733
Maatschappijleer	4525	-	2911	-	2040
Filosofie	-	-	276	262	1659
Muziek	746	-	1186	-	713
Tekenen	7259	-	5023	-	2685
Handenarbeid	3293	-	1717	-	460
Textiele werkvormen	268	-	182	-	33
CKV 2	-	-	3398	-	1700

¹⁾ In deze aantallen zijn ook de examenkandidaten van de Antillen opgenomen. De cijfers voor Nederland liggen derhalve iets lager. Daarnaast ligt het feitelijk aantal kandidaten gemiddeld enkele procenten lager dan het opgegeven aantal, omdat scholen een zekere veiligheidsmarge in hun bestellingen inbouwen.

²⁾ Inschrijvingen voor de specifiek oude stijl examens.

Tabel 2 Overzicht van toegepaste vraagvormen in de centrale examens 2002

	VBO/MAVO-C	VBO/MAVO-D	HAVO oude stijl	HAVO nieuwe stijl	VWO oude stijl	VWO nieuwe stijl
Latijn					<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Grieks					<input type="checkbox"/>	<input type="checkbox"/>
Nederlands	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Fries	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Frans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Duits	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Engels	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Spaans	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Russisch	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Turks	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
Arabisch	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
Wiskunde	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Natuurkunde	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Scheikunde	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Biologie	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Economie	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Handelswetenschappen			<input type="checkbox"/>			
Management & Organisatie				<input type="checkbox"/>		<input type="checkbox"/>
Aardrijkskunde	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Geschiedenis / staatsinrichting	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Filosofie				<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Maatschappijleer	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Muziek	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Tehatex	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
CKV 2				<input type="checkbox"/>		<input type="checkbox"/>

open vragen

open en gesloten vragen

Bij de constructie van examenopgaven en correctievoorschriften wordt gewerkt met examenmodellen. In een examenmodel zijn de productspecificaties ten aanzien van het uiteindelijke examen opgenomen, zoals:

- de verdeling van de vragen over de examenstof;
- de mate waarin het examen vragen met een reproductief en/of productief karakter moet bevatten;
- de keuze van teksten en contextmateriaal;
- de toe te passen vraagvormen en vaardigheidsvragen;
- de wijze waarop de correctievoorschriften opgesteld dienen te zijn.

De examenmodellen worden door de CEVO vastgesteld en fungeren voor de Citogroep en de constructiegroepen als opdracht voor de constructie. Een examenmodel bevordert dat elk examen een evenwichtige representatie van het examenprogramma is en dat examens van verschillende tijdvakken en verschillende jaren inhoudelijk als gelijkwaardig beschouwd kunnen worden.

De conceptexamens worden van commentaar voorzien en beoordeeld door de vaksecties van de CEVO. Aan de commentaarrondes wordt onder verplichting van strikte geheimhouding ook deelgenomen door niet-schoolgebonden deskundigen, zoals universitaire deskundigen voor speciale onderwerpen en voor de moderne vreemde talen deskundigen die de desbetreffende taal als moedertaal spreken.

De CEVO-vaksectie stelt ten slotte de opgaven en het correctievoorschrift in hun definitieve vorm vast.

De constructie van opgaven en correctievoorschriften betreft elk jaar drie tijdvakken: het eerste tijdvak, dat in de tweede helft van mei afgenomen wordt, het tweede tijdvak (de herkansing), dat in juni wordt afgenomen, en het derde tijdvak (afname in augustus) dat onder meer bedoeld is voor kandidaten die – om geldige redenen – niet aan de eerste twee tijdvakken konden deelnemen.

In totaal zijn er dit jaar voor de algemene vakken vwo/havo en vbo/mavo C- en D-niveau 311 examens geconstrueerd, 120 voor het eerste tijdvak, 107 voor het tweede tijdvak en 84 voor het derde tijdvak.

1.4 De afname van de examens

De periodes waarin de Centrale Examens afgenomen worden, zijn ruim voor het begin van het cursusjaar aan de scholen bekendgemaakt. Het examenrooster 2002 vermeldde 15 mei tot en met 31 mei als de periode van het eerste tijdvak, 19 juni als de dag van het tweede tijdvak. Voor enkele weinig gekozen vakken vond de afname van het tweede tijdvak plaats op vrijdag 21 juni. De periode voor het derde tijdvak was van 12 augustus tot en met 16 augustus. De organisatie van de afname van de examens is in handen van de schoolleiding. Tijdens de examenperiode begint ook het correctiewerk voor docenten/examinatoren. De correctieperiode loopt door tot enkele dagen vóór de datum waarop de uitslag gegeven wordt. In dezelfde periode komen bij de Citogroep dagelijks duizenden optisch leesbare formulieren binnen om machinaal verwerkt te worden. Daarover meer in de volgende paragraaf.

1.5 De verzameling van afnamegegevens

De activiteiten voor de verzameling van afnamegegevens ten behoeve van de normering beginnen lang voor de examenperiode. In november vindt de inschrijving van kandidaten bij de Productgroep examendiensten van de Informatie Beheer Groep in Groningen plaats. Deze gegevens worden ook door de Citogroep gebruikt. Het merendeel van de afnamegegevens is op de gebruikelijke wijze via optisch leesbare formulieren verzameld. Een deel van de gegevens is via elektronische weg verzameld. Beide wijzen van het verzamelen van gegevens worden hieronder kort aan de orde gesteld.

Gegevensverzameling via optisch leesbare formulieren

Begin maart zijn de pakketten met formulieren voor de correctoren aan de scholen toegestuurd. Op deze formulieren worden de gegevens van vijf kandidaten (of voor enkele vakken tien of twintig) per school verzameld. De gegevens uit deze steekproef worden op de normeringvergaderingen gebruikt om na te gaan hoe de examens door de kandidaten gemaakt zijn.

Voor zover de examens vwo oude stijl identiek waren aan de nieuwe-stijl-examens, zijn er geen afzonderlijke oude-stijl-gegevens opgevraagd. Voor deze examens zijn de bezemkandidaten naar rato vertegenwoordigd onder de nieuwe-stijl-kandidaten. Voor de specifieke oude-stijl-vwo-examens zijn de gegevens opgevraagd van de op alfabetische volgorde eerste vijf kandidaten.

Voor het tweede tijdvak werden, net als vorig jaar, op formulieren voor de corrector de gegevens van vijf kandidaten of minder per school verzameld. Tevens werd gevraagd het cijfer voor het eerste tijdvak aan te strepen. Deze gegevens werden gebruikt voor een vergelijking tussen eerste en tweede tijdvak.

Op het formulier voor de corrector diende per open vraag de behaalde score aangestreept te worden en per gesloten vraag het gekozen alternatief. Dit laatste gegeven is essentieel om de kwaliteit van de afleiders (de onjuiste alternatieven) te kunnen beoordelen. Sinds 2001 bestaat de mogelijkheid om het vraagnummer door te strepen in geval een kandidaat de vraag in het geheel niet beantwoord had. De mate waarin met name vragen aan het eind van het examen door kandidaten niet beantwoord zijn, kan geïnterpreteerd worden als een indicatie voor tijdnoed en kan er op wijzen dat een examen als geheel te bewerkelijk is geweest.

Het aantal formulieren van de correctoren (inclusief de elektronische) dat verwerkt werd, bedroeg voor het eerste tijdvak ruim 175.000 en voor het tweede tijdvak 23.000.

Gegevensverzameling via elektronische weg

Er zijn technisch gesproken ook andere mogelijkheden beschikbaar voor het verzamelen van afnamegegevens. Zeker nu sinds 2001 de antwoordbladen (dit zijn de optisch leesbare formulieren die de leerlingen zelf invulden bij de 100%-meerkeuze-examens) definitief tot het verleden behoren, is de tijd rijp voor vernieuwing.

De verzameling van afnamegegevens via optisch leesbare formulieren is niet alleen kostbaar, bij sommige vakken treedt ook uitval van gegevens op omdat bij het invullen van de formulieren door de correctoren fouten worden gemaakt.

Door elektronische invoer en uitwisseling van afnamegegevens wordt het mogelijk om invoerfouten te minimaliseren en de snelheid van de gegevensverzameling en -verwerking te verhogen.

Dit jaar hebben 200 scholen meegedaan aan het experiment WOLF2002. Wolf staat voor: Windows Optisch Leesbaar Formulier. Dat betekent dat de docenten van examenklassen na het examen de scoregegevens van de alfabetisch eerste vijf (of tien) leerlingen niet op papieren (optisch leesbare) formulieren voor correctoren zetten, maar via een computerprogramma hebben ingevoerd. De school- en examengegevens zijn via een pincode te downloaden van internet en de ingevulde scores worden weer via het web teruggestuurd naar de Citogroep. WOLF2002 neemt de gebruiker bij de hand; het is te vergelijken met het elektronisch aangifteformulier van de belastingdienst. Wordt er iets fout ingevuld of overgeslagen, dan meldt het programma dat direct.

Het voordeel voor de docent is dat hij met behulp van WOLF2002 tegelijkertijd een scoreoverzicht voor de tweede corrector kan samenstellen. Er hoeven dus maar één keer gegevens te worden ingevoerd. Ook voor de Citogroep zal (op termijn) het programma zijn vruchten afwerpen: de data die we op deze manier binnenkrijgen zullen eerder beschikbaar zijn en gebleken is dat er geen fouten meer in de gegevens zitten.

De ervaringen en reacties van de deelnemende scholen zijn overwegend positief. Wel moet worden opgemerkt dat het programma ten opzichte van eerste prototype met nieuwe functionaliteiten was uitgerust en deze bleken nog niet helemaal vrij te zijn van kinderziektes. In overleg met de CEVO en VVO zal bekeken worden of in de campagne 2003 op grotere schaal gebruik gemaakt gaat worden van de elektronische gegevensverzameling. Op termijn wordt gestreefd naar invoering bij alle scholen.

1.6 Fouten in examenopgaven

Ondanks het feit dat elke examenopgave door verschillende groepen deskundigen onafhankelijk van elkaar bekeken wordt en dat drukproefcorrecties door verschillende betrokkenen uitgevoerd worden, is het nooit helemaal uitgesloten dat er in een of enkele examens een storende fout over het hoofd wordt gezien. Wanneer deze fout nog bij de laatste controle, enkele weken voor het examen, geconstateerd wordt, ontvangen de scholen een erratumblad. Een fout in de opgaven moet bij het begin van het examen worden voorgelezen en is dus zeer storend voor de kandidaten. Een fout in het correctievoorschrift moet ter kennis worden gebracht van de examinatoren/correctoren. Wanneer de keuzemogelijkheid aanwezig is, wordt de voorkeur gegeven aan een erratumblad voor het correctievoorschrift boven een erratumblad voor de opgaven.

Dit jaar moesten er voor het eerste tijdvak 7 erratumbladen voor de opgaven en 5 erratumbladen voor de correctievoorschriften verzonden worden. Voor het tweede tijdvak moesten er 3 erratumbladen voor de opgaven en 11 voor de correctievoorschriften verzonden worden. Scholen krijgen van de Informatie Beheer Groep een overzicht van de examens waarvoor een erratumblad verschenen is. Dit wordt ook vermeld op teletekst en internet.

Blijkt er bij de afname desondanks nog een fout te staan in een examen of in een correctievoorschrift, dan stuurt de CEVO de scholen een tussentijds bericht over de wijze waarop deze vraag beoordeeld moet worden. Dit bericht wordt ook op teletekst en internet vermeld. Dit jaar was dat bij 6 examens in het eerste tijdvak nodig, en bij 1 examen in het tweede tijdvak. Een tussentijds bericht kan alleen verzonden worden wanneer de fout kort na afname van het examen ontdekt is. Wordt tijdens de normeringvergadering besloten dat een vraag toch op vakinhoudelijke gronden minder gelukkig was, dan kan dat alleen leiden tot aanpassing van de normeringsterm. Wanneer een bepaalde druk- of zetfout geen probleem oplevert bij de beantwoording van de vraag, vindt er geen correctie plaats.

Het aantal errata was dit jaar iets lager dan vorig jaar en dat was al weer lager dan het aantal in 2000 toen dat op 0,5 % van het aantal vragen lag. In 2002 lag het aantal errata op iets minder dan 0,4% van het aantal vragen. Voor de Citogroep geldt echter dat elke fout er een te veel is en daarom blijven we al het mogelijke doen om dit soort fouten te voorkomen.

1.7 Normering

De normeringvergaderingen voor de eerste-tijdvak-examens vonden dit jaar plaats op 11 en 12 juni. De verzending van de examenresultaten eerste tijdvak vond voor havo/vwo op woensdag 12 en voor vbo/mavo op donderdag 13 juni plaats. De resultaten van de

normering werden op dezelfde dagen bekend gemaakt via de websites www.eindexamen.nl en www.citogroep.nl.

Onder normering van examens verstaan we in het algemeen de procedure die gevolgd wordt bij de omzetting van scores in cijfers. Daarbij is de vaststelling van de normeringsterm de belangrijkste beslissing. De in 2000 ingevoerde normeringmethodiek is (m.u.v. Nederlands stelopdracht) voor alle examens toepasbaar ongeacht de gehanteerde vraagvormen of de lengte van de scoreschaal.

Door de nieuwe normeringmethodiek zijn de bonuspunten (10 scorepunten vooraf) vervallen. In plaats daarvan wordt het cijfer verhoogd met de normeringsterm. De normeringsterm is een ijkvariabele die zo wordt vastgesteld dat in verschillende jaren en tijdvakken even zware eisen aan kandidaten gesteld kunnen worden om een voldoende voor het centrale examen te halen. De normeringsterm kan variëren van 0,0 tot en met 2,0. Een waarde hoger dan 1,0 duidt niet zonder meer op een versoepeling van de norm maar op een ijking van een examen dat aan de moeilijke kant is. Op de normeringvergaderingen worden door de CEVO de normeringstermen vastgesteld.

In de besluitvorming ten aanzien van de normering zijn globaal genomen drie stappen te onderscheiden. De eerste stap betreft de advisering van de Cito(vak)medewerker op basis van de geanalyseerde afnamegegevens. De Cito(vak)medewerker stelt het advies op voorafgaand aan de normeringsvergadering van de vaksectie.

De tweede stap betreft de advisering door de vaksectie aan het DB van de CEVO.

Op basis van het Cito-advies alsook een vakinhoudelijke beoordeling van het examen en weging van binnen gekomen reacties en commentaren stelt de CEVO-vaksectie een normeringsadvies op voor het DB van de CEVO.

Het DB van de CEVO verzamelt alle adviezen en beoordeelt deze o.a. op consistentie over vakken en schooltypen en stelt vervolgens de definitieve normering vast.

De advisering door de Cito(vak)medewerker berust voornamelijk op de interpretatie van de analyses van de afnamegegevens. Het advies van de Cito(vak)medewerker is in eerste instantie vooral een technisch advies. Het is technisch in die zin dat op basis van de verzamelde afnamegegevens een zodanige normering wordt voorgesteld dat over jaren en tijdvakken heen gelijke eisen aan kandidaten worden gesteld. De norm die als uitgangspunt wordt genomen is door de CEVO bepaald door voor elk vak een referentie-examen aan te wijzen met een daarbij toe te passen referentienorm. De Cito(vak)medewerker adviseert voor het nieuwe examen een normering die equivalent is aan de referentienorm op het referentie-examen.

De gegevensverzameling en de daarop aansluitende normeringprocedures zijn niet voor alle vakken volledig identiek. Voor sommige examens zijn er meer en andersoortige gegevens beschikbaar om de normeringbeslissing te ondersteunen. Voor die examens waarbij de gegevensverzameling minder volledig is moeten noodzakelijkerwijze enkele aannamen worden gedaan. De waarde van het Cito-advies is afhankelijk van de juistheid van die aannamen. Maar gegeven het uitgangspunt is de redeneertrant transparant. Het Cito-advies heeft geen 'absolute' pretenties, maar als de geëxpliciteerde aannamen juist zijn dan zou ook het resultaat juist moeten zijn. We zullen dat hieronder nader toelichten.

Bij de gegevensverzameling worden globaal de volgende varianten onderscheiden:

1. de standaardprocedure van gegevensverzameling die voor (nagenoeg) alle examens geldt. Dit houdt in dat er bij een steekproef van ongeveer 2000 kandidaten deelscores per vraag worden verzameld;
2. sommige examens met overwegend de gesloten vraagvorm zijn naast de standaardprocedure ook nog opgenomen in de procedure 'normvergelijking achteraf' (dit zijn de examens Frans, Duits en Engels voor alle schooltypen en Nederlands leesvaardigheid en Muziek voor vbo/mavo);
3. sommige examens met uitsluitend of overwegend open vragen zijn naast de standaardprocedure ook nog opgenomen in de procedure 'normvergelijking vooraf' (dit zijn m.n. de exacte vakken op havo en vwo-niveau).
4. Tot slot zijn er de examens die naast de standaardprocedure ook nog overlap vertonen met een ander examen en waarbij die examens in relatie tot elkaar worden genormeerd. Dit geldt bijvoorbeeld voor de vbo/mavo-C- en D-examens waarbij op basis van overlap tussen beide examens aanvullende analyses worden uitgevoerd.

Ad 1) De standaardprocedure is als volgt. Het definitieve correctievoorschrift wordt vóór afname van het examen vastgesteld. Tegelijk met de opgaven wordt dit definitieve voorschrift naar de scholen gestuurd. Bij de constructie van het merendeel van de examens is een gemiddelde moeilijkheidsgraad nagestreefd zodat in de regel een normeringsterm tussen 0,8 en 1,3 kan gelden. De docent/examinator kan onmiddellijk na afname van het examen met de correctie beginnen en op grond van het correctievoorschrift een score bepalen. De normeringsterm wordt definitief vastgesteld op basis van een steekproef van de scores van kandidaten na eerste correctie – de alfabetisch eerste 5 kandidaten per vak – en tevens op basis van reacties van docenten. Bij het opstellen van het Cito-advies wordt aangenomen dat de populaties van jaar tot jaar een vergelijkbaar prestatieniveau hebben. Een verschil in de gemiddelde score met het referentie-examen wordt derhalve gezien als een gevolg van een verschil in moeilijkheidsgraad met het referentie-examen. Aan de hand van de analyseresultaten wordt vervolgens nagegaan met welke normeringsterm het verschil in moeilijkheidsgraad adequaat gecompenseerd kan worden.

Ad 2) De procedure voor examens die zijn opgenomen in de 'normvergelijking achteraf' verloopt als volgt. Bij de samenstelling en de vaststelling van de examens wordt een zelfde moeilijkheidsgraad nagestreefd als die van een vooraf gekozen referentie-examen met de daarbij behorende referentienorm. Na afname van de examens wordt de feitelijke moeilijkheidsgraad van het nieuwe examen vergeleken met die van het referentie-examen en wordt de norm van het referentie-examen overgebracht op het nieuwe examen. Dit gebeurt op basis van de normvergelijgingsgegevens. Bij deze werkwijze wordt er derhalve niet voorshands van uitgegaan dat populaties over jaren heen een zelfde prestatieniveau hebben. Voor de moderne vreemde talen gold in 1999 (bij een schaallengte van 50) een normering die overeenkwam met een normeringsterm die ligt tussen 0,0 (cesuur 30/31) en 1,0 (cesuur 24/25). Sinds 2000 vindt de correctie plaats op school en niet meer zoals voorheen bij de Citogroep. In de Maartmededelingen 2002 heeft de CEVO voor deze vakken een marge van de normeringsterm aangegeven. Deze marge was geschat op basis van de resultaten uit het verleden.

Ad 3) De gegevens van de 'normvergelijking vooraf' worden verzameld door middel van een pretest. Deze gegevens zijn reeds verwerkt bij de samenstelling van de definitieve examens en het opstellen van de correctievoorschriften. Deze examens zijn meestal met een normeringsterm van 1,0 equivalent aan het aangewezen referentie-examen. Ook voor de examens wordt niet voorshands aangenomen dat populaties over jaren heen een zelfde prestatieniveau hebben. Alleen wanneer na de pretest nog ingrijpende veranderingen in het examen zijn aangebracht geldt een vergelijkbare werkwijze als hierboven beschreven onder 1.

Ad 4) Bij het vbo/mavo speelt ook nog de afstand tussen de C- en D-norm een rol. Op basis van de overlap tussen beide examens wordt het verschil in moeilijkheidsgraad tussen de examens geschat en wordt er bij normering gestreefd naar een substantieel niveau-onderscheid tussen de C- en D-kwalificatie. Bij de advisering door de Cito(vak)medewerker wordt een verschil tussen C en D van één cijferpunt nagestreefd. Dat wil zeggen dat de adviesnormering van het C- en D-examen zodanig is dat een 5,5 op het D-examen overeenkomt met een 6,5 op het C-examen.

Ook bij de experimentele examens (zie § 1.10) speelt de overlap met de reguliere examens een belangrijke rol. Hier wordt expliciet niet naar een afstand tussen beide examens gestreefd maar wordt op basis van de overlap getracht de beide examens zo te normeren dat aan beide groepen kandidaten dezelfde eisen worden gesteld.

tweede-tijdvak-normering

In principe heeft het tweede tijdvak dezelfde normeringsterm als het eerste tijdvak. Wanneer bij de vaststelling van de normeringsterm voor het eerste tijdvak het feit van één of meer onjuiste vragen in dat examen is meegewogen, dan werkt dit niet door naar het tweede tijdvak. Niet de normeringsterm van het eerste tijdvak geldt dan als 'voorlopige' normeringsterm voor het tweede tijdvak, maar de normeringsterm die vastgesteld zou zijn als er geen onjuiste vraag in het eerste tijdvak was gebleken. De aanname daarbij is dat de eerste- en tweede-tijdvak-examens, die gelijktijdig geconstrueerd en vastgesteld worden, gelijkwaardige examens zijn.

De voorlopige normeringstermen voor het tweede tijdvak zijn voorlopig in die zin dat ze nooit ten nadele van de kandidaten zullen worden bijgesteld. De normeringstermen voor het tweede tijdvak zullen niet lager worden vastgesteld. Ze worden alleen gewijzigd wanneer uit

de vergelijking van het eerste en tweede tijdvak onomstotelijk blijkt, dat het tweede tijdvak moeilijker is dan het eerste tijdvak. De vergelijking van het eerste en tweede tijdvak vindt plaats aan de hand van de scores van het tweede tijdvak en het cijfer voor het eerste tijdvak. Deze worden direct na de correctie op optisch leesbare formulieren naar de Citogroep opgestuurd.

Score-cijfer-transformatie

Behalve de normeringsterm is ook de wijze van omzetten van score naar cijfer bij de normering van belang. Sinds 2000 geldt ook hiervoor een nieuwe systematiek.

Uitgangspunten bij de omzetting van score naar cijfer zijn:

- elk volgend gescoord punt leidt altijd tot een hoger examencijfer (afgezien van afrondingen);
- een score van 0% correspondeert altijd met het examencijfer 1;
- een score van 100% correspondeert altijd met het examencijfer 10;
- over een zo breed mogelijk centraal interval van de scoreschaal is sprake van een constante stijging van het cijfer per extra scorepunt (onafhankelijk van de normering).

Doordat de lengte van de scoreschaal variabel is geworden, is de Score-cijfer-transformatie iets minder transparant geworden. De omzetting van scores naar cijfers dient op basis van door de CEVO vastgestelde formules te worden uitgevoerd (zie Uitleg Gele Katern nr. 8 van 22 maart 2000).

Voor elk examen zijn na de normering omzettingstabellen gegenereerd. Deze zijn alleen per e-mail verzonden aan die scholen die zich daarvoor hebben aangemeld. Daarnaast werden de omzettingstabellen gepubliceerd op de sites www.eindexamen.nl en www.citogroep.nl. Deze omzettingstabellen zijn berekend volgens de vastgestelde formules en de vastgestelde normeringstermen.

In de hierna volgende tabellen wordt per schooltype een overzicht gepresenteerd van de feitelijke per vak genomen normeringsbeslissing en het daarmee corresponderende percentage onvoldoendes.

Vbo/mavo C-examens

Bij de normering van de C-examens deden zich geen uitzonderlijke zaken voor. De C-examens worden altijd in samenhang met de D-examens genormeerd. Alle vastgestelde normeringstermen liggen binnen de vooraf aangegeven marges (zie tabel 3).

Evenals vorig jaar valt op dat bij Engels het percentage onvoldoendes hoog ligt in verhouding tot Frans. Voor Engels geldt dat het verschil tussen de C- en D-populaties, relatief gezien, een stuk groter is dan bij Frans (zie ook § 1.12). Het percentage onvoldoendes voor wiskunde is relatief ten opzichte van de andere examens aan de hoge kant maar ligt in dezelfde orde van grootte als afgelopen jaar.

Vbo/mavo D-examens

Ook bij de normering van de D-examens deden zich geen uitzonderlijke zaken voor. De D-examens worden altijd in samenhang met de C-examens genormeerd. Ook voor de D-examens geldt dat alle vastgestelde normeringstermen binnen de vooraf aangegeven marges liggen (zie tabel 4).

Voor Frans bleek het D-examen 2002 evenals vorig jaar een stuk makkelijker te zijn dan het referentie-examen. Bij de normering werd hiervoor ten dele gecompenseerd.

Net als vorig jaar vertoont het vak Nederlands leesvaardigheid met 12% onvoldoendes zowel in vergelijking met de andere vakken als in vergelijking met de resultaten in het verleden een opmerkelijk resultaat.

Bij de overige examens zijn er geen ongebruikelijke percentages onvoldoendes. Bij een aantal vakken zijn de percentages in vergelijking met vorig jaar iets hoger en bij andere weer iets lager. Het is niet mogelijk om dit ondubbelzinnig toe te schrijven aan een geringere of grotere vaardigheid bij de kandidaten van dit jaar. Een dergelijke conclusie is alleen mogelijk op basis van de afnameresultaten van examens waarvoor normvergelijking wordt uitgevoerd (zie ook paragraaf 1.8).

Tabel 3 Normeringstermen en percentages onvoldoendes examens C-niveau, 1e tijdvak

Vak	L (lengte van de scoreschaal)	N (normeringsterm)	% onvoldoendes	incidentele aanpassing
Nederlands, leesvaardigheid	51	1,3	18	-
Nederlands, functioneel schrijven	90	0,8	20	-
Frans	46	0,6	18	-
Duits	46	0,6	24	-
Engels	49	0,6	34	-
Spaans	49	1,1	*	-
Turks	50	1,2	3	-
Arabisch	47	1,1	*	-
Wiskunde	90	1,1	37	-
Natuurkunde	90	1,3	32	-
Scheikunde	90	1,2	21	-
Biologie	95	1,0	20	-
Economie	85	1,2	20	-
Aardrijkskunde	90	0,8	10	-
Geschiedenis en staatsinrichting	85	0,8	15	-
Maatschappijleer	90	1,1	21	-
Muziek	50	0,5	15	-
Tekenen	90	1,3	24	-
Handenarbeid	90	1,3	27	-
Textiele werkvormen	90	1,3	*	-

*: onvoldoende gegevens

Tabel 4 Normeringstermen en percentages onvoldoendes examens D-niveau, 1e tijdvak

Vak	L (lengte van de scoreschaal)	N (normeringsterm)	% onvoldoendes	incidentele aanpassing
Nederlands, leesvaardigheid	48	0,8	12	-
Nederlands, functioneel schrijven	90	0,8	18	-
Frans	46	0,1	23	-
Duits	49	1,0	28	-
Engels	46	0,4	18	-
Spaans	48	1,0	43	-
Turks	52	0,5	16	-
Arabisch	46	0,1	0	-
Wiskunde	90	1,1	33	-
Natuurkunde	90	0,9	27	-
Scheikunde	90	1,0	17	-
Biologie	86	0,8	18	-
Economie	85	1,0	18	-
Aardrijkskunde	90	0,8	9	-
Geschiedenis en staatsinrichting	85	0,8	10	-
Maatschappijleer	90	1,3	18	-
Muziek	50	0,0	16	-
Tekenen	90	1,3	25	-
Handenarbeid	90	1,0	24	-
Textiele werkvormen	90	1,1	21	-

Havo nieuwe stijl

In 2001 deden zich vooral bij biologie en geschiedenis problemen voor bij de havo-nieuwe-stijl examens en leek voor de andere vakken de afstemming van de moeilijkheidsgraad van de nieuwe examens op het prestatieniveau van de leerlingenpopulatie goed geslaagd. Dit jaar bleken geen problemen bij biologie en geschiedenis maar bij enkele andere examens.

Met name de wiskunde-examens bleken over de hele linie voor de kandidaten moeilijkheden op te leveren.

Enkele examens bleken aan de makkelijke kant te zijn. Voor deze vakken is een normeringsterm vastgesteld die kleiner is dan 1.0. Dit geldt o.a. voor de vakken natuurkunde 1 en 1,2 scheikunde, filosofie en textiele wiskunde A1,2.

Voor de specifieke examens havo oude werkvormen. Het hoogste percentage onvoldoendes valt dit jaar bij economie1,2 en bij stijl zijn geen afnamegegevens verzameld. In tabel 6 worden derhalve geen percentage onvoldoendes vermeld.

Tabel 5 Normeringstermen en percentages onvoldoendes examens havo nieuwe stijl, 1e tijdvak

Vak	L (lengte van de scoreschaal)	N (normeringsterm)	% onvoldoendes	incidentele aanpassing
Nederlands	50	1,1	18	-
Frans	48	1,2	37	-
Duits	50	0,8	37	-
Engels	50	0,8	26	-
Engels complex	50	1,5	27	-
Spaans	49	0,8	34	-
Russisch	49	0,7	-	-
Turks	50	0,8	0	-
Arabisch	45	0,7	0	-
Wiskunde A 1,2	90	1,8	41	-
Wiskunde B 1	83	2,0	31	-
Wiskunde B 1,2	88	1,8	27	-
Natuurkunde 1	78	0,8	12	-
Natuurkunde 1,2	84	0,8	13	-
Scheikunde	81	0,8	19	-
Biologie	84	0,9	26	-
Biologie complex	88	1,5	19	-
Economie 1	63	1,2	31	-
Economie 1,2	60	1,2	43	-
Management & organisatie	65	1,0	16	-
Aardrijkskunde	67	1,2	30	-
Geschiedenis en staatsinrichting	81	1,2	33	-
Maatschappijleer	83	1,1	33	-
Filosofie	59	0,8	9	-
Muziek	90	1,1	15	-
Tekenen	90	1,0	25	-
Handenarbeid	90	1,3	22	-
Textiele werkvormen	90	0,8	7	-
CKV 2 ¹⁾	66	0,9	15	-

¹⁾ Geen Centraal Examen

Tabel 6 Normeringstermen en schaallengte examens havo oude stijl, 1e tijdvak

Vak	L (lengte van de scoreschaal)	N (normeringsterm)	incidentele aanpassing
Aardrijkskunde	74	1,2	-
Wiskunde A	90	1,5	-
Wiskunde B	90	1,0	-
Scheikunde	80	0,8	-
Handelswetenschappen	90	1,0	-

Vwo-examens nieuwe stijl

Bij het normeren van een nieuw schooltype met enkele nieuwe vakken kan men niet, zoals bij de andere schooltypen, terugvallen op een in de traditie gewortelde norm. Desalniettemin was de insteek om vergelijkbare eisen te stellen als in 2001 voor de voorhoede scholen. Slechts bij één examen speelde bij de normering de aanwezigheid van een minder geslaagde vraag expliciet een rol. Bij wiskunde A1,2 betrof dit een vraag die weliswaar inhoudelijk correct was maar waarbij in het correctievoorschrift – achteraf gezien ten onrechte – een aftrekpunt aan een bepaalde aanpak werd toegekend. Enkele examens bleken achteraf toch iets moeilijker te zijn dan beoogd. Dit betrof de examens Wiskunde A1 en A1,2 en de examens economie 1 en 1,2. Globaal genomen zijn de percentages onvoldoendes redelijk vergelijkbaar met die van 2001. Bij Frans 1,2 en Duits 1,2 vallen relatief de hoogste percentages onvoldoende.

Vwo examens oude stijl

Voor zes vakken werden er specifieke oude-stijl-examens aangeboden. Ook hier bleken achteraf enkele examens moeilijker te zijn dan beoogd en werd een normeringsterm buiten de vooraf aangegeven marge vastgesteld. Met name wiskunde B bleek voor veel kandidaten moeilijkheden op te leveren.

Het percentage onvoldoendes ligt over het geheel genomen aan de hoge kant maar daarbij moet in ogenschouw worden genomen dat de populatie voor een deel bestaat uit recidivisten.

Tabel 7 Normeringstermen en percentages onvoldoendes examens vwo nieuwe stijl, 1e tijdvak

Vak	L (lengte van de scoreschaal)	N (normeringsterm)	% onvoldoendes	incidentele aanpassing
Nederlands	49	1,0	20	-
Latijn	90	1,1	28	-
Grieks	84	1,3	23	-
Frans 1,2	46	0,1	36	-
Duits 1,2	49	0,9	36	-
Engels	47	0,7	22	-
Spaans	51	0,5	21	-
Russisch	51	0,6	-	-
Turks	55	0,8	-	-
Arabisch	45	0,7	-	-
Wiskunde A 1	81	1,6	23	-
Wiskunde A 1,2	90	1,4	29	0,1
Wiskunde B 1	84	0,8	26	-
Wiskunde B 1,2	86	0,8	24	-
Natuurkunde 1	78	1,0	25	-
Natuurkunde 1,2	83	0,8	8	-
Moderne Natuurkunde	83	1,2	5	-
Scheikunde 1	66	1,3	26	-
Scheikunde 1,2	71	1,0	18	-
Biologie 1,2	82	0,9	19	-
Economische wetenschappen 1	62	1,5	29	-
Economische wetenschappen 1,2	61	1,5	24	-
Management & organisatie	74	1,1	19	-
Aardrijkskunde	75	0,8	17	-
Geschiedenis	90	0,8	23	-
Maatschappijleer	82	0,8	15	-
Filosofie	44	0,8	9	-
Muziek	90	1,0	17	-
Tekenen, handvaardigh., text. werkv.	68	0,8	17	-
CKV 2 ¹⁾	62	0,6	18	-

¹⁾ Geen Centraal Examen

Tabel 8 Normeringstermen en percentages onvoldoendes examens vwo oude stijl, 1e tijdvak

Vak	L (lengte van de scoreschaal)	N (normeringsterm)	% onvoldoendes	incidentele aanpassing
Wiskunde A	90	1,3	29	-
Wiskunde B	90	2,0	46	-
Scheikunde	70	1,0	43	-
Economische wetenschappen II	90	1,4	48	-
Aardrijkskunde	86	1,2	30	-
Filosofie	45	1,0	43	-

1.8 Normvergelijking

In de vorige paragraaf is aangegeven dat de gegevensverzameling niet voor alle examens identiek is. Sommige examens zijn opgenomen in procedures van normvergelijking. Daarbij werd onderscheid gemaakt in 'normvergelijking achteraf' en 'normvergelijking vooraf'. Normvergelijkingsprocedures 'achteraf' zijn in 2002 toegepast bij de volgende eerste-tijdvak-examens:

- de examens Frans, Duits en Engels op alle niveaus (m.u.v. vbo/mavo-C)
- Nederlands en muziek op vbo/mavo-D;

Normvergelijkingsprocedures 'vooraf' zijn voor de examens 2002-1 toegepast bij

- wiskunde A en B, natuurkunde, scheikunde (zowel voor havo als vwo);
- economie 1,2 (havo) en Nederlands leesvaardigheid (havo);
- biologie (vwo).

Het systeem van normvergelijking moet ervoor zorgen dat er over de jaren en tijdvakken heen gelijke eisen gesteld worden aan de kandidaten. Het is niet gewenst dat examen-opgaven en de normering van de examens ongemerkt met veranderingen in de vaardigheid van de kandidaten meebewegen.

In deze procedures worden onderdelen van de nieuwe examens gecombineerd met onderdelen uit een referentie-examen afgenomen. Het referentie-examen is tevoren door de CEVO met zorg uitgekozen en voorzien van een referentienorm die men als een redelijke eis beschouwt bij het betreffende referentie-examen.

De afnamen maken uitspraken mogelijk over de moeilijkheidsgraad van het examen 2002 in vergelijking met het referentie-examen. Ook kunnen uitspraken worden gedaan over de populatie 2002 in vergelijking met de populatie van het referentiejaar. De CEVO gebruikt deze gegevens bij de vaststelling van het examen of bij de vaststelling van de normeringsterm na de afname van het examen.

Normvergelijking 'achteraf'

Voor de talenexamens vbo/mavo en havo en muziek vbo/mavo werd de normvergelijking reeds voor het negende achtereenvolgende jaar uitgevoerd. Voor de vwo-examens Duits, Engels en Frans is deze normvergelijking dit jaar voor de zesde maal uitgevoerd.

Tabel 8 geeft een overzicht van de verschillen in moeilijkheidsgraad tussen de examens van dit jaar en de referentie-examens.

Omdat de lengte van de scoreschaal geen vast gegeven meer is, worden verschillen uitgedrukt in een procentuele score (ook wel p-waarde genoemd; zie bijlage 1). Een procentuele score van 50 wil zeggen dat de kandidaten gemiddeld genomen de helft van het maximum hebben behaald.

Tabel 9 Vergelijking tussen de referentie-examens en de examens 2002 en de afstand C/D

Vak	verschil	Afstand C/D in cijferpunten
Duits D	2% moeilijker	0,8
Engels D	even moeilijk	1,3
Frans D	14% makkelijker	0,9
Muziek D	9% makkelijker	0,8
Nederlands leesvaardigheid D	even moeilijk	0,6
Duits havo	1% makkelijker	
Engels havo	2% moeilijker	
Frans havo	5% moeilijker	
Duits vwo	4% moeilijker	
Engels vwo	1% moeilijker	
Frans vwo	3% makkelijker	

De normvergelijking vindt bij deze examens plaats direct na de afname van het examen. Dit betekent dat eventuele verschillen in moeilijkheidsgraad niet vooraf gereduceerd worden, maar pas achteraf door vaststelling van een juiste normeringsterm gecompenseerd worden. Tabel 8 laat zien in hoeverre zich verschillen in moeilijkheidsgraad voordeden tussen het examen 2002 en het referentie-examen. In de regel wordt bij de constructie gemikt op een gelijke moeilijkheidsgraad als die van het referentie-examen.

Alhoewel de normvergelijking is uitgevoerd met oudere referentie-examens is kort voor de normering besloten om 2001 als referentie-norm te kiezen. Deze keuze hield mede verband met de veranderingen die sinds 2000 en 2001 in de MVT examens zijn doorgevoerd zoals het opnemen van eindterm 1 en het gebruik van een woordenboek. De CEVO heeft voor zover mogelijk bij deze examens de normeringstermen zodanig vastgesteld dat deze equivalent zijn aan de in 2001 gestelde eisen.

Bij het vbo/mavo vindt indirect ook normvergelijking plaats voor de C-examens van de genoemde vakken. De normvergelijking vindt hier plaats via de afstand ten opzichte van het D-examen. Tabel 8 geeft de afstanden tussen de C- en D-examens bij de vastgestelde normeringstermen. Deze afstanden mogen variëren tussen 0,5 en 1,5 en zijn gemiddeld genomen ongeveer 1,0 cijferpunt groot.

Alleen bij Nederlands is het verschil relatief klein. Nadat in 1996 was geconstateerd dat er bij het examen Nederlands leesvaardigheid een te geringe afstand bestond tussen C en D, wordt sinds 1997 een iets hogere moeilijkheidsgraad beoogd voor het D-niveau. Voor dit vak is de lat dus bewust iets hoger gelegd om een betekenisvolle afstand tussen de C- en D-examens te verkrijgen. Desondanks was het D-examen 2002 even moeilijk als het referentie-examen en dus toch makkelijker dan beoogd met als gevolg nog steeds een kleine afstand tussen het C- en D-examen.

Aangezien de normvergelijking een aantal jaren achtereenvolgend wordt met hetzelfde referentie-examen, zijn de populaties uit deze jaren ook onderling vergelijkbaar.

Het systeem van normvergelijking maakt het – onder bepaalde aannamen – mogelijk om uitspraken te doen over ontwikkelingen in de populaties over de jaren heen.

De staafdiagrammen hieronder brengen de ontwikkeling in de populatie in beeld. Op de verticale as zijn de gemiddelde procentuele scores afgezet. De kolommen voor de verschillende jaren geven de gemiddelde procentuele scores aan die de opeenvolgende populaties op hetzelfde referentie-examen behaald zouden hebben.

In de diagrammen voor de moderne vreemde talen zijn de gegevens van vbo/mavo-D, havo en vwo opgenomen. Voor vbo/mavo C geldt ongeveer hetzelfde als voor D, omdat deze examens via de overlap aan de D-examens gekoppeld zijn.

Het zij uitdrukkelijk opgemerkt dat deze ontwikkeling slechts geconstateerd wordt en dat de verzamelde informatie niets zegt over de feitelijke oorzaken die daaraan ten grondslag liggen. Daarnaast dient nog te worden opgemerkt dat de methodiek van normvergelijking

veronderstelt dat zowel het referentie-examen als het examen 2002 precies dezelfde eigenschap meten. Er wordt met andere woorden geen rekening gehouden met:

- veranderingen in de examens zoals de opname van open vragen in het examen en de toetsing van eindterm 1 (extensief lezen van teksten);
- veranderingen in de examencondities zoals bijvoorbeeld het gebruik van het woordenboek (voor vbo/mavo en havo vanaf 2000 en voor vwo vanaf 2001).

Uiteraard compliceren dit soort veranderingen de interpretatie van de gegevens. De vergelijking van de mvt-examens is uitgevoerd in een afnameconditie waarbij geen gebruik mocht worden gemaakt van een woordenboek. Bij de afname van het examen zelf was dat wel toegestaan. Bij de vergelijking van de score van de populatie 2002 met die van eerdere jaren dient men dus wel te bedenken dat het (al dan niet vermeende) voordeel van het woordenboekgebruik reeds in die score is verdisconteerd. De vergelijking van het prestatieniveau van de populaties van de elkaar opvolgende jaren is dan ook mede door deze complicaties enigszins speculatief van aard.

Geringe schommelingen in het prestatieniveau van jaar tot jaar kunnen zowel door verschillen in prestatieniveau als door de meetprocedure veroorzaakt zijn.

Diagram 1 *Prestatieniveau van de opeenvolgende populaties Duits uitgedrukt in een score op de referentietoets*

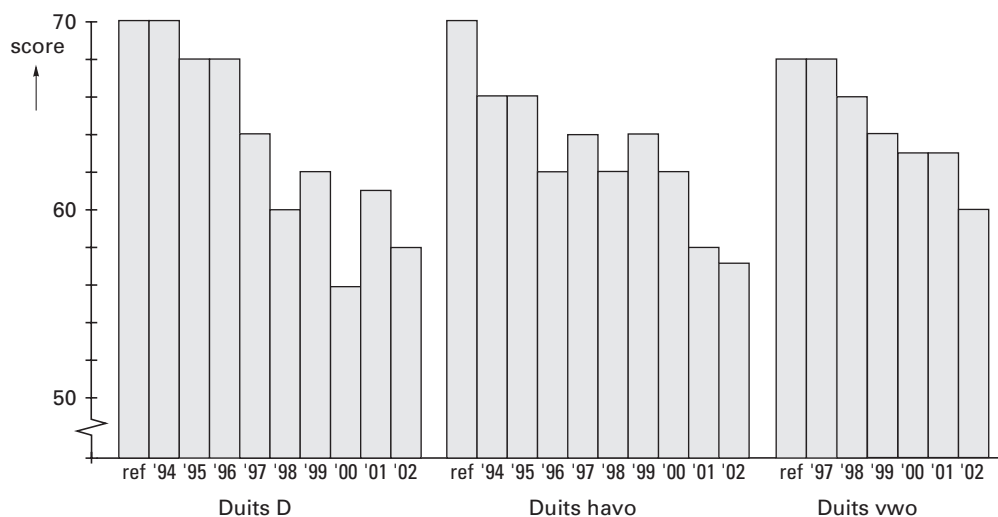


Diagram 2 *Prestatieniveau van de opeenvolgende populaties Engels uitgedrukt in een score op de referentietoets*

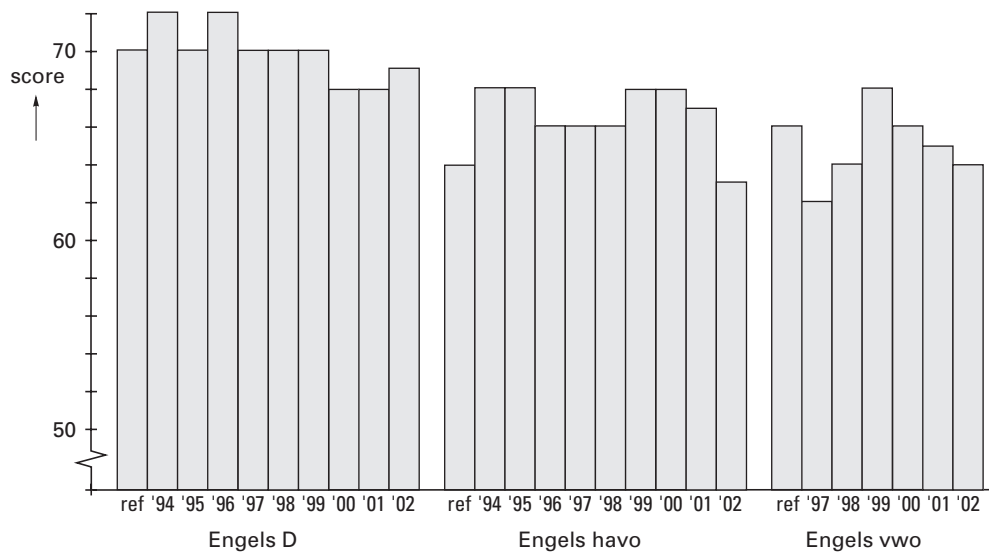


Diagram 3 Prestatieniveau van de opeenvolgende populaties Frans uitgedrukt in een score op de referentietoets

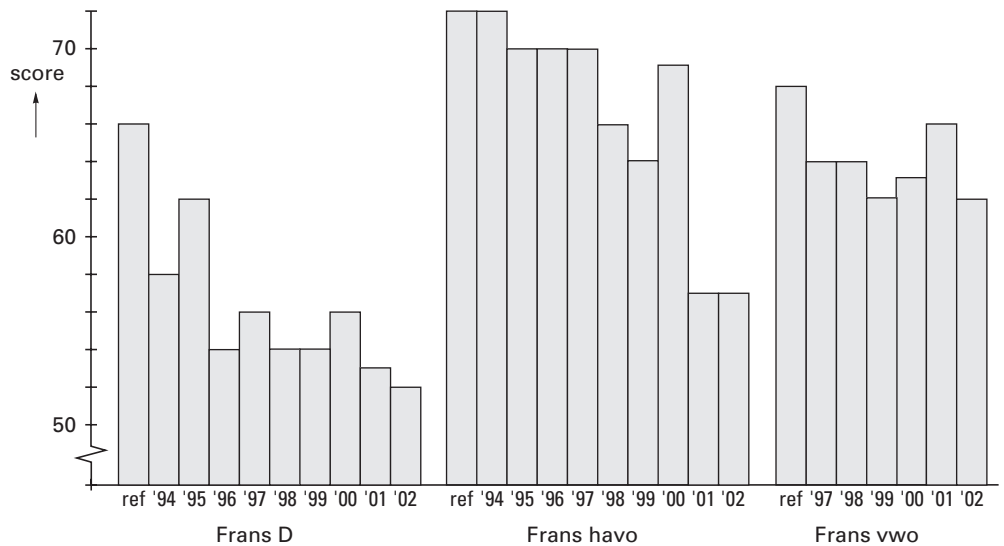
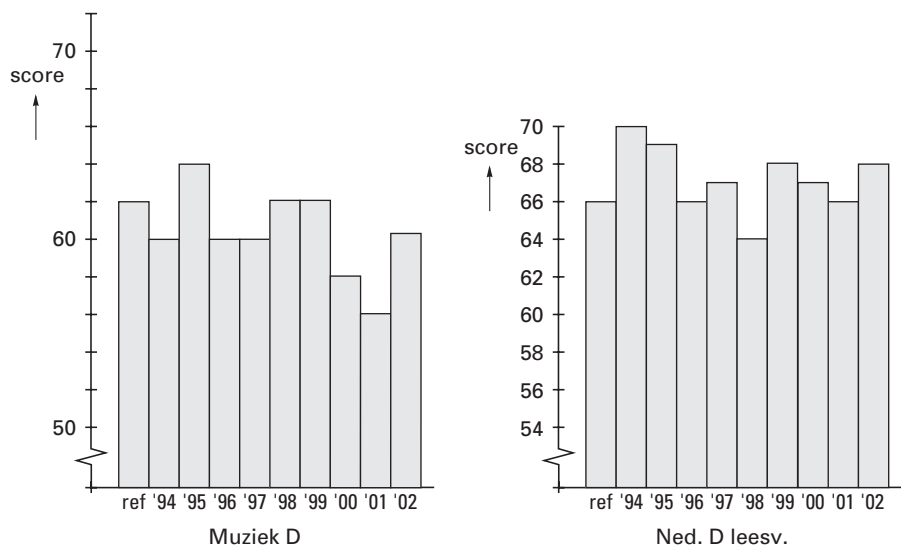


Diagram 4 Prestatieniveau van de opeenvolgende populaties muziek en Nederlands vbo/mavo-D uitgedrukt in een score op de referentietoets



Normvergelijking 'vooraf'

Voor havo en vwo zijn in 2002 de examens wiskunde A en B, natuurkunde, scheikunde, en biologie vwo en economie 1,2 en Nederlands havo vastgesteld met gebruikmaking van normvergelijgingsgegevens. Dit betekent dat er bij de samenstelling van de examens naar gestreefd is dat ze grotendeels dezelfde moeilijkheidsgraad zouden hebben als een door de CEVO gekozen referentie-examen met een door de CEVO vastgestelde referentienorm. Als de procedure voor het gehele examen is toegepast, kan voor deze examens de normeringsterm van 1,0 aangehouden worden. Bij veel van de gepreteste examens was dit echter niet het geval. In tabel 10 zijn de resultaten van 2002 voor deze vakken bij elkaar gezet. Bij de invoering van de vernieuwde tweede fase heeft een uitgebreide oriëntatie plaatsgevonden m.b.t. de consequenties daarvan voor de continuering van de normvergelijgingsprocedures. Daarbij deed zich de paradox voor dat juist in een complexe en nieuwe situatie referentiegegevens over nieuwe examens zeer wenselijk zijn maar dat vanwege de vele veranderingen in het onderwijs het maken van valide vergelijkingen buitengewoon gecompliceerd is. Bij het continueren van de procedures is het uitgangspunt geweest dat er weliswaar minder strikte vergelijkingen gemaakt zouden kunnen worden maar dat getracht zou moeten worden zo relevant mogelijke gegevens te verzamelen over het niveau en de maakbaarheid van de toekomstige examens.

Tabel 10 Aantal kandidaten, N-termen, gemiddelde cijfers en percentages onvoldoendes examens 2002

Vak	aantal kandidaten	N-term	gemiddeld cijfer	percentage onvoldoendes
Nederlands havo	1234	1,1	6,6	19
Wiskunde A1,2 havo	2064	1,8	5,8	41
Wiskunde A 1,2 vwo	2127	1,4	6,3	29
Wiskunde B1,2 havo	1987	1,8	6,4	27
Wiskunde B 1,2 vwo	1900	0,8	6,6	24
Natuurkunde 1,2 havo	1941	0,8	6,7	13
Natuurkunde 1,2 vwo	1859	0,8	7,1	8
Scheikunde havo	1839	0,8	6,4	19
Scheikunde 1,2 vwo	1976	1,0	6,8	18
Biologie 1,2 vwo	1784	0,9	6,3	19
Economie 1,2 havo	2051	1,2	5,7	43

1.9 Vergelijking van deelvak- en totaalvak-examens en -populaties

In tabel 12 worden de vergelijkingen van deelvak- en totaalvakpopulaties voor de vakken wiskunde A en B, natuurkunde en scheikunde weergegeven. Zowel voor vwo als voor havo geldt dat de deelvakpopulaties aanmerkelijk lager blijken te scoren op de overlap. Hierbij dient te worden aangetekend dat deelvak en totaalvak gebaseerd zijn op een ander programma, verschil in studielast kennen en op verschillende vervolgopleidingen voorbereiden.

Tabel 12 Vergelijking van totaalvak- en deelvakpopulaties, totaalvak- en deelvakexamens en de normering

Vak	schooltype	verschil in populaties	verschil in moeilijkheidsgraad	afstand na normering in cijferpunten
Wiskunde A 1,2 / Wiskunde A 1	vwo	10	8	1,0
Wiskunde B 1,2 / Wiskunde B 1	vwo	10	8	0,7
Natuurkunde 1,2 / Natuurkunde 1	vwo	9	-2	0,0
Scheikunde 1,2 / Scheikunde 1	vwo	6	-5	-0,1
Wiskunde B 1,2 / Wiskunde B 1	havo	13	8	0,9
Natuurkunde 1,2 / Natuurkunde 1	havo	7	7	0,6

Bij vier van de zes vergelijkingen blijkt het heelvakexamen moeilijker te zijn dan het deelvak examen. Bij scheikunde en natuurkunde vwo blijkt het deelvak-examen (iets) moeilijker te zijn. Kijken we naar het resultaat na de normering dan worden voor de meeste vakken aan de heelvakkers de hoogste eisen gesteld. Alleen bij scheikunde en natuurkunde vwo zijn de prestatie-eisen om aan een voldoende te komen voor deel- en heelvakkers nagenoeg gelijk.

1.10 Vergelijking van reguliere en experimentele examens en -populaties

Voor de vakken Engels havo, Biologie havo en Natuurkunde 1,2 zijn naast de reguliere examens ook op beperkte schaal experimentele examens afgenomen. In § 2.2 wordt de inhoudelijke kant daarvan nader belicht. Hier wordt ingegaan op de resultaten. In alle gevallen bestond er een aanzienlijke overlap in vragen tussen het reguliere en experimentele examen. Op basis van deze overlap worden de populaties, de examens en de normering op dezelfde wijze als in de beide voorgaande paragrafen vergeleken. Tabel 13 geeft de resultaten van deze vergelijking weer.

Tabel 13 *Vergelijking van populaties, examens en normering regulier / experimenteel*

Vak	schooltype	verschil in populaties	verschil in moeilijkheidsgraad	afstand na normering in cijferpunten
Engels / Engels complex *	havo	3	-4	0,4
Biologie / Biologie complex	havo	0	-6	0,1
Natuurkunde 1,2 / Moderne Natuurkunde	vwo	-1	-5	0,0

* vergelijking gebaseerd op mc-vragen

In alle gevallen bleken de experimentele examens iets moeilijker te zijn dan de reguliere examens. Qua prestatieniveau ontlieden de populaties elkaar niet veel. Bij Engels blijken de kandidaten die het reguliere examen hebben afgelegd op de meerkeuzevragen die voor beide examens identiek zijn iets hoger te scoren. Voor de andere examens scoren de populaties nagenoeg gelijk op de overlap.

Het is uitdrukkelijk niet de bedoeling dat er bij de experimentele examens hogere eisen worden gesteld dan bij de reguliere examens. Daarom wordt er voor het verschil in moeilijkheidsgraad bij de normering gecompenseerd. Voor biologie en natuurkunde worden na de normering ook nagenoeg gelijke eisen gesteld aan de kandidaten van het reguliere en het experimentele examen. Alleen bij Engels ligt de vergelijking wat moeilijker. De overlap bestaat weliswaar uit dezelfde vragen maar deze zijn niet in beide condities op dezelfde manier aangeboden. Wanneer er geen rekening gehouden wordt met dat verschil wordt het experimentele examen iets soepeler genormeerd.

1.11 Niveaoverschillen C- en D-examens

De C- en D-examens worden zo geconstrueerd dat er een redelijk verschil in moeilijkheidsgraad tussen beide examens te verwachten is. Een verschil in moeilijkheidsgraad wordt expliciet in de constructieopdracht beoogd. De examenprogramma's bieden voor dit verschil in niveau bij de meeste vakken weinig houvast. De examenprogramma's C en D zijn geheel of vrijwel geheel gelijk. De examenconstructeurs moeten op basis van hun ervaring met de C-kandidaten en de D-kandidaten in de verschillende schooltypen de verschillende opgaven maken.

Ondanks een zorgvuldige constructie waarbij de mogelijkheden om een niveaoverschil aan te brengen goed benut worden, kan men er vooraf niet zonder meer zeker van zijn dat een niveaoverschil ook echt gerealiseerd is. In de meeste C- en D-examens is de mogelijkheid ingebouwd om dit na afname van de examens te controleren. Alle examens behalve Nederlands schrijfvaardigheid kennen een overlappedeelte. In de C- en D-examens zijn tien à twaalf vragen van een gemiddelde moeilijkheidsgraad geheel gelijk. Dit overlappedeelte maakt het mogelijk het niveaoverschil na normering in termen van cijferpunten te kwantificeren.

Tijdens de normeringvergaderingen zijn deze gegevens over het niveaoverschil beschikbaar. Als voor een vak het aanbrengen van een niveaoverschil bij de constructie van de examens in onvoldoende mate is gelukt, kan zodoende door de vaststelling van de normeringsterm van het C- en/of D-examen alsnog een 'niveaoverschil' gecreëerd worden. Bij de normering gaat de CEVO ervan uit dat het niveaoverschil niet groter mag zijn dan 1,5 cijferpunt en in de meeste gevallen niet kleiner dan 0,5 cijferpunt.

Tabel 14 biedt een overzicht van de verschillen tussen de C- en D-populaties, -examens en -normering.

In de kolom 'verschil in populaties' wordt het verschil in procentuele score op het overlapdeel tussen de C- en D-populatie aangegeven. Zoals te verwachten blijken de D-kandidaten over de gehele linie sterker te presteren. Voor Muziek, Spaans, Turks en Arabisch worden geen resultaten gepresenteerd omdat de vergelijkingen op een te beperkt aantal kandidaten zou moeten worden gebaseerd. Daardoor zouden de uitkomsten te gevoelig zijn voor toevalligheden.

De verschillen in moeilijkheidsgraad liggen voor alle vakken in de verwachte richting. Voor enkele vakken zijn de verschillen echter klein te noemen. In de ideale situatie zijn de verschillen in examens substantieel, stabiel en zodanig dat ze goed zijn afgestemd op het verschil in prestatieniveau tussen de C- en D-populatie.

Bij Nederlands, Engels en natuurkunde zien we echter dat het verschil tussen de populaties veel groter is dan het verschil in moeilijkheid tussen de beide examens.

In de kolom 'afstand na normering' zijn de afstanden tussen de C- en D-examens weergegeven voor de vastgestelde normeringstermen. De afstand geeft aan hoeveel cijferpunten een kandidaat met een gemiddelde score op het D-examen meer behaald zou hebben als hij het C-examen had afgelegd. Tijdens de normeringvergaderingen is bekend wat de consequenties van verschillende normeringstermen zijn voor de afstand tussen het C- en D-examen. Bij alle vakken is sprake van een correcte afstand, dat wil zeggen dat de afstand ligt tussen 0,5 en 1,5 cijferpunt.

Tabel 14 Vergelijking van de C- en D-populaties, de C- en D-examens en de normering

Vak	verschil in populaties	verschil in moeilijkheidsgraad	afstand na normering in cijferpunten
Nederlands, leesvaardigheid	10	1	0,6
Frans	6	5	0,9
Duits	8	13	0,8
Engels	24	12	1,3
Spaans	-	-	-
Turks	-	-	-
Arabisch	-	-	-
Wiskunde	22	17	1,5
Natuurkunde	13	6	1,0
Scheikunde	11	6	0,7
Biologie	10	8	0,9
Economie	10	6	0,5
Aardrijkskunde	7	5	0,8
Geschiedenis en staatsinrichting	9	6	0,6
Maatschappijleer	9	11	0,8
Muziek	-	-	-
Tekenen	9	9	0,8
Handenarbeid	10	7	0,9
Textiele werkvormen	-	-	-

1.12 Verschillen tussen de resultaten van jongens en meisjes

Jongens en meisjes presteren op de meeste examens verschillend. De gemiddelden van de jongens zijn vaak hoger dan die van meisjes. De examens Nederlands schrijfvaardigheid, filosofie en kunstbeschouwing worden meestal door meisjes beter gemaakt. De examens moderne vreemde talen en de examens exacte vakken laten vaak een hoger gemiddeld cijfer voor jongens zien. In tabel 15 zijn alle significante verschillen tussen de gemiddelden van jongens en meisjes opgenomen, uitgedrukt in een percentage van de maximumscore die op elk examen te behalen was. De streepjes in de tabel geven examens aan waarbij zich geen significant verschil voordeed. De gearceerde vlakken geven aan dat een vak niet in dat bepaalde schooltype wordt geëxamineerd.

Tabel 15 Significante verschillen in p²-waarde tussen jongens en meisjes

	vwo oud		vwo nieuw		havo nieuw		vbo/mavo-D		vbo/mavo-C	
	meisjes beter	jongens beter	meisjes beter	jongens beter	meisjes beter	jongens beter	meisjes beter	jongens beter	meisjes beter	jongens beter
Nederlands (leesv.)			3		3		-	-		3
Nederlands opstel	-	-					-	-	-	-
Nederlands fs							7		7	
Frans (1,2)				4		2	-	-	3	
Duits (1,2)				4		2		2		2
Engels			-	-		3		4		1
Spaans (1,2)			-	-	-	-	-	-	-	-
Russisch (1,2)			-	-	-	-				
Turks			-	-	-	-	-	-	-	-
Arabisch			-	-	-	-	-	-	-	-
Latijn			7							
Grieks			3							
Wiskunde (A1) (A)	-	-	-	-				5		4
Wiskunde A 1,2			2			3				
Wiskunde B	-	-								
Wiskunde B1			2		2					
Wiskunde B1,2			-	-	-	-				
Natuurkunde (1)			-	-	-	-		5		3
Natuurkunde (1,2)				2		4				
Scheikunde (1)		6	3			1		3		3
Scheikunde 1,2			-	-						
Biologie (1,2)			-	-		3		2		1
Economie (I) (1, 2)				3		3		2	-	-
Econ.II/handel/M&O	-	-		1		4				
Economie 1				7		8				
Gesch.en staatsinr.			-	-	-	-		3		4
Aardrijkskunde	-	-		3		5		4		5
Maatschappijleer			-	-	-	-	-	-		2
Filosofie	-	-	-	-	7					
Muziek			-	-	-	-		3	-	-
Tekenen (tehatex)			2		-	-	-	-		3
Handenarbeid					-	-	2		-	-
Textiele werkvormen					-	-	-	-	-	-
CKV 2			-	-	4					

De verschillen zijn het grootst bij economie 1 havo. De examens wiskunde, natuurkunde, scheikunde, biologie en economie gaven in het verleden bijna altijd wel voor een of meer van de schooltypen significante verschillen in het voordeel van de jongens te zien. Dat patroon was de laatste altijd redelijk stabiel. Dit jaar zien we echter dat het patroon deels doorbroken wordt. Op de examens wiskunde A1,2 en B1 en scheikunde 1 scoren meisjes significant hoger. Voor vbo/mavo C en D en in iets mindere mate voor havo vertoont het patroon nog steeds grote overeenkomst met de resultaten van 2001 en voorgaande jaren.

Over de oorzaken van de verschillen valt weinig met zekerheid te zeggen. Uit onderzoek blijkt dat meisjes gemiddeld jonger zijn wanneer ze examen doen; ze lopen minder studievertraging op en jongens verlaten de school vaker vóór het eindexamen dan meisjes. Meisjes worden vaker naar 'hogere' schooltypen doorverwezen dan jongens. Bij een gelijke score op de eindtoets basisonderwijs worden meisjes meer toegelaten tot moeilijker geachte schooltypen dan jongens.

Behalve deze mogelijke oorzaken, die met de schoolloopbaan te maken hebben, kunnen oorzaken misschien ook gevonden worden in kenmerken van examens zelf. Meisjes scoren

hoger op traditioneel vrouwelijke onderwerpen, zoals huishouden, opvoeding en onderwerpen als literatuur, kunst, intermenselijke relaties en gevoelens. Jongens scoren hoger op 'typische mannenonderwerpen' zoals oorlog, misdaad, politiek, sport en op economische, financiële en natuurwetenschappelijke onderwerpen. Overigens moet benadrukt worden dat als het te toetsen leerdoel aan het examenprogramma ontleend is, de toetsing zonder meer gerechtvaardigd is. Voor alle kandidaten geldt immers hetzelfde examenprogramma. Van de examenconstructeurs mag gevraagd worden rekening te houden met de verschillende interesses van jongens en meisjes als het factoren betreft die niet essentieel zijn voor het te toetsen leerdoel, zoals de teksten bij talenexamens en het uitgangs- of contextmateriaal in overige examens. Hoewel bij de selectie van teksten en contexten hier expliciet rekening mee wordt gehouden, bleken de verschillen in resultaten van jongens en meisjes over een reeks van jaren toch redelijk consistent te zijn. De vraag is of de verandering daarin die we dit jaar bij het vwo hebben gesignaleerd een incident is of een trendbreuk zal blijken.

1.13 De examens tweede tijdvak

Op 19 juni 2002 (en voor enkele vakken op 21 juni) werd het tweede tijdvak afgenomen. De deelname aan het tweede tijdvak staat voor alle kandidaten voor één vak open. Een ieder die meent dat hij of zij voor één vak een beter cijfer zal kunnen behalen, mag aan het tweede tijdvak deelnemen.

Alleen kandidaten die tijdens het eerste tijdvak ziek waren, mogen in twee vakken examen doen. Wie in meer dan twee vakken nog examen moet doen, kan in augustus, in het derde tijdvak, het examen afronden.

De scholen is gevraagd voor elk examen maximaal van vijf kandidaten (de in alfabetische volgorde eerste vijf) de scores van het tweede tijdvak op speciaal daarvoor toegezonden optisch leesbare formulieren aan te strepen, samen met het cijfer dat de kandidaat voor het eerste-tijdvak-examen behaalde, en deze formulieren op te sturen naar de Citogroep. Voor de scholen die aan het experiment WOLF2002 meededen gold een vergelijkbare verplichting. De verwerkingstijd is zowel voor docenten als voor de Citogroep bijzonder kort omdat de resultaten nog voor de datum van de uitslag – zes dagen na de afname – gebruikt moeten kunnen worden. De scholen is gevraagd de formulieren twee dagen na de afname – op vrijdag 21 juni – al op te sturen. In de tabellen met resultaten is te zien dat deze snelle inzending en verwerking met succes zijn verlopen.

De gegevens werden verzameld om een vergelijking tussen het eerste en tweede tijdvak mogelijk te maken.

De resultaten van kandidaten die een onvoldoende op het eerste tijdvak behaald hebben, worden afgezet tegen hun resultaten op het tweede tijdvak. De aanname is dat het gemiddelde cijfer van deze kandidaten in het tweede tijdvak hoger moet zijn dan in het eerste tijdvak. Er zijn twee redenen om dit aan te nemen. In de eerste plaats zullen kandidaten zich nog speciaal kunnen richten op één vak waardoor leerwinst kan optreden. In de tweede plaats heeft de regressie naar het gemiddelde, die optreedt bij een herhaalde meting bij kandidaten die op het eerste tijdvak een lage score hebben behaald, een positief effect op de hoogte van de score. Dit regressie-effect kan ook begrepen worden als het effect dat veroorzaakt wordt doordat met name die kandidaten herkansen die vinden dat ze pech hebben gehad of, met andere woorden, dat door toevalsfluctuaties hun score op het eerste tijdvak een onderwaardering van hun niveau weergeeft. Is de gemiddelde score van deze groep kandidaten op het tweede tijdvak niet hoger dan die op het eerste tijdvak, dan moet dit veroorzaakt worden doordat het tweede tijdvak een hogere moeilijkheidsgraad heeft dan het eerste tijdvak. Een aanpassing van de normeringsterm wordt dan overwogen.

Bij de tweede-tijdvak-examens van 2002 is de definitieve normeringsterm van het tweede tijdvak voor de volgende examens aangepast ten opzichte van de 'voorlopige' normeringsterm:

Duits C	met 0,4	tot 1,0
Natuurkunde D	met 0,2	tot 1,1
Economie 1,2 havo	met 0,5	tot 1,7
Wiskunde B1,2 vwo	met 0,5	tot 1,3

Bij economie 1,2 hield de aanpassing verband met enkele minder geslaagde vragen. Bij de overige examens betrof het een compensatie voor een moeilijker tweede-tijdvak-examen. De verzameling van de gegevens betekent ook dat de resultaten van het tweede tijdvak voor een groot aantal examens beschikbaar zijn. In de tabellen 16 tot en met 19 worden deze resultaten gepresenteerd. Als toelichting bij deze tabellen het volgende:

- Niet alle examens zijn in de tabel opgenomen. Van een aantal vakken zijn geen gegevens verzameld, omdat vooraf reeds te voorzien was dat er te weinig kandidaten zouden zijn voor een zinvolle gegevensverzameling.
- De eerste kolom wordt het vak/examen vermeld.
- De tweede kolom betreft alle voor dat vak door de Citogroep verwerkte formulieren, hetgeen overeenkomt met het aantal kandidaten in de steekproef.
- In de derde kolom wordt het gemiddelde cijfer vermeld gebaseerd op de gehele steekproef van tweede-tijdvak-kandidaten.
- In de vierde kolom wordt vermeld welk percentage van de kandidaten in de steekproef op het tweede tijdvak een onvoldoende haalt. Over het algemeen scoort de populatie die deelneemt aan het tweede tijdvak gemiddeld lager dan de populatie die deelneemt aan het eerste tijdvak. Dit wordt veroorzaakt doordat in de populatie van het tweede tijdvak relatief meer zwakke kandidaten zitten. Voor de vbo/mavo-D-examens geldt nog specifiek dat hier een groot aantal profielverbeteraars aan deelneemt die in het eerste tijdvak op C-niveau examen hebben gedaan.

Zoals te verwachten liggen de gemiddelde cijfers van het tweede tijdvak lager dan die van het eerste tijdvak. Een groot deel van de tweede-tijdvakpopulatie bestaat immers uit herkansers die op het eerste tijdvak een onvoldoende resultaat hadden behaald. Van deze groep haalt (gemiddeld over vakken) ruim 40% alsnog een voldoende in het tweede tijdvak.

Tabel 16 Gemiddeld cijfer, aantallen kandidaten examens vbo/mavo C 2e tijdvak

Vak	steekproefgrootte	gemiddeld cijfer	% onvoldoendes
Nederlands, leesvaardigheid	337	6,2	31
Nederlands, functioneel schrijven	-	-	-
Nederlands opstel	-	-	-
Frans	61	6,3	15
Duits	291	5,1	55
Engels	774	5,2	57
Spaans	-	-	-
Turks	-	-	-
Arabisch	-	-	-
Wiskunde	942	5,1	60
Natuurkunde	192	5,7	42
Scheikunde	62	5,8	32
Biologie	190	5,4	55
Economie	141	5,3	55
Aardrijkskunde	58	5,5	41
Geschiedenis en staatsinrichting	23	6,8	4
Maatschappijleer	-	-	-
Muziek	-	-	-
Tekenen/handenaarbeid/text.werkv.	-	-	-

Tabel 17 Gemiddeld cijfer, aantallen kandidaten examens vbo/mavo D 2e tijdvak

Vak	steekproefgrootte	gemiddeld cijfer	% onvoldoendes
Nederlands, leesvaardigheid	486	6,2	21
Nederlands, functioneel schrijven	-	-	-
Nederlands opstel	-	-	-
Frans	294	5,5	49
Duits	858	5,8	36
Engels	1396	4,9	69
Spaans	-	-	-
Turks	-	-	-
Arabisch	-	-	-
Wiskunde	1728	5,0	66
Natuurkunde	514	5,0	72
Scheikunde	349	5,2	58
Biologie	473	5,1	67
Economie	703	5,1	66
Aardrijkskunde	196	5,2	58
Geschiedenis en staatsinrichting	98	5,9	28
Maatschappijleer	-	-	-
Muziek	-	-	-
Tekenen/handenaarbeid/text.werkv.	-	-	-

Tabel 18 Gemiddeld cijfer, aantallen kandidaten examens havo nieuwe stijl 2e tijdvak¹⁾

Vak	steekproefgrootte	gemiddeld cijfer	% onvoldoendes
Nederlands	388	5,4	49
Frans (1,2)	217	5,7	39
Duits (1,2)	199	4,7	76
Engels	480	5,5	52
Spaans (1,2)	-	-	-
Russisch (1,2)	-	-	-
Turks 1,2	-	-	-
Arabisch 1,2	-	-	-
Wiskunde A 1,2	499	5,6	41
Wiskunde B 1	189	6,3	25
Wiskunde B 1,2	186	6,2	23
Natuurkunde 1	77	5,8	31
Natuurkunde (1,2)	116	5,9	28
Scheikunde	256	5,4	47
Biologie	173	5,6	45
Economie 1	465	5,7	33
Economie (1,2)	761	5,8	33
Management & organisatie	44	6,0	32
Aardrijkskunde	339	5,2	57
Geschiedenis	476	5,7	34
Muziek	-	-	-
Tekenen	-	-	-
Handenaarbeid	-	-	-
Textiele werkvormen	-	-	-
Filosofie	-	-	-
Maatschappijleer	-	-	-
CKV 2 ²⁾	-	-	-

1) Geen gegevens over havo oude stijl

2) Geen Centraal Examen

Tabel 19 Gemiddeld cijfer, aantallen kandidaten examens vwo nieuwe stijl 2e tijdvak¹⁾

Vak	steekproefgrootte	gemiddeld cijfer	% onvoldoendes
Nederlands	376	5,7	40
Latijn	-	-	-
Grieks	-	-	-
Frans 1,2	212	5,0	73
Duits 1,2	212	6,0	34
Engels	472	6,0	35
Spaans	-	-	-
Russisch	-	-	-
Turks	-	-	-
Arabisch	-	-	-
Wiskunde A 1	52	4,8	73
Wiskunde A 1,2	198	4,5	75
Wiskunde B 1	173	4,6	74
Wiskunde B 1,2	243	4,8	67
Natuurkunde 1	201	5,7	43
Natuurkunde 1,2	64	5,4	55
Scheikunde 1	179	6,3	18
Scheikunde 1,2	95	6,0	29
Biologie 1,2	180	5,9	33
Economie 1	60	5,8	45
Economie 1,2	304	5,6	45
Management & organisatie	55	6,0	33
Aardrijkskunde	73	5,4	42
Geschiedenis en staatsinrichting	162	6,1	23
Maatschappijleer	-	-	-
Muziek	-	-	-
Tekenen/handenaarbeid/text.werkv.	-	-	-
CKV 2 ²⁾	-	-	-

1) Geen gegevens over vwo oude stijl

2) Geen Centraal Examen

2 Nieuwe examenvormen

In dit hoofdstuk wordt aandacht besteed aan veranderingen in de examens die dit jaar voor het eerst of die betrekkelijk recent gerealiseerd werden. In 2.1 worden de veranderingen gerelateerd aan de ontwikkeling van nieuwe examenprogramma's. In 2.2 wordt een beknopte beschrijving gegeven van het experiment computers en examens. In 2.3 worden de ervaringen met de vernieuwde examens besproken. Hierbij wordt met name ingegaan op die examens die mede n.a.v. de ervaringen uit 2001 zijn aangepast. Paragraaf 2.4 biedt een overzicht van de wisselende onderwerpen en in 2.5 wordt tot slot een korte vooruitblik gegeven op 2003.

2.1 Nieuwe examenprogramma's

In het kader van de vernieuwde tweede fase havo/vwo zijn enkele jaren geleden nieuwe examenprogramma's ontwikkeld en vastgesteld. Door Citogroep en CEVO zijn per vak syllabi opgesteld waarin de benodigde informatie wordt gegeven over de veranderingen in de centrale examens. Deze veranderingen zijn tweërlei. Enerzijds worden veranderingen in vakinhouden aangegeven, anderzijds zijn de nieuwe examenprogramma's sterker dan voorheen gericht op vaardigheden. Ook in de centrale examens worden (voor zover mogelijk binnen een schriftelijk examen) vaardigheidsvragen opgenomen. Dit jaar zijn voor het eerst bij alle vwo-kandidaten de centrale examens volgens de nieuwe programma's afgenomen. Vorig jaar was hier al bij 125 voorloop-scholen ervaring mee opgedaan. In dit verslag zullen de vernieuwde examens voor vwo in 2.3.2 aan de orde komen.

In 2002 werden de vbo- en mavo-examens voor de zesde maal afgenomen volgens nieuwe examenprogramma's. De programma's werden indertijd aangepast met het oogmerk een betere aansluiting met de basisvorming te realiseren.

Bij alle examens vbo en mavo werd meer aandacht besteed aan vaardigheden. Het examineren van vaardigheden in een centraal examen kent evenwel de beperking van de schriftelijke toets en van de beperkte afnametijd.

2.2 Computers en examens

Achtergrond en doelstelling

In de tweede fase van het havo/vwo en in het vernieuwde VMBO wordt in de leerplannen expliciet het aspect van ICT-vaardigheden genoemd. Bij de CEVO leefde de wens in 2002 de eerdere experimenten met het gebruik van de computer bij de afname van centrale examens te intensiveren en te gebruiken als een start voor een beleidsplan voor de jaren 2003 tot en met 2005. Daartoe heeft zij het Cito gevraagd om opgaven te ontwikkelen voor twee vakken op havo-niveau, die geschikt zijn voor afname met een computer in een examensituatie. Het doordenken van de opzet van dit experiment riep veel vragen op met betrekking tot de organisatie rondom de afname van de examens. De antwoorden op deze vragen kunnen alleen verkregen worden door daadwerkelijk deze opgaven in een examen op te nemen. Vanuit die gedachte is een aantal scholen per vak aangezocht om aan dit experiment mee te werken.

Opzet

Het ging in dit experiment om twee vakken op havo-niveau: Engels en biologie. De experimentele afnames beperkten zich tot het eerste tijdvak. De wijze waarop de computer bij de afname van het examen werd ingezet was voor de onderscheiden vakken zeer verschillend.

Engels

Voor Engels werd de computer ingezet om het examen geheel met behulp van de computer af te nemen. De teksten en de vragen werden via de computer aangeboden. De leerling-antwoorden, zowel op open als gesloten vragen werden in de computer of op het netwerk

opgeslagen. De open vragen werden door de docent op de computer gescoord en vervolgens werden de scores naar de Citogroep verzonden. De scholen ontvingen van de Citogroep vervolgens een rapport van de resultaten. De wijze waarop de computer bij het examen Engels is ingezet wordt aangeduid met de term CBT (**C**omputer **B**ased **T**esting). Aan deze afname namen zes scholen deel met in totaal ongeveer 340 leerlingen. Het CBT-examen Engels was voor het grootste deel inhoudelijk identiek aan het reguliere schriftelijke examen.

Biologie

Biologie is een examen dat wordt aangeduid met de term IMEX (Ict en **M**ultimedia in **E**Xamens).

De ongeveer 300 leerlingen van de negen deelnemende scholen maakten voor een deel van het examen (ongeveer 40%) gebruik van computerprogramma's en Multimedia toepassingen zoals film met geluid en interactieve java-applets. De overige 60% van het examen bestond uit vragen die ook in het reguliere examen voorkwamen. De opgaven werden op papier aangeleverd en de kandidaten schreven de antwoorden ook op papier.

Om de condities op de scholen zo veel mogelijk gelijk te houden, is een afnameprotocol ontworpen. Voor Engels is een uitgebreide handleiding naar de scholen gegaan.

De afname

De afname heeft plaatsgevonden op de scholen zelf. Alle scholen beschikten over een computernetwerk. Van tevoren is er een informatiebijeenkomst geweest.

De afname van het CBT-examen Engels op de scholen is goed verlopen. Er deden zich geen problemen voor met het vastlopen van de computers of het verdwijnen van antwoorden van kandidaten. Veel scholen hebben in verband met gebrek aan voldoende faciliteiten om alle kandidaten tegelijk te kunnen bedienen de afname in twee sessies gedaan en ook dat was goed te organiseren.

De afname Biologie complex is bijzonder goed verlopen. Bij biologie kwam het wel op een van de scholen een aantal keren voor dat er een PC vast liep en dat de kandidaat naar de reserve-computer verwezen moest worden. Alle kandidaten zijn in de gelegenheid geweest alle opgaven te beantwoorden die met de computer moesten worden gemaakt. De CEVO heeft toegestaan dat de kandidaten een half uur langer over het examen mochten doen.

Daar is met name bij biologie veel gebruik van gemaakt.

Gezien de kwetsbaarheid van computers in een schoolsituatie (veel handen aan de knoppen) is één reserve-computer per 10 kandidaten zowel bij Engels als bij biologie voorgeschreven.

De scholen zijn bereid om energie in zo'n experiment te steken, maar zij moeten wel de indruk hebben dat het inhoudelijk (bijvoorbeeld bij biologie) of procedureel (bijvoorbeeld correctiewerk uithanden nemen bij Engels) iets oplevert.

Op alle scholen was de systeembeheerder op de achtergrond aanwezig. Per lokaal waren minstens twee surveillanten aanwezig, waarvan er tenminste één 'computervaardig' was. Wanneer een kandidaat een technisch probleem had, kon de ene surveillant helpen zonder dat het overzicht over de rest van de kandidaten verloren ging.

De gegevens uit de evaluaties

Engels

De leerlingen zijn overwegend positief over deze manier van examineren. De docenten twijfelen aan de meerwaarde van de computer in het geval van puur leesvaardigheids-toetsing. Indien er meer geïntegreerd getoetst zou worden, zoals in een experiment bij Frans BB digitaaltoets, waarbij lees- en kijk- en luisterfragmenten geïntegreerd via de computer worden aangeboden dan wordt de mening een stuk positiever. De correctieprocedure van open vragen op de computer had te kampen met veel technische tegenslag en was derhalve geen succes.

Biologie

Zowel leerlingen als docenten waren na afname van het biologie examen enthousiast: IMEX biologie wordt als een verrijking van het reguliere examen beschouwd

Aan de leerlingen en de systeembeheerders is door middel van een korte enquête gevraagd naar hun eerste indrukken. Een eerste indruk is dat de meeste leerlingen het examen moeilijk(er) vonden (zie ook § 1.10).

De docenten hebben vooraf oefenmateriaal gekregen en zijn van mening dat zij de leerlingen goed hebben kunnen voorbereiden. De leerlingen vonden in meerderheid dat zij goed voorbereid zijn op dit examen. De betrokken docenten biologie vinden dat deze opgaven een

stap in de goede richting zijn van het gebruik van ICT bij examens. Het gebruik van Multimedia biedt naar hun mening goede mogelijkheden voor de toepassing van de computer in een examensituatie. Het is goed mogelijk gebleken de correctie op de 'oude voet' door de laten gaan.

De leerlingprestaties

De resultaten zijn in de normale procedure van examenverwerking meegenomen. Op basis van de overlapvragen met het reguliere centrale examen is een vergelijking gemaakt met de resultaten van de landelijke afnamegroep. Daaruit bleek dat de kandidaten uit de experimentele groepen op de overlap een vergelijkbaar of hoger prestatieniveau vertoonden. Bij biologie en bij Engels was het experimentele examen iets slechter gemaakt dan het reguliere examen (zie ook § 1.10).

Bij de normering van de experimentele examens is rekening gehouden met deze verschillen en zijn kandidaten indien nodig gecompenseerd met een iets hogere normeringsterm.

Voorlopige conclusies:

- het is voor scholen moeilijk en lastig een examensituatie te creëren in het computerlokaal of mediatheek, wel is er een duidelijke progressie ten opzichte van het eerste experiment in 2000;
- de te gebruiken programmatuur moet absoluut foutloos onder WIN9x, WINme, WIN2000 en WIN NT4.0 werken;
- een goede systeembeheerder is onontbeerlijk in een dergelijke situatie;
- het experiment heeft de nodige gegevens opgeleverd over de problemen die er rijzen bij de organisatie van een examenafname waarbij gebruik moet worden gemaakt van computerprogramma's;
- het is voor de Citogroep goed om ervaring op te doen met het maken van examenopgaven die alleen op te lossen of te beantwoorden zijn door gebruikmaking van een computerprogramma dat in het onderwijs al ingeburgerd is of snel ingeburgerd zal zijn;
- het opschalen van de experimenten met de computer in een examensituatie, zowel in de breedte (meer vakken) al in de diepte (meer scholen) is een gigantische opdracht en zal dus stapsgewijs moeten gebeuren.
- het is voor de Citogroep vakmedewerkers en de Cevo-vaksectieleden goed om ervaring op te doen aan de hand van zulke experimenten met het examineren van vaardigheden die te maken hebben met het gebruik van de computer.

2.3 Vernieuwingen in de examens 2002

2.3.1 Introductie open vraagvormen en woordenboekgebruik bij de MVT

In 2002 bevatten de examens moderne vreemde talen voor vwo, behalve de gebruikelijke meerkeuzevragen, voor de tweede maal ook open vragen. Het vbo/mavo en het havo zijn het vwo in dit opzicht een jaar eerder voorgegaan.

Voor alle examens mvt is, naar aanleiding van problemen bij het correctiemodel in 2001, extra zorg besteed aan de scoretoekenning, en aan de formulering van 'stellingvragen'. Dit om de correctoren niet in problemen te brengen. Op dit gebied kwamen dit jaar dan ook geen vragen binnen.

De reactie op open vragen is niet onverdeeld positief; bij sommige vragen geven docenten aan dat zij het correctiemodel te wijds vinden (vooral wanneer het antwoordmodel wordt vooraf gegaan door 'een antwoord met de volgende strekking ...'); bij andere vragen vinden zij het correctiemodel te weinig flexibel. Daarnaast werd opgemerkt dat veel van de open vragen ook in de meerkeuzevorm hadden kunnen staan. Van voornamelijk constructiegroepleden werd vernomen dat men constateerde dat de antwoorden van de kandidaten op de meerkeuzevragen soms ten onrechte als fout werden beoordeeld.

CEVO en Citogroep zullen n.a.v. bovenstaande bezien of voor 2003 nog aanvullende voorlichting over de (correctie van) open vragen in de mvt-examens nodig is.

Naar aanleiding van reacties van enkele scholen is dit jaar de opzet van de modelantwoordbladen aangepast. Met de nieuwe opzet wordt een besparing aan papier en correctietijd beoogd en hopelijk wordt ook de kans op fouten bij de correctie van de meerkeuzevragen er door verminderd.

Gebruik woordenboek

Van constructiegroepleden werd vernomen dat een aanzienlijk aantal kandidaten veelvuldig van het woordenboek gebruik (moest) maken. Bij de samenstelling van het examen wordt er echter van uitgegaan dat kandidaten slechts beperkt van het woordenboek gebruik hoeven te maken.

2.3.2 Nieuwe en vernieuwde examens voor het vwo

2.3.2.1 Culturele en kunstzinnige vorming 2

Culturele en kunstzinnige vorming 2 (CKV2) is het algemeen theoretisch gedeelte van het samenhangende programma CKV 2,3 binnen het profiel Cultuur en Maatschappij. Het CKV3-gedeelte wordt in het schoolexamen getoetst, CKV2 zou een centraal examen zijn, maar wordt tijdelijk als schoolexamen afgenomen, zowel op havo als op vwo.

Het CKV2-examen wordt afgenomen met behulp van een cd-rom. Hierop staan de opgaven en de contexten: leesteksten, afbeeldingen, muziek- en videofragmenten. Kandidaten maken het examen elk aan een eigen pc in hun eigen tempo. Wanneer nodig zetten zij een koptelefoon op. Zij schrijven hun antwoorden op proefwerkpapier.

Er waren examens voor twee tijdvakken geproduceerd en scholen konden deze bestellen om voor het schoolexamen te gebruiken. Er waren 136 vwo-scholen die dit deden. Systeembeheerders waren goed geïnstrueerd en voorbereid en de afnamen op de computers verliepen nagenoeg vlekkeloos.

De twee examens bestonden elk uit drie vragenblokken. De hofcultuur werd getoetst aan de hand van de ontwikkeling van de hofdans in Frankrijk en het hofleven in Versailles onder Lodewijk XIV. De negentiende eeuw kwam aan bod in het theater van Wagner, de Opéra in Parijs en het toneelstuk *Peer Gynt* van Hendrik Ibsen. Elektronische muziek van Varèse tot Zappa en wereldtentoonstellingen waren onderwerpen uit de twintigste eeuw, evenals de dansvoorstelling *Lara* van Krisztina de Chatel, gebaseerd op het computerspelletje *Tomb Raider*.

Docenten op de landelijke evaluatie toonden zich zeer tevreden over de inhoud van het examen. De kandidaten behaalden uitstekende resultaten.

Het schoolexamen CKV2 is op een vergelijkbare wijze geëvalueerd als bij de centrale examens gangbaar is. De resultaten van een steelproef van kandidaten zijn verzameld en geanalyseerd. Op basis daarvan heeft de CEVO de scholen geadviseerd om voor het vwo-examen een normeringsterm van 0,6 te hanteren.

2.3.2.2 Talen

Klassieke talen

Zowel het examen Grieks als het examen Latijn zijn positief ontvangen. Met name de vertaalopgave bij Latijn vond veel waardering; de vertaling is redelijk goed gemaakt; de gemiddelde P-waarde week nauwelijks af van die voor de vragen. Bij het Grieks waren de vragen duidelijk beter gemaakt dan de vertaling.

Dit jaar werd op alle scholen geëxamineerd volgens het nieuwe examenprogramma. Het programma vertoont een aantal wijzigingen. Zo vallen de eindtermen voor het centrale examen explicieter dan voorheen binnen de domeinen Taalreflectie, Tekstreflectie en Cultuurreflectie. Daarnaast is het jaarlijks voorgeschreven pensum van teksten gekozen vanuit een thematische invalshoek. Ook nieuw ten opzichte van de oude situatie is dat het aantal pagina's in het Grieks / Latijn te lezen tekst is teruggebracht van 25 naar 20. Daar tegenover staat het verplicht lezen van 50 voorgeschreven pagina's teksten in vertaling. Nog steeds heerst er onder de docenten bezorgdheid dat de sterke reductie van het aantal contacturen tot een daling van de vertaalvaardigheid zal leiden. Voor Latijn werd de N-term vastgesteld op 1,1 met 28% onvoldoendes en voor Grieks op 1,3 met 23% onvoldoendes.

Nederlands (leesvaardigheid)

Het examen Nederlands vwo (leesvaardigheid) 2002 voor oude en nieuwe stijl bestond net als het overgangsexamen in 2001 uit twee teksten die voorzien waren van een reeks van open en gesloten vragen en een samenvattingsopdracht.

Bij de open en gesloten vragen werd er, zij het nog steeds met terughoudendheid, een beroep gedaan op argumentatieve vaardigheden (een zelfstandig domein in het vernieuwde examenprogramma), maar op zo'n wijze dat ook de kandidaten oude stijl ermee uit de voeten moesten kunnen. De samenvattingsopgave had een geleid karakter. Het correctievoorschrift inzake incorrecte formuleringen en onjuist taalgebruik in de samenvatting kende voor het eerst een categorie "overige taalfouten" en een bepaling die het mogelijk maakte herhaalde taalfouten slechts één keer als fout mee te tellen.

De lengte van de scoreschaal bedroeg 49 scorepunten. Uitgangspunt bij de scoretoekenning was afzonderlijke antwoordelementen te laten corresponderen met één scorepunt.

Het examen 2002 werd, net als in 2001, genormeerd met een N-term van 1.0, hetgeen een geschat percentage onvoldoendes opleverde van 20%.

Een punt van zorg bij het examen Nederlands is de geringe betrouwbaarheidscoëfficiënt. De grote vraag is of dit een gevolg is van de gevolgde schattingprocedure of dat het een juiste weergave is van een gebrekkig onderscheidend vermogen van het examen als meetinstrument. Hiervoor is echter nader onderzoek vereist.

Moderne vreemde talen

Het examen oude stijl en het examen nieuwe stijl was voor alle moderne vreemde talen identiek. Dit betekent dat, in tegenstelling tot 2001, nu in alle examens ook eindterm-1 opgaven voorkwamen. Deze opgaven bestaan uit één vraag bij een vrij lange tekst en vergen het gericht zoeken naar informatie.

Bij een vrij groot aantal examens werd door de kandidaten geklaagd met betrekking tot de lengte ervan. Wellicht wordt dit veroorzaakt door de teksttoename die samenhangt met de toetsing van eindterm-1. Een andere oorzaak zou kunnen zijn dat het beantwoorden van open vragen meer tijd kost dan het beantwoorden van meerkeuzevragen. Op dit gebied zou nader onderzoek interessant zijn.

2.3.2.3 Maatschappijvakken

Aardrijkskunde

Het vwo-examen aardrijkskunde nieuwe stijl omvatte drie domeinen: Politiek en ruimte, Migratie en mobiliteit en Mens en milieu. De 45% overlap met het examen oude stijl was hier verdeeld over de laatstgenoemde twee onderwerpen.

Zo kon de vereiste overlap tussen oude en nieuwe stijl worden gerealiseerd. Die overlap is echter traditioneel getoetst, dus zonder gebruik te maken van nieuwe vraagtypen en zonder expliciet een appèl te doen op vaardigheden. Kandidaten oude stijl hebben de leerstof immers op traditionele wijze verwerkt en dus nauwelijks ervaring opgedaan met de vaardigheden zoals deze omschreven zijn in het nieuwe examenprogramma.

Waaruit blijkt het 'nieuwe' in de examens nieuwe stijl?

In de specifieke nieuwe-stijl-vragen zijn de volgende nieuwe vaardigheden/geografische werkwijzen getoetst waarbij atlasgebruik een centrale rol speelde:

- Informatievaardigheden: drie van de vier vraagtypen uit de syllabus, waarbij informatiebegrip en informatieselectie nogal vaak zijn getoetst. Er is slechts enkele keren naar de atlas verwezen, zonder dat er een beperking in het te selecteren aantal kaarten was aangebracht. Daardoor is tijdnood bij kandidaten voorkomen. Bij de andere vragen is de selectie ingeperkt tot maximaal 10 à 12 kaartbladen. Er is geen informatiebewerking getoetst, omdat dit ook tijdrovend kan uitpakken.
- Standpuntbepaling: er zijn argumentatievragen en een afwegings/betoogvraag opgenomen, standpuntbepaling (in enge zin) is niet getoetst.
- Onderzoeksvaardigheden: deze vraagtypen zijn niet getoetst.

De vijf geografische werkwijzen (gebieden vergelijken, relaties leggen, dimensies confronteren, wisselen van analyseniveau en veranderen van ruimtelijke schaal) zijn alle in het examen terug te vinden. Om kandidaten op het spoor van een correct antwoordgedrag te zetten is boven elke vraag of vragencluster de gewenste geografische werkwijze vermeld. Of dit tot een beter antwoordgedrag heeft geleid is echter niet uit de toets- en itemanalyse af te leiden. Uit het veld zijn veel reacties gekomen waarin gevraagd wordt om deze verwijzingen niet meer in het examen op te nemen omdat de kandidaten er door in verwarring zouden kunnen raken. Dit wekt wat bevreemding omdat de geografische werkwijzen de kern van het vak aardrijkskunde zijn en de kandidaten er dus vertrouwd mee dienen te zijn.

Bij de samenstelling van het vwo-examen 2002 is verder rekening gehouden met de ervaringen van afgelopen jaren waardoor:

- het toetsen van vaardigheden bewust wat beperkt;
- pure kaartselectie, met de gehele atlas als zoekdomein mondjesmaat is getoetst;
- bij andere kaartselectievragen het zoekgebied beperkt is.

Enkele conclusies uit de voorlopige examenanalyse:

- er bleek geen indicatie te zijn voor tijdnoed bij de afname;
- de examens nieuwe stijl zijn duidelijk veel meer geografisch dan de traditionele examens;
- het vwo-examen was te gemakkelijk.

Voor het vwo-nieuwe stijl examen werd de N-term vastgesteld op 0,8 met een percentage onvoldoendes van 17.

Economie 1,2 (totaalvak)

Het examen economie 1,2 vwo was het eerste examen voor het totaalvak economie dat door alle leerlingen in het profiel Economie en Maatschappij werd gemaakt. Daarnaast was er nog een kleiner aantal kandidaten uit de andere drie profielen, waarvoor Economie 1,2 een keuzevak was in het vrije deel. In dit examen waren de vernieuwingen verwerkt die in de Syllabus Economie voor de centrale examens in de tweede fase havo/vwo (juni 1998) waren aangekondigd en in een voorbeeldexamen waren uitgewerkt. Het examen had dientengevolge de volgende kenmerken:

- Vrijwel elke opgave begint met de beschrijving van een concreet maatschappelijk vraagstuk waarna enkele economische aspecten van dat vraagstuk aan de orde worden gesteld. Doel van deze opzet is een grotere aansluiting bij de beleevingswereld van de leerling te realiseren.
- De opgaven zijn korter dan bij economie gebruikelijk was. Door het aantal vragen per opgave te beperken, kunnen in een examen meer onderwerpen aan de orde worden gesteld. De verwachting is dat leerlingen daardoor meer thema's voorgelegd krijgen die hun belangstelling hebben.
- De (context)informatie is meer dan vroeger gebundeld en losgemaakt van de vragen. Daardoor wordt een groter beroep gedaan op de informatievaardigheid van de leerling. Het gaat daarbij vooral om het begrijpen, beoordelen, selecteren en bewerken van economische informatie.
- Er is nadrukkelijker dan voorheen aandacht voor het toetsen van standpuntbepalingsvaardigheden van de kandidaten.
- Er is een aparte opgave opgenomen die de afwegingsvaardigheid van de kandidaten toetst: de betoogvraag.

Voor economie 1,2 vwo werd de N-term vastgesteld op 1,5 met een percentage onvoldoendes van 24.

Geschiedenis

De eindexamens geschiedenis zijn vóór de tweede fase al sterk gemoderniseerd onder invloed van de sinds 1994 ingevoerde veranderingen, doorgevoerd door de WIEG (de Werkgroep Implementatie Eindexamens Geschiedenis). Binnen deze verandering nam de toetsing van vaardigheden een belangrijke plaats in.

De tweede-fase-vernieuwing leidde daarom niet tot een ingrijpende verandering in de centrale examens geschiedenis en staatsinrichting. Het examen nieuwe stijl is volkomen gelijk aan het oude-stijl-examen. Er zijn nog steeds twee onderwerpen per tijdvak, die elk in twee achtereenvolgende jaren getoetst worden en waar elk jaar één nieuw onderwerp bij zit. Een stelopdracht maakt deel uit van het examen. Argumentatie- en standpuntbepalingsvaardigheden komen daarbij aan bod. Informatievaardigheden worden getoetst aan de hand van authentiek bronmateriaal zoals teksten, tabellen, cartoons en foto's. In de vorm van een casus gebeurt dat met wat omvangrijker bronmateriaal.

Er wordt van uitgegaan dat ook in de tweede fase de stofomschrijving in ongeveer 2x27 studielasturen door de kandidaten zal worden doorgewerkt (overeenkomstig de traditionele 18-20 lesuren per onderwerp in de oude stijl).

Het examen was dit jaar makkelijker dan beoogd. De N-term is vastgesteld op $N = 0,8$ waardoor de kandidaten op een gemiddeld cijfer van 6,3 kwamen. Bij deze N-term scoorde 23% onvoldoende.

Over de omvang van het examen hebben ons dit jaar geen inhoudelijke klachten bereikt. Vanuit de vakvereniging, de docenten en de wetenschap was er nauwelijks kritiek op en zelfs

voorzichtige lof over de inhoud van dit examen. Wel werden kanttekeningen geplaatst bij de stelopdracht, die moeilijk te corrigeren bleek.

De vorig jaar ingezette pogingen de communicatie met de docenten te verbeteren worden voortgezet. Er zal weer een overleg plaatsvinden tussen de vakvereniging, Cito/CEVO en de voorzitters van de regiovergaderingen.

2.3.2.4 Exacte vakken

Wiskunde A1 en A1,2.

Het vak wiskunde A is in de vernieuwde tweede fase vervangen door wiskunde A1 (profiel C&M) en wiskunde A1,2 (E&M). *Consistentie met de (betere) formulering van wiskunde B)* Beide vakken worden afgesloten met een centraal examen.

Wiskunde A1,2

Het examenprogramma van wiskunde A1,2 is voor een belangrijk deel een voortzetting van wiskunde A. De leerstofgebieden Toegepaste Analyse, Toegepaste Algebra en Waarschijnlijkheidsrekening en statistiek komen ook in het nieuwe examenprogramma voor hoewel de naamgeving enigszins is aangepast: het betreft bij A1,2 de gebieden Toegepaste Analyse, Discrete wiskunde en Kansrekening en statistiek. Met name bij het gebied Discrete wiskunde is er een verandering te constateren: oorspronkelijk bevatte dit leerstofgebied de onderwerpen Discrete analyse, Discrete dynamische modellen, Lineair programmeren en Grafen en matrices. Het onderwerp Grafen en matrices, dat ook al onderdeel van het vak wiskunde A vormde, is echter uit de verzameling onderwerpen voor het centraal examen geschrapt.

Een belangrijke verandering betreft het gebruik van de grafische rekenmachine. De mogelijkheden van dit apparaat stellen de examenmakers in staat onderzoeksvaardigheden uitgebreider in examens te toetsen. Per slot van rekening betrof een van de belangrijke veranderingen van de Tweede Fase meer aandacht voor het verschijnsel vaardigheden. Bij wiskunde werd en wordt daar dus een belangrijke rol voor de grafische rekenmachine voor gereserveerd.

In de ervaringen met de eerste examens nieuwe stijl in 2001, gemaakt voor die scholen die al in 1998 met de invoering van de vernieuwde Tweede Fase in 4 VWO begonnen waren, is geen aanleiding gezien de examens 2002 aan te passen. Met andere woorden: toetsmatrijs en andere uitgangspunten in de examenmodellen waren voor de examens 2002 gelijk aan die van 2001.

Vorig jaar leek al geconstateerd te kunnen worden dat A1,2-leerlingen over minder algebraïsche vaardigheden lijken te beschikken dan hun wiskunde-A-voorgangers. Die tendens lijkt zich voort te zetten. Enkele vragen (vragen 7 en 14), die heel specifiek gericht waren op deze vaardigheden doen het minder dan verwacht. Als voorbeeld in dat kader moge vraag 7 dienen.

Bij deze vraag moest uitgegaan worden van een formule in de vorm
$$\frac{f_{\text{hout}}}{1 - f_{\text{hout}}} = 3,03 \cdot 0,96^t .$$

We zien in deze betrekking dat de variabele f_{hout} op directere wijze in de variabele t kan worden uitgedrukt. Bedoeling is dan dat deze directe vorm $f_{\text{hout}} = \dots$ opgesteld wordt, uitgaande van de gegeven formule. Aan de hand van de analyse kan geconstateerd worden dat 59% van de kandidaten voor deze vraag geen enkel punt (van de maximumscore 5) bleek te halen. Nog eens 30% van de kandidaten scoorde hier slechts 1 punt. Al met al geeft de beantwoording van deze vraag aan dat het aspect 'algebraïsche vaardigheden' in de toekomst meer aandacht in het onderwijs verdient dan nu het geval is, zeker gezien het feit dat dit aspect in het huidige examenprogramma een duidelijk aanwijsbare en herkenbare plaats heeft.

Een opgave die de veranderingen in het examenprogramma illustreerde, was de opgave Aardbeien. Bij deze opgave werd in een discreet dynamisch model (DDM) de fluctuatie van prijzen en aangeboden hoeveelheden van aardbeien bestudeerd. Deze opgave was een concretisering van het 'nieuwe' domein DDM, gecombineerd met een economische invalshoek waarvoor het examenprogramma van A1,2 ook uitdrukkelijk aandacht vraagt. De opgave bestond uit vier vragen waarvan de eerste drie standaard waren. Leerlingen die het onderwerp DDM bestudeerd hadden, zouden al heel snel, alleen al op grond van kennisaspecten, tot een redelijke score op deze vragen moeten komen, zo was het gevoelen van examenmakers en leden van de CEVO-vaksectie. De p'-waarden van deze vragen (56, 68 respectievelijk 67) gaven na afloop echter toch aan dat deze vragen moeilijker bleken dan

verwacht. De vierde vraag van deze opgave, vraag 20, was van een ander kaliber, iets dat ook op voorhand reeds verwacht was. Deze vraag was bewust in deze opgave opgenomen om daarmee te illustreren dat een opgave als deze, gebaseerd op een dergelijk dynamisch model, meer dan alleen maar standaardvragen kan oproepen. Bij deze vraag werd van leerlingen daadwerkelijk inzicht in het model gevraagd en werd er 'gesleuteld' aan specifieke, bij het model horende randvoorwaarden die eerder in de opgave als constant beschouwd werden. De p' -waarde van deze vraag ($p' = 22$) gaf aan dat dit voor heel veel leerlingen duidelijk een brug te ver was.

Aan de hand van de analyse is na afloop van het examen besloten voor een normeringsterm $N = 1,4$. Hierbij moet opgemerkt worden dat de vaksectie deze vaststelling in twee delen gesplitst wilde zien. Op grond van de achteraf gebleken moeilijkheid zou men tot een $N = 1,3$ zijn gekomen. Achteraf was men van oordeel dat het toekennen van één aftrekpunt bij vraag 12 niet terecht was waarvoor derhalve alsnog een incidentele aanpassing van 0,1 werd voorgesteld.

Wiskunde A1

Wiskunde A1 vormt een deelvak bij het vak wiskunde A1,2. Belangrijke zaken die niet in het A1-programma en wel in A1,2 zijn opgenomen zijn Differentiaalrekening met toepassingen, Lineair programmeren en, sinds de aanpassingen van juli 1999 van de staatssecretaris, ook het onderwerp Toetsen van hypothesen.

Voor het examen A1 van 2002 was vastgesteld dat er een overlap van circa 50% met het A1,2-examen moest zijn. Omdat de resterende 50% van de vragen, als gevolg van de deelvak/heelvak-structuur, betrekking heeft op onderwerpen die ook in het A1,2-programma zitten, was het zaak om deze 'profiel specifieke' opgaven niet op meer van hetzelfde neer te laten komen. Zo vinden we in dit deelvakexamen een opgave als 'Sparen' terug die niet in het A1,2-examen is opgenomen. Uiteraard zou dit wel tot de mogelijkheden behoord hebben omdat immers alle kennis en vaardigheden horend bij het programma A1 ook deel uit maken van het programma A1,2. In de opgave 'Sparen' wordt het probleem van een langetermijn-spaarrekening aan de orde gesteld waarbij het de bedoeling is op zeker moment, uitgaande van rente op jaarbasis, een bedrag van 10 000 euro op de rekening aan te treffen. Het probleem werd van verschillende kanten benaderd: er werd gevraagd naar een eenmalige beginstorting zowel als naar de jaarlijks constante inleg om te komen tot dit bedrag. De opgave 'Sparen' was een invulling van het subdomein 'Rijen' zoals dat in het examenprogramma voor wiskunde A1 is opgenomen.

Op basis van de analyses en uitgaande van het criterium dat bij dit deelvak het percentage onvoldoendes van vorig jaar bij de voorhoedescholen min of meer maatgevend zou moeten zijn voor het percentage onvoldoendes van 2002 is besloten tot een N-term van 1,6. Verder is, uitgaande van deze $N = 1,6$, de bij het A1,2-examen gekozen N-term $N = 1,4$ en de overlap tussen beide examens te constateren dat de gemiddelde A1,2-leerling die het A1-werk gemaakt zou hebben voor dat A1-werk een volle scorepunt hoger gescoord zou hebben. Anders gezegd: de A1,2-leerling met een cijfer 6,3 (gemiddelde cijfer uitgaande van $N = 1,4$) zou voor het A1-werk een 7,3 gehaald hebben (uitgaande van $N = 1,6$ aldaar). Een dergelijke afstand tussen beide examens doet recht, zo ervaren betrokkenen, aan het niveauverschil tussen beide vakken.

Wiskunde B1 en B1,2

Het vak wiskunde B is in de profielen vervangen door de vakken wiskunde B1 (profiel N&G) en wiskunde B1,2 (profiel N&T). Beide programma's zijn door het gebruik van de grafische rekenmachine en enkele veranderingen in het programma niet goed te vergelijken met het oude-stijl-programma. Daarom hebben het oude wiskunde-B-examen enerzijds en de wiskunde B1- en -B1,2-examens anderzijds geen overlap. Wiskunde B1 is een deelvak van B1,2. Deze twee nieuwe-stijl-examens vertonen een overlap van 50%, hetgeen overigens niet overeenkomt met het verschil in domeinen en studielasturen. Bij wiskunde B1 en B1,2 hoort het domein 'Statistiek en kansrekening' waardoor de examens iets meer een 'wiskunde A'-karakter krijgen. De domeinen 'Voortgezette analyse' en 'Voortgezette meetkunde' bij wiskunde B1,2 compenseren dat ruimschoots.

De resultaten van het wiskunde B1-examen waren goed, evenals die van het B1,2-examen. Het B1,2-examen is zelfs nog iets beter gemaakt als het B1-examen. Leerlingen die Wiskunde B1,2 volgen kunnen door de voortgezette analyse meer tijd aan de analyse besteden. Daardoor hebben de B1,2-kandidaten ondanks het gelijke aantal studielasturen op het examen toch een voorsprong. Wel valt op dat de grafische rekenmachine nog niet door iedere leerling wordt gebruikt wanneer dat nodig of zinvol is. Daarnaast lijkt niet elke leerling precies

te weten wanneer de rekenmachine wel en wanneer niet gebruikt mag worden. Dit zal de komende jaren hopelijk verbeteren als de grafische rekenmachine meer ingeburgerd raakt. Voor B1 werd de N-term vastgesteld op 0,8 met een percentage onvoldoendes van 26 en voor B1,2 ook op 0,8 met een percentage onvoldoendes van 24.

Natuurkunde

Het vak natuurkunde vwo is in de vernieuwde Tweede Fase opgesplitst in de vakken natuurkunde 1 (profiel N&G) en natuurkunde 1,2 (profiel N&T). Natuurkunde 1 is een deelvak van natuurkunde 1,2.

De vakinhoudelijke verschillen tussen het oude examenprogramma en het nieuwe programma van natuurkunde 1,2 zijn niet zo groot. Het nieuwe programma is meer dan voorheen gericht op het ontwikkelen van vaardigheden; ook is het de bedoeling dat er minder gerekend moet worden tijdens het examen.

De examenprogramma's natuurkunde 1 en natuurkunde 1,2 hebben ongeveer 50% overlap. In de examens is de overlap ietsje groter.

De deelvakkandidaten presteerden bij de overlapvragen duidelijk minder dan de heelvak-kandidaten (zie § 1.9). Desalniettemin bleek het specifieke deel van het deelvak-examen iets moeilijker te zijn waardoor het deelvak als geheel net iets moeilijker was dan het heelvakexamen.

De resultaten van de natuurkunde 1,2-kandidaten waren uitgesproken goed. Net als vorig jaar blijkt dat in de vernieuwde tweede fase de groep kandidaten met belangstelling voor Natuurkunde ten opzichte van de oude situatie wordt opgesplitst in een sterke N1,2-groep en een iets zwakkere N1-groep.

Voor Natuurkunde 1 vwo werd de N-term vastgesteld op 1,0 met een percentage onvoldoendes van 25 en voor natuurkunde 1,2 op 0,8 met een percentage onvoldoendes van 8.

Scheikunde

Voor het vak scheikunde vwo worden twee examens nieuwe stijl gemaakt: scheikunde 1 (profiel N&G) en scheikunde 1,2 (profiel N&T). Het programma voor scheikunde 1,2 lijkt nog het meest op het programma oude stijl. Toch bestaan er vrij grote verschillen: een flink aantal onderwerpen is vervallen, bijvoorbeeld entropie, mesomerie, elektronenformules, sterische hindering en de wet van Nernst. Daarvoor in de plaats hebben enkele onderwerpen een uitbreiding ondergaan (bijvoorbeeld brandstoffen en biochemie) en zijn ook enkele nieuwe onderwerpen aan het programma toegevoegd (bijvoorbeeld kringlopen van koolstof en stikstof, silicaten, infrarood- en massaspectrometrie). Verder ligt er een groter accent dan voorheen op het ontwikkelen van vaardigheden. Scheikunde 1 is een deelvak van scheikunde 1,2. Het examen scheikunde 1,2 bestond voor 50% uit dezelfde vragen als het examen scheikunde oude stijl. Het examen scheikunde 1 bestond voor 42% uit dezelfde vragen als het examen scheikunde 1,2.

De examens scheikunde nieuwe stijl onderscheiden zich vooral van het oude-stijl-examen doordat in de nieuwe-stijl-examens een groter beroep wordt gedaan op vaardigheden.

Het profielspecifieke gedeelte van het examen scheikunde 1,2 betrof vooral de chemische techniek, terwijl in het profielspecifieke gedeelte van het examen scheikunde 1 een aantal vragen voorkomt dat een sterk biochemisch karakter heeft.

Problemen als in 2001, toen in de nieuwe stijl examens een aantal vragen voorkwam over stof die niet in de leerboeken behandeld is, deden zich dit jaar niet voor. Wel werd door veel docenten als knelpunt genoemd dat de examens te lang en te "talig" zijn. Hierdoor zouden vooral veel scheikunde 1 kandidaten in tijdnood zijn gekomen. Uit de toets- en itemanalyse bleek overigens die tijdnood niet: naar het eind van het examen kwam score nul niet vaker voor dan aan het begin. Ook bleek dat bij geen van de vragen meer dan twee procent van de kandidaten de beantwoording ervan had overgeslagen.

Het examen scheikunde 1,2 bleek voor de kandidaten goed te maken. De kandidaten die scheikunde 1 hebben gevolgd, presteerden iets minder dan de kandidaten die scheikunde 1,2 hebben gevolgd. Dat blijkt uit de resultaten op de overlap tussen beide examens: kandidaten die scheikunde 1 hebben gedaan scoorden daarop gemiddeld 51% van de punten, terwijl de scheikunde-1,2-kandidaten gemiddeld 57% scoorden (zie ook § 1.9). Verder bleek dat het profielspecifieke deel voor de scheikunde 1 kandidaten ook moeilijker was dan het profielspecifieke deel voor de scheikunde 1,2 kandidaten.

Bij de normering is voor het examen scheikunde 1,2 een N-term van 1,0 vastgesteld. Met deze N-term is voor scheikunde 1,2 het gemiddelde cijfer 6,8 en krijgt 18% van de kandidaten een onvoldoende.

Voor scheikunde 1 is de N-term 1,3 geworden. Daarbij is het gemiddelde cijfer 6,2 en krijgt 23% van de kandidaten een onvoldoende.

2.4 Vakken met wisselende onderwerpen

In een aantal examenprogramma's van vooral de maatschappijgerichte vakken is een keuze gemaakt voor z.g. wisselende onderwerpen. Meestal twee jaar voor de afname van het examen, doch uiterlijk één jaar voor de afname, maakt de minister de onderwerpen bekend. De bekendmaking van de onderwerpen gaat meestal vergezeld van een beschrijving van de examenstof.

In het hiernavolgende zijn de examenonderwerpen die dit jaar geëxamineerd zijn voor de verschillende vakken weergegeven.

VWO

Latijn (nieuwe stijl) en Latijn (oude stijl)

Literair genre: epische poëzie; thema: Schuld en onschuld, kernauteur: Ovidius

Grieks (nieuwe stijl) en Grieks (oude stijl)

Literair genre: historiografie; thema: Hoogmoed komt voor de val, kernauteur: Herodotus

Geschiedenis (nieuwe stijl) en geschiedenis en staatsinrichting (oude stijl)

1. Nederland en Indonesië. Vier eeuwen contact en beïnvloeding
2. De Sovjet-Unie. Een systeem onder spanning

Aardrijkskunde (nieuwe stijl)

1. Politiek en ruimte
2. Migratie en mobiliteit
3. Mens en milieu

Aardrijkskunde (oude stijl)

1. Het Nederlandse landschap – zandlandschap
2. Het Nederlandse landschap – rivierkleilandschap
3. Migratie en mobiliteit
4. Nederland en de internationale milieuproblemen
5. De Nederlander en de milieugebruiksruimte
6. Marokko
7. Internationalisering: Nederland in Europa, Europa in de wereld

Maatschappijleer (nieuwe stijl en oude stijl)

1. Politieke besluitvorming
2. Criminaliteit en rechtsstaat
3. Ontwikkelingssamenwerking

Filosofie (nieuwe stijl) en Filosofie (oude stijl)

Het demarcatieprobleem in de wetenschappen

Tekenen, handenarbeid, textiele werkvormen (nieuwe stijl en oude stijl)

Van beroep kunstenaar

CKV-2 (nieuwe stijl)

1. Hofcultuur in de zestiende en zeventiende eeuw
2. Cultuur van Romantiek en Realisme in de negentiende eeuw
3. Cultuur van het moderne en massacultuur in de twintigste eeuw

HAVO

Geschiedenis (nieuwe stijl) en geschiedenis en staatsinrichting (oude stijl)

1. Nederland en Indonesië. Vier eeuwen contact en beïnvloeding
2. De Sovjet-Unie. Een systeem onder spanning

Aardrijkskunde (nieuwe stijl)

1. Migratie en vervoer
2. Politiek en ruimte

Aardrijkskunde (oude stijl)

1. Het Nederlandse landschap – zandlandschap
2. Het Nederlandse landschap – rivierkleinlandschap
3. Migratie en vervoer
4. Frankrijk

Maatschappijleer (nieuwe stijl en oude stijl)

1. Politieke besluitvorming
2. Multiculturele samenleving
3. Criminaliteit en rechtsstaat

Filosofie

'Macht', n.a.v. Hobbes en Foucault.

Tekenen, handenarbeid, textiele werkvormen (nieuwe stijl en oude stijl)

Tekenen

1. Nederlandse 17^e-eeuwse schilderkunst
2. Architectuur, beeldende kunst en vormgeving vanaf 1850
3. Schilderkunst van de Romantiek

Handenarbeid

1. Nederlandse 17^e-eeuwse schilderkunst
2. Architectuur, beeldende kunst en vormgeving vanaf 1850
3. Middeleeuwse beeldhouwkunst
4. Beeldhouwkunst van de Barok

Textiele werkvormen

1. Nederlandse 17^e-eeuwse schilderkunst
2. Architectuur, beeldende kunst en vormgeving (kleding) vanaf 1850
3. Middeleeuwse kunst en kleding
4. 18^e-eeuwse kleding

CKV-2 (nieuwe stijl)

1. Cultuur van Romantiek en Realisme in de negentiende eeuw
2. Cultuur van het moderne in de eerste helft van de twintigste eeuw; massacultuur in de tweede helft van de twintigste eeuw

MAVO/VBO algemene vakken C en D

Geschiedenis en staatsinrichting

1. Nederland en Indonesië. Vier eeuwen contact en beïnvloeding
2. Dwang en overtuiging. Het communistisch systeem onder Stalin, 1928-1953

Aardrijkskunde

1. Spanje
2. De Nederlanders en hun vakantiebestemmingen

Maatschappijleer

1. Politiek en beleid
2. Criminaliteit en rechtsstaat
3. Arbeid en samenleving

2.5 Vooruitblik 2003

VMBO

In 2003 zullen voor het eerst centrale examens voor het vmbo worden afgenomen. De examens voor het C- en D-niveau zullen worden vervangen door examens voor de kaderberoepsgerichte leerweg (KB), de gemengde leerweg (GL) en de theoretische leerweg (TL). Inhoudelijk zal er voor de meeste examens niet veel veranderen, zij het dat de leerstof uit het examenprogramma anders geordend is en vooral dat het deel van het examenprogramma waarover centraal wordt geëxamineerd bij veel vakken minder omvangrijk is. Ook de populatie van de centrale examens zal veranderen. Konden de leerlingen in het verleden nog op verschillende niveaus examen doen, bij het vmbo kiezen de leerlingen voor een leerweg en worden alle examens in het bij die leerweg passende niveau afgelegd. Verder is de keuze van het vakkenpakket iets ingeperkt door de keuze voor een sector waarbij twee sectorvakken verplicht zijn naast Engels en Nederlands. Er is nog geen zicht op de gevolgen van de vmbo operatie op het vaardigheidsniveau van de examenkandidaten en de samenstelling van de populaties bij de verschillende examens.

Overigens zijn bij de algemene vakken de examens bij de theoretische leerweg identiek aan de examens bij de gemengde leerweg. Het verschil in de leerwegen zit in de samenstelling van het vakkenpakket. In plaats van een algemeen vak is bij de gemengde leerweg een beroepsgericht programma opgenomen.

Er zijn enkele opvallende veranderingen bij vakken: bij Engels (TL en GL) wordt naast leesvaardigheid ook schrijfvaardigheid getoetst. Bij Nederlands (TL en GL) krijgt het onderdeel schrijfvaardigheid het karakter van gedocumenteerd schrijven waarbij ook de resultaten van het sectorwerkstuk worden betrokken. Bij de KB-examens worden de leesvaardigheid en de schrijfvaardigheid in een examen getoetst. Bij de vakken Spaans, Turks, Arabisch en natuur- en scheikunde-2 sluiten de kandidaten het KB-examen uitsluitend af met een schoolexamen. Voor GL en TL zijn er centrale examens.

Bij de kunstvakken zijn de veranderingen aanzienlijk: Er zijn uitsluitend centrale examens voor GL en TL. Er zijn nieuwe examens voor de vakken dans, drama en audiovisueel. Bij de beeldende vakken tekenen, handvaardigheid, textiele werkvormen en audiovisuele vormgeving is er een centraal schriftelijk examen rond de kunstbeschouwing. Verder is er voor elke differentiatie bij de beeldende vakken een centraal praktisch examen.

Van de examens met ingrijpende veranderingen zijn voorbeeldexamens gepubliceerd op de internetsite www.vmbo-examengids.nl.

Voor de leerlingen die zijn gezakt voor het examen in 2002 is er de mogelijkheid deel te nemen aan de bezemexamens van 2003. Veel van deze examens zijn nagenoeg gelijk aan de vmbo-examens. Soms zijn er specifieke bezemexamens, met name bij die vakken waarbij er geen pendant van het C-examen is.

examens met ICT

De inzet van computers bij de examinering is voor havo/vwo vastgelegd in een beleidsnotitie van de CEVO voor de periode 2003-2005. Daarin wordt voorgesteld om elk jaar en voor elk profiel bij één van de profielvakken een examen te ontwikkelen waarbij de computer bij de examenafname zal worden ingezet.

In 2002 is de constructie gestart en zijn de scholen aangezocht dit experiment in 2003 zullen deelnemen. Gemiddeld zal per examen bij 30 à 40% van de vragen gebruik moeten worden gemaakt van de computer om de vragen te beantwoorden. De overige vragen zijn identiek aan die in de reguliere examens voor die vakken.

Ook in het vmbo zal de computer in toenemende mate bij de examinering worden ingezet. In 2003 kunnen de kandidaten bij het examen Nederlands schrijfvaardigheid GL en TL gebruik maken van een computer indien zij dit wensen en de school het mogelijk maakt. Verder zullen de kandidaten de examens dans en drama met behulp van de computer maken. Op enkele scholen zullen kandidaten bij de examens biologie, aardrijkskunde en natuur- en scheikunde 1 gebruik maken van de computer. Ook voor de basisberoepsgerichte leerweg (die buiten dit verslag valt) zullen examens met behulp van de computer worden afgenomen met name bij Frans en misschien Duits.

Normvergelijking

De normvergelijkingsprocedures zullen in 2003 gecontinueerd worden voor de examens Frans, Duits en Engels op alle niveaus, voor Nederlands op havo- en D-niveau en voor muziek alleen op D-niveau.

Voor havo/vwo vindt normvergelijking ook plaats voor de examens wiskunde A, wiskunde B, natuurkunde en scheikunde, Nederlands havo, economie 1,2 havo en biologie vwo. Ook zal in 2003 de vergelijking tussen de resultaten van het eerste en tweede tijdvak weer uitgevoerd worden.

3 Examengegevens per vak

3.1 Cito-vakmedewerkers

Bij de samenstelling van de examenopgaven werkten de verschillende constructiegroepen onder leiding van de volgende Cito-medewerkers:

Latijn	vwo	H. Terra
Grieks	vwo	F. Partouns
Nederlands	havo/vwo vbo/mavo	A. v.d. Kerkhof K. Schreuder, L. Kruze, S. Sluiter
Frans	havo/vwo vbo/mavo	C. Peer C. Peer, I. v/d Laar
Duits	havo/vwo vbo/mavo	L. Melse, O. de Vries H. Rauschenberg
Engels	havo/vwo vbo/mavo	D. Samson, N. v. Zuijlen M. Egberink, E. Reichard,
Spaans	havo/vwo vbo/mavo	F. Agerkop M. Welling
Russisch	havo/vwo	O. Petri
Turks	havo/vwo vbo/mavo	T. Duindam Ö. Konak
Arabisch	havo/vwo vbo/mavo	A. de Graaf F. El Azzarouali
Wiskunde A	havo/vwo	K. Lagerwaard, G. Limpens
Wiskunde B	havo/vwo	E. van Kervel, E. Severins G. Stroomer, H. Boertien
Wiskunde	vbo/mavo	P. Boon
Natuurkunde	havo/vwo vbo/mavo	J. Hendricx, B. Kneepkens P. Smeets
Scheikunde	havo/vwo vbo/mavo	K. Beers, D. Witte B. Schut
Biologie	havo/vwo vbo/mavo	J. Brûens, M. Lieverse M. Heideveld
Economische vakken	havo/vwo	K. Blokker, F. Denie, N. Dieteren, A. Geurtsen, L. Knobben, M. Nijpels N. Dieteren
Aardrijkskunde	vbo/mavo havo/vwo	B. v. Erp Taalman Kip, R. van Roon
Geschiedenis	vbo/mavo havo/vwo	H. Post S. Boom
Filosofie	vbo/mavo	H. Kurstjens
Maatschappijleer	havo/vwo vbo/mavo	H. Wessels V. Gijsselhart N. Bogaard
Muziek	havo/vwo vbo/mavo	T. Doevendans H. Denissen
Tehatex	havo/vwo vbo/mavo	M. Wensing, M. Knüppe T. Nijhof, V. v Woerkom
CKV 2	havo/vwo	T. Doevendans, M. Wensing

3.2 Tabellen

In de hiernavolgende tabellen volgen per vak de algemene gegevens van de examens eerste tijdvak, verdeeld over de verschillende schooltypen. Een toelichting op een aantal psychometrische begrippen staat in de bijlage.

De gegevens over de examens zijn bij alle vakken ontleend aan de verwerking van de afnameresultaten van een steekproef van kandidaten. In deze steekproef zijn uitsluitend kandidaten opgenomen uit het reguliere VO.

De genoemde aantallen van de kandidaten zijn gebaseerd op de gegevens van de inschrijving. Deze aantallen kunnen een overschatting van enkele honderden kandidaten te zien geven. De aantallen C- en D-kandidaten kunnen niet vermeld worden omdat deze bij de inschrijving niet onderscheiden worden. Daarom worden in de voorkomende gevallen, de aantallen gegeven van de gehele vbo- en mavo-populatie voor dat bepaalde vak.

In de regel worden per examen gegevens gerapporteerd uit de toets- en itemanalyse. Wanneer deze is uitgevoerd op een aantal leerlingen kleiner dan 25 worden gegevens als gemiddeld cijfer, percentage onvoldoende, betrouwbaarheid en standaardmeefout niet vermeld.

Tabel 20 Algemene gegevens van het CE Nederlands 2002

	leesvaardigheid		schrijfvaardigheid			
	C	D	fs C	fs D	opstel C	opstel D
Totaal aantal kandidaten	84.344		83.360		83.360	
Steekproefgrootte	2.925	3.434	1.774	1.780	365	656
Aantal vragen *	40	42	15	14	10	10
P'-waarde van het examen jongens + meisjes	0,55	0,64	0,63	0,64	-	-
P'-waarde, alleen jongens	0,57	0,64	0,60	0,60	-	-
P'-waarde, alleen meisjes	0,54	0,64	0,66	0,67	-	-
Maximumscore	51	48	90	90	100	100
Gemiddelde score	28	31	56	57	-	-
Normeringsterm	1,3	0,8	0,8	0,8	n.v.t.	n.v.t.
Gemiddeld cijfer	6,3	6,5	6,4	6,5	-	-
Percentage onvoldoendes (<5.5)	18	12	20	18	-	-
Standaardafwijking	5,3	5,1	12,3	12,4	-	-
Betrouwbaarheid	0,56	0,54	-	-	-	-
Standaardmeetfout	3,6	3,5	-	-	-	-

* het vermelde aantal geeft het aantal afzonderlijk gescoorde beoordelingscomponenten weer

Tabel 21 Algemene gegevens van het CE Nederlands 2002¹⁾

	leesvaardigheid	
	HAVO	VWO
Totaal aantal kandidaten	42.101	29085
Steekproefgrootte	1.905	1903
Aantal vragen*	32	30
P'-waarde van het examen jongens + meisjes	0,59	0,59
P'-waarde, alleen jongens	0,58	0,57
P'-waarde, alleen meisjes	0,61	0,60
Maximumscore	50	49
Gemiddelde score	30	29
Normeringsterm	1,1	1,0
Gemiddeld cijfer	6,4	6,3
Percentage onvoldoendes (<5.5)	18	20
Standaardafwijking	5,4	5,3
Betrouwbaarheid	0,39	0,45
Standaardmeetfout	4,2	3,9

¹⁾ Geen gegevens over Nederlands HAVO oude stijl en VWO oude stijl)

* het vermelde aantal geeft het aantal afzonderlijk gescoorde beoordelingscomponenten weer

Tabel 22 Algemene gegevens van het CE Latijn, Grieks 2002

	la	gr
	VWO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	4.912	1.846
Steekproefgrootte	510	770
Aantal vragen	49	46
P'-waarde van het examen jongens + meisjes	0,6	0,59
P'-waarde, alleen jongens	0,56	0,57
P'-waarde, alleen meisjes	0,62	0,60
Maximumscore	90	84
Gemiddelde score	54	50
Normeringsterm	1,1	1,3
Gemiddeld cijfer	6,5	6,6
Percentage onvoldoendes (<5.5)	28	23
Standaardafwijking	16,1	14,4
Betrouwbaarheid	0,90	0,89
Standaardmeetfout	5,1	4,9

Tabel 23 Algemene gegevens van het CE Frans 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	11.508		8.691	7.718
Steekproefgrootte	1.543	2.299	1.910	1.849
Aantal vragen	44	42	45	44
P'-waarde van het examen jongens + meisjes	0,63	0,69	0,51	0,65
P'-waarde, alleen jongens	0,61	0,69	0,53	0,69
P'-waarde, alleen meisjes	0,64	0,69	0,51	0,65
Maximumscore	46	46	48	46
Gemiddelde score	29	32	25	30
Normeringsterm	0,6	0,1	1,2	0,1
Gemiddeld cijfer	6,3	6,4	5,8	6,0
Percentage onvoldoendes (<5.5)	18	23	37	36
Standaardafwijking	5,1	5,8	5,5	6,0
Betrouwbaarheid	0,67	0,76	0,68	0,76
Standaardmeetfout	2,9	2,9	3,1	2,9

Tabel 24 Algemene gegevens van het CE Duits 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	34.699		10.089	7.428
Steekproefgrootte	2.944	2.923	1.866	1.788
Aantal vragen	42	45	50	44
P'-waarde van het examen jongens + meisjes	0,62	0,57	0,55	0,55
P'-waarde, alleen jongens	0,63	0,58	0,57	0,59
P'-waarde, alleen meisjes	0,61	0,56	0,54	0,55
Maximumscore	46	49	50	49
Gemiddelde score	29	28	27	27
Normeringsterm	0,6	1,0	0,8	0,9
Gemiddeld cijfer	6,2	6,1	5,7	5,9
Percentage onvoldoendes (<5.5)	24	28	37	36
Standaardafwijking	5,9	6,0	5,9	6,5
Betrouwbaarheid	0,76	0,75	0,70	0,77
Standaardmeetfout	2,9	3,0	3,2	3,1

Tabel 25 Algemene gegevens van het CE Engels 2002

	C	D	HAVO nieuwe stijl	HAVO compex	VWO nieuwe stijl
Totaal aantal kandidaten	82.643		41.989	0	29.359
Steekproefgrootte	3.814	3.662	2.019	334	2.000
Aantal vragen	44	42	40	32	43
P'-waarde van het examen jongens + meisjes	0,6	0,68	0,61	0,54	0,63
P'-waarde, alleen jongens	0,61	0,7	0,62	0	0,64
P'-waarde, alleen meisjes	0,6	0,66	0,6	0	0,63
Maximumscore	49	46	50	50	47
Gemiddelde score	30	31	30	27	30
Normeringsterm	0,6	0,4	0,8	1,5	0,7
Gemiddeld cijfer	6	6,5	6,3	6,3	6,4
Percentage onvoldoendes (<5.5)	34	18	26	27	22
Standaardafwijking	7,5	6,2	7,1	8,3	6,6
Betrouwbaarheid	0,83	0,78	0,79	0,81	0,78
Standaardmeetfout	3,1	2,9	3,3	3,6	3,1

Tabel 26 Algemene gegevens van het CE Spaans 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	722		1.062	700
Steekproefgrootte	17	30	68	96
Aantal vragen	40	45	44	49
P'-waarde van het examen jongens + meisjes	-	0,59	0,61	0,67
P'-waarde, alleen jongens	-	0,62	0,64	0,68
P'-waarde, alleen meisjes	-	0,58	0,59	0,67
Maximumscore	49	48	49	51
Gemiddelde score	-	28	30	34
Normeringsterm	1,1	1,0	0,8	0,5
Gemiddeld cijfer	-	6,3	6,3	6,5
Percentage onvoldoendes (<5.5)	-	43	34	21
Standaardafwijking	-	11,4	7,6	6,7
Betrouwbaarheid	-	0,94	0,83	0,81
Standaardmeetfout	-	2,9	3,1	3,0

Tabel 27 Algemene gegevens van het CE Russisch 2002

	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	1	14
Steekproefgrootte	0	9
Aantal vragen	44	49
P'-waarde van het examen jongens + meisjes	-	-
P'-waarde, alleen jongens	-	-
P'-waarde, alleen meisjes	-	-
Maximumscore	49	51
Gemiddelde score	-	-
Normeringsterm	0,7	0,6
Gemiddeld cijfer	-	-
Percentage onvoldoendes (<5.5)	-	-
Standaardafwijking	-	-
Betrouwbaarheid	-	-
Standaardmeetfout	-	-

Tabel 28 Algemene gegevens van het CE Turks 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	322	18	6	
Steekproefgrootte	34	49	7	2
Aantal vragen	48	49	44	47
P'-waarde van het examen jongens + meisjes	0,70	0,67	-	-
P'-waarde, alleen jongens	0,70	0,67	-	-
P'-waarde, alleen meisjes	0,69	0,67	-	-
Maximumscore	50	52	50	55
Gemiddelde score	35	35	-	-
Normeringsterm	1,2	0,5	0,8	0,8
Gemiddeld cijfer	7,5	6,6	-	-
Percentage onvoldoendes (<5.5)	3	16	-	-
Standaardafwijking	5,1	6,9	-	-
Betrouwbaarheid	0,68	0,81	-	-
Standaardmeetfout	2,9	3,0	-	-

Tabel 29 Algemene gegevens van het CE Arabisch 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	87		13	5
Steekproefgrootte	6	32	1	0
Aantal vragen	43	43	41	40
P'-waarde van het examen jongens + meisjes	-	0,78	-	-
P'-waarde, alleen jongens	-	0,75	-	-
P'-waarde, alleen meisjes	-	0,8	-	-
Maximumscore	47	46	45	45
Gemiddelde score	-	36	-	-
Normeringsterm	1,1	0,1	0,7	0,7
Gemiddeld cijfer	-	7,2	-	-
Percentage onvoldoendes (<5.5)	-	0	-	-
Standaardafwijking	-	4,8	-	-
Betrouwbaarheid	-	0,72	-	-
Standaardmeetfout	-	2,5	-	-

Tabel 30 Algemene gegevens van het CE Wiskunde 2002

	C	D
Totaal aantal kandidaten	65.036	
Steekproefgrootte	4.099	3.548
Aantal vragen	25	24
P'-waarde van het examen jongens + meisjes	0,52	0,55
P'-waarde, alleen jongens	0,54	0,56
P'-waarde, alleen meisjes	0,50	0,52
Maximumscore	90	90
Gemiddelde score	47	49
Normeringsterm	1,1	1,1
Gemiddeld cijfer	5,8	6,0
Percentage onvoldoendes (<5.5)	37	33
Standaardafwijking	14,0	12,7
Betrouwbaarheid	0,80	0,75
Standaardmeetfout	6,3	6,3

Tabel 31 Algemene gegevens van het CE Wiskunde A 2002

	wa 1, 2 HAVO nieuwe stijl	wa VWO oude stijl	wa 1 VWO nieuwe stijl	wa 1, 2 VWO nieuwe stijl
Totaal aantal kandidaten	16.051	3.241	4.703	9.292
Steekproefgrootte	2.064	331	1.900	2.127
Aantal vragen	21	19	19	20
P'-waarde van het examen jongens + meisjes	0,44	0,55	0,53	0,54
P'-waarde, alleen jongens	0,45	0,56	0,54	0,53
P'-waarde, alleen meisjes	0,43	0,54	0,53	0,55
Maximumscore	90	90	81	90
Gemiddelde score	40	50	43	49
Normeringsterm	1,8	1,3	1,6	1,4
Gemiddeld cijfer	5,8	6,3	6,4	6,3
Percentage onvoldoendes (<5.5)	41	29	23	29
Standaardafwijking	13,2	13,7	11,6	14,4
Betrouwbaarheid	0,73	0,70	0,69	0,76
Standaardmeetfout	6,8	7,5	6,4	7,1

Tabel 32 Algemene gegevens van het CE Wiskunde B 2002

	wb1 HAVO nieuwe stijl	wb1, 2 HAVO nieuwe stijl	wb VWO oude stijl	wb1 VWO nieuwe stijl	wb1, 2 VWO nieuwe stijl
Totaal aantal kandidaten	6.012	5.601	1.533	6.669	6.039
Steekproefgrootte	1.968	1.987	151	1.856	1.900
Aantal vragen	21	19	13	18	18
P'-waarde van het examen jongens + meisjes	0,46	0,52	0,39	0,63	0,65
P'-waarde, alleen jongens	0,45	0,52	0,38	0,61	0,64
P'-waarde, alleen meisjes	0,46	0,52	0,41	0,64	0,65
Maximumscore	83	88	90	84	86
Gemiddelde score	38	45	35	53	56
Normeringsterm	2,0	1,8	2,0	0,8	0,8
Gemiddeld cijfer	6,1	6,4	5,5	6,4	6,6
Percentage onvoldoendes (<5.5)	31	27	46	26	24
Standaardafwijking	12,4	14,9	13,5	14,2	14,7
Betrouwbaarheid	0,73	0,80	0,69	0,78	0,78
Standaardmeetfout	6,5	6,6	7,5	6,7	7,0

Tabel 33 Algemene gegevens van het CE Natuurkunde 2002

	C	D	HAVO N1 nieuwe stijl	HAVO N1,2 nieuwe stijl	VWO N1 nieuwe stijl	VWO N1,2 nieuwe stijl	VWO moderne natuur- kunde
Totaal aantal kandidaten	27.181		6.055	6.735	6.919	7.199	345
Steekproefgrootte	2.655	2.622	1.926	1.941	1.958	1.859	144
Aantal vragen	45	40	27	27	23	24	24
P'-waarde van het examen jongens + meisjes	0,51	0,57	0,66	0,65	0,58	0,7	0,66
P'-waarde, alleen jongens	0,52	0,58	0,66	0,66	0,58	0,7	0,66
P'-waarde, alleen meisjes	0,49	0,53	0,66	0,62	0,58	0,69	0,67
Maximumscore	90	90	78	84	78	83	83
Gemiddelde score	46	52	51	55	45	58	55
Normeringsterm	1,3	0,9	0,8	0,8	1,0	0,8	1,2
Gemiddeld cijfer	5,9	6,1	6,7	6,7	6,2	7,1	7,1
Percentage onvoldoendes (<5.5)	32	27	12	13	25	8	5
Standaardafwijking	10,1	10,1	9,4	10,4	10,0	10,4	9,4
Betrouwbaarheid	0,68	0,68	0,70	0,75	0,72	0,77	0,72
Standaardmeetfout	5,7	5,7	5,1	5,2	5,3	5,0	5,0

Tabel 34 Algemene gegevens van het CE Scheikunde 2002

	C	D	HAVO nieuwe stijl	VWO oude stijl	VWO sk 1 nieuwe stijl	VWO sk 1, 2 nieuwe stijl
Totaal aantal kandidaten	16.504		10.882	1.650	6.225	5.842
Steekproefgrootte	2.075	2.271	1.839	149	1.933	1.976
Aantal vragen	46	45	35	24	24	25
P'-waarde van het examen jongens + meisjes	0,56	0,61	0,63	0,52	0,54	0,64
P'-waarde, alleen jongens	0,57	0,62	0,63	0,55	0,52	0,64
P'-waarde, alleen meisjes	0,55	0,59	0,62	0,49	0,55	0,64
Maximumscore	90	90	81	70	66	71
Gemiddelde score	50	55	51	36	36	45
Normeringsterm	1,2	1,0	0,8	1,0	1,3	1,0
Gemiddeld cijfer	6,2	6,5	6,4	5,7	6,2	6,8
Percentage onvoldoendes (<5.5)	21	17	19	43	26	18
Standaardafwijking	10,0	10,8	10,4	8,1	9,1	10,8
Betrouwbaarheid	0,67	0,76	0,77	0,61	0,73	0,80
Standaardmeetfout	5,8	5,3	5,0	5,0	4,8	4,8

Tabel 35 Algemene gegevens van het CE Biologie 2002

	C	D	HAVO nieuwe stijl	HAVO complex	VWO nieuwe stijl
Totaal aantal kandidaten	37.694		14.501	284	10.865
Steekproefgrootte	2.927	2.975	1.638	140	1.784
Aantal vragen	55	50	49	53	40
P'-waarde van het examen jongens + meisjes	0,57	0,61	0,57	0,52	0,60
P'-waarde, alleen jongens	0,58	0,62	0,59	0,53	0,60
P'-waarde, alleen meisjes	0,57	0,61	0,56	0,51	0,60
Maximumscore	95	86	84	88	82
Gemiddelde score	54	53	48	45	49
Normeringsterm	1,0	0,8	0,9	1,5	0,9
Gemiddeld cijfer	6,2	6,3	6,1	6,1	6,3
Percentage onvoldoendes (<5.5)	20	18	26	19	19
Standaardafwijking	9,6	8,7	8,3	7,4	8,6
Betrouwbaarheid	0,63	0,61	0,62	0,55	0,63
Standaardmeetfout	5,8	5,4	5,1	5,0	5,2

Tabel 36 Algemene gegevens van het CE Economie 2002

	C	D
Totaal aantal kandidaten	38.432	
Steekproefgrootte	2.246	2.312
Aantal vragen	43	43
P'-waarde van het examen jongens + meisjes	0,56	0,59
P'-waarde, alleen jongens	0,56	0,60
P'-waarde, alleen meisjes	0,55	0,59
Maximumscore	85	85
Gemiddelde score	47	50
Normeringsterm	1,2	1,0
Gemiddeld cijfer	6,2	6,3
Percentage onvoldoendes (<5.5)	20	18
Standaardafwijking	8,6	8,7
Betrouwbaarheid	0,60	0,63
Standaardmeetfout	5,5	5,3

Tabel 37 Algemene gegevens van het CE Economische vakken HAVO 2002

	ec1 HAVO nieuwe stijl	ec1,2 HAVO nieuwe stijl	M&O HAVO
Totaal aantal kandidaten	14.667	18.526	11.650
Steekproefgrootte	1.980	2.051	1.831
Aantal vragen	34	26	33
P'-waarde van het examen jongens + meisjes	0,53	0,50	0,61
P'-waarde, alleen jongens	0,59	0,51	0,63
P'-waarde, alleen meisjes	0,51	0,48	0,59
Maximumscore	63	60	65
Gemiddelde score	34	30	40
Normeringsterm	1,2	1,2	1,0
Gemiddeld cijfer	6,0	5,7	6,5
Percentage onvoldoendes (<5.5)	31	43	16
Standaardafwijking	8,1	7,5	7,9
Betrouwbaarheid	0,70	0,63	0,72
Standaardmeetfout	4,4	4,6	4,2

Tabel 38 Algemene gegevens van het CE Economische vakken VWO 2002

	ec II VWO oude stijl	Ec 1 VWO nieuwe stijl	Ec 1, 2 VWO nieuwe stijl	M&O VWO
Totaal aantal kandidaten	1.281	5.913	11.946	7.746
Steekproefgrootte	178	1.754	2.087	1.759
Aantal vragen	40	31	26	33
P'-waarde van het examen jongens + meisjes	0,45	0,53	0,53	0,58
P'-waarde, alleen jongens	0,44	0,57	0,54	0,59
P'-waarde, alleen meisjes	0,46	0,49	0,51	0,57
Maximumscore	90	62	61	74
Gemiddelde score	40	33	32	43
Normeringsterm	1,4	1,5	1,5	1,1
Gemiddeld cijfer	5,4	6,2	6,2	6,3
Percentage onvoldoendes (<5.5)	48	29	24	19
Standaardafwijking	11,1	9,6	7,8	9,0
Betrouwbaarheid	0,78	0,79	0,69	0,74
Standaardmeetfout	5,1	4,4	4,4	4,5

Tabel 39 Algemene gegevens van het CE Aardrijkskunde 2002

	C	D	HAVO nieuwe stijl	VWO oude stijl	VWO nieuwe stijl
Totaal aantal kandidaten	22.902		18.794	1.672	11.505
Steekproefgrootte	1.433	2.102	1.771	159	1.934
Aantal vragen	45	45	26	38	30
P'-waarde van het examen jongens + meisjes	0,64	0,66	0,53	0,53	0,62
P'-waarde, alleen jongens	0,68	0,68	0,55	0,54	0,64
P'-waarde, alleen meisjes	0,63	0,64	0,50	0,52	0,60
Maximumscore	90	90	67	86	75
Gemiddelde score	58	59	35	46	46
Normeringsterm	0,8	0,8	1,2	1,2	0,8
Gemiddeld cijfer	6,6	6,7	6,0	6,0	6,4
Percentage onvoldoendes (<5.5)	10	9	30	30	17
Standaardafwijking	9,1	9,0	7,6	8,9	8,0
Betrouwbaarheid	0,63	0,66	0,64	0,63	0,62
Standaardmeetfout	5,6	5,2	4,6	5,4	4,9

Tabel 40 Algemene gegevens van het CE Geschiedenis en staatsinrichting 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten		17.027	29.187	16.733
Steekproefgrootte	1.166	1.962	1.952	1.895
Aantal vragen	36	35	24	22
P'-waarde van het examen jongens + meisjes	0,63	0,66	0,53	0,61
P'-waarde, alleen jongens	0,65	0,67	0,54	0,61
P'-waarde, alleen meisjes	0,61	0,64	0,52	0,61
Maximumscore	85	85	81	90
Gemiddelde score	53	56	43	55
Normeringsterm	0,8	0,8	1,2	0,8
Gemiddeld cijfer	6,4	6,7	6,0	6,3
Percentage onvoldoendes (<5.5)	15	10	33	23
Standaardafwijking	9,0	9,5	10,5	11,1
Betrouwbaarheid	0,70	0,75	0,76	0,73
Standaardmeetfout	5,0	4,7	5,1	5,7

Tabel 41 Algemene gegevens van het CE Maatschappijleer 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	4.525		2.911	2.040
Steekproefgrootte	430	274	353	342
Aantal vragen	45	44	34	34
P'-waarde van het examen jongens + meisjes	0,56	0,54	0,54	0,65
P'-waarde, alleen jongens	0,58	0,55	0,52	0,66
P'-waarde, alleen meisjes	0,55	0,54	0,54	0,64
Maximumscore	90	90	83	82
Gemiddelde score	50	49	45	53
Normeringsterm	1,1	1,3	1,1	0,8
Gemiddeld cijfer	6,1	6,2	5,9	6,6
Percentage onvoldoendes (<5.5)	21	18	33	15
Standaardafwijking	8,7	8,5	9,9	9,6
Betrouwbaarheid	0,55	0,57	0,73	0,74
Standaardmeetfout	5,9	5,6	5,1	4,9

Tabel 42 Algemene gegevens van het CE Filosofie 2002

	HAVO	VWO oude stijl	VWO nieuwe stijl
Totaal aantal kandidaten	276	262	1.659
Steekproefgrootte	181	7	1.106
Aantal vragen	13	13	13
P'-waarde van het examen jongens + meisjes	0,69	-	0,71
P'-waarde, alleen jongens	0,65	-	0,70
P'-waarde, alleen meisjes	0,71	-	0,71
Maximumscore	59	45	44
Gemiddelde score	41	-	31
Normeringsterm	0,8	1,0	0,8
Gemiddeld cijfer	7,0	-	7,2
Percentage onvoldoendes (<5.5)	9	-	9
Standaardafwijking	7,0	-	6,3
Betrouwbaarheid	0,67	-	0,68
Standaardmeetfout	4,0	-	3,5

Tabel 43 Algemene gegevens van het CE Muziek 2002

	C	D	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	746		1.186	713
Steekproefgrootte	34	212	359	253
Aantal vragen	50	47	50	52
P'-waarde van het examen jongens + meisjes	0,64	0,71	0,61	0,61
P'-waarde, alleen jongens	0,67	0,73	0,63	0,61
P'-waarde, alleen meisjes	0,63	0,70	0,60	0,60
Maximumscore	50	50	90	90
Gemiddelde score	32	35	55	54
Normeringsterm	0,5	0,0	1,1	1,0
Gemiddeld cijfer	6,3	6,4	6,6	6,4
Percentage onvoldoendes (<5.5)	15	16	15	17
Standaardafwijking	4,4	4,8	10,4	10,5
Betrouwbaarheid	0,57	0,62	0,76	0,79
Standaardmeetfout	2,9	3,0	5,1	4,8

Tabel 44 Algemene gegevens van het CE Tekenen, handvaardigheid en textiele werkvormen 2002

	te C	te D	ha C	ha D	tx C	tx D
Totaal aantal kandidaten	7.259		3.293		268	
Steekproefgrootte	650	1.262	415	747	18	99
Aantal vragen	42	45	40	44	41	43
P'-waarde van het examen jongens + meisjes	0,53	0,53	0,53	0,56	-	0,56
P'-waarde, alleen jongens	0,55	0,53	0,54	0,54	-	0,56
P'-waarde, alleen meisjes	0,53	0,52	0,52	0,56	-	0,56
Maximumscore	90	90	90	90	90	90
Gemiddelde score	48	47	47	50	-	50
Normeringsterm	1,3	1,3	1,3	1,0	1,3	1,1
Gemiddeld cijfer	6,1	6,0	6,0	6,0	-	6,1
Percentage onvoldoendes (<5.5)	24	25	27	24	-	21
Standaardafwijking	10,7	8,9	10,3	8,7	-	8,4
Betrouwbaarheid	0,83	0,75	0,81	0,74	-	0,71
Standaardmeetfout	4,4	4,4	4,5	4,5	-	4,5

Tabel 45 Algemene gegevens van het CE Tekenen, handvaardigheid en textiele werkvormen 2002

	te HAVO nieuwe stijl	ha HAVO nieuwe stijl	tx HAVO nieuwe stijl	tehatex VWO nieuwe stijl
Totaal aantal kandidaten	5.023	1.717	182	3.178
Steekproefgrootte	1.684	748	28	1.646
Aantal vragen	43	44	42	35
P'-waarde van het examen jongens + meisjes	0,56	0,53	0,61	0,63
P'-waarde, alleen jongens	0,55	0,54	-	0,61
P'-waarde, alleen meisjes	0,56	0,53	0,61	0,63
Maximumscore	90	90	90	68
Gemiddelde score	50	48	57	43
Normeringsterm	1,0	1,3	0,8	0,8
Gemiddeld cijfer	6,0	6,1	6,5	6,4
Percentage onvoldoendes (<5.5)	25	22	7	17
Standaardafwijking	8,6	9,6	8,0	7,6
Betrouwbaarheid	0,73	0,78	0,72	0,76
Standaardmeetfout	4,4	4,5	4,3	3,8

Tabel 46 Algemene gegevens van het CE Cultureel Kunstzinnige Vorming 2 (CKV 2) 2002

	HAVO nieuwe stijl	VWO nieuwe stijl
Totaal aantal kandidaten	3398	1700
Steekproefgrootte	683	581
Aantal vragen	36	34
P'-waarde van het examen jongens + meisjes	0,62	0,64
P'-waarde, alleen jongens	0,59	0,64
P'-waarde, alleen meisjes	0,63	0,64
Maximumscore	66	62
Gemiddelde score	41	40
Normeringsterm	0,9	0,6
Gemiddeld cijfer	6,5	6,4
Percentage onvoldoendes (<5.5)	15	18
Standaardafwijking	7,0	6,8
Betrouwbaarheid	0,69	0,73
Standaardmeetfout	3,9	3,6

Bijlage

Bijlage 1 Enkele begrippen

betrouwbaarheid

De betrouwbaarheid is 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.

Bij een betrouwbare meting zal het resultaat dus niet beïnvloed mogen zijn door storende factoren met betrekking tot de toets, kandidaat en beoordelaar zoals het tijdstip van de toetsafname, de specifieke vormgeving van de toets, of pech of geluk met de opgaven. De mate waarin een meting ongevoelig is voor bepaalde, verstorende factoren, kan worden geschat door berekening van een betrouwbaarheidscoëfficiënt. De waarde van een betrouwbaarheidscoëfficiënt ligt tussen 0 en 1. De coëfficiënt kan op verschillende manieren worden vastgesteld; elke manier richt zich op bepaalde storende factoren:

- a via de test-hertest-methode;
- b via de parallelle-vormen-methode;
- c via een generaliseerbaarheidscoëfficiënt;
- d via een berekening van de interne consistentie (KR20-formule voor dichotome scoring of berekening van alpha (α) wanneer de scoring polytoom is).

De berekening van de interne consistentie wordt in de praktijk het meest gebruikt omdat deze uitgevoerd kan worden op één toetsafname. De interne consistentie geeft vooral aan in hoeverre van item tot item hetzelfde wordt gemeten. Dit gegeven wordt gebruikt om te schatten in hoeverre iemand dezelfde score zou behalen bij het maken van een denkbeeldige, andere vergelijkbare toets over dezelfde leerstof.

cesuur

De cesuur is de grens tussen de hoogste toetsscore waaraan een onvoldoende en de laagste toetsscore waaraan een voldoende wordt toegekend. Wanneer bijvoorbeeld aan een score van 30 punten een voldoende wordt toegekend en aan 29 punten een onvoldoende dan ligt de cesuur tussen 29 en 30. Sinds 2000 wordt de cesuur niet meer expliciet vastgesteld, maar wordt het examen genormeerd door de vaststelling van een normeringsterm. Na vaststelling van de normeringsterm kan men vervolgens de cesuur berekenen door na te gaan welke score het cijfer 5,5 oplevert.

normeringsterm (N-term)

De normeringsterm is een ijkvariabele die na de afname van het examen zodanig wordt vastgesteld dat het examen vergelijkbare eisen stelt als in eerdere jaren. De normeringsterm kan variëren tussen 0,0 en 2,0.

p'-waarde open vraag

De p'-waarde is een getal tussen 0 en 1 dat de moeilijkheidsgraad van een opgave weergeeft. De p'-waarde wordt berekend door de gemiddelde score op een opgave te delen door de maximaal haalbare score op die opgave.

Een opgave met een p'-waarde van .10 is erg moeilijk; een opgave met een p'-waarde van .90 is erg gemakkelijk.

p-waarde gesloten vraag

De p-waarde van een meerkeuzevraag is de proportie kandidaten die het goede antwoord heeft gekozen. Met dit getal wordt de moeilijkheidsgraad van een item weergegeven.

p- of p'-waarde van het examen

Zie: gemiddelde p-waarde (\bar{P})

gemiddelde p-waarde (\bar{P})

De gemiddelde p-waarde (\bar{P}) is het gemiddelde van een reeks p-waarden, verkregen door alle p-waarden op te tellen en de som te delen door het aantal p-waarden of, wat nauwkeuriger is, de gemiddelde score (\bar{X}) te delen door het aantal items van een toets.

Bij gewogen scores, die bij open vragen kunnen voorkomen, wordt de p'-waarde van de gehele toets berekend door de gemiddelde toetsscore te delen door het maximaal haalbare aantal punten.

standaardafwijking

De standaardafwijking (S_x) is een maat voor de spreiding van getallen rondom hun gemiddelde. In het geval van een toets gaat het om de spreiding van toetsscores rondom de gemiddelde score.

De formule voor het berekenen van de standaardafwijking van het gemiddelde is:

$$S_x = \sqrt{\frac{\sum(X_i - \bar{X})^2}{N}}$$

S_x = standaardafwijking

X_i = toetsscore van kandidaat i

\bar{X} = gemiddelde toetsscore

N = aantal kandidaten

standaardmeetfout

De standaardmeetfout (S_e) is een indicatie voor de onnauwkeurigheid van een meting. De grootte van de standaardmeetfout van een toets hangt af van de betrouwbaarheid en de standaardafwijking van de toetsscores.

De formule voor het berekenen van de standaardmeetfout is:

$$S_e = S_x \sqrt{(1-r)}$$

S_e = standaardmeetfout

S = standaardafwijking toetsscores

r = toetsbetrouwbaarheid