

Talenexamens 1999

Te moeilijk of te makkelijk?

In deze bijdrage wordt een overzicht van de resultaten van de centrale examens voor de talen gepresenteerd. De resultaten kunnen beschouwd worden als de resultante van de moeilijkheidsgraad van de examens en het vaardigheidsniveau van de examenkandidaten. Voor zover de gegevens dat toelaten zal op beide aspecten worden ingegaan.

P-waarden, cesuren, % onvoldoenden, gemiddeld cijfer

In tabel 1 op de volgende pagina worden de primaire kenmerken van de eerste tijdvakexamens gepresenteerd. Deze gegevens zijn bij Frans, Duits en Engels gebaseerd op een steekproef van 4000 kandidaten uit het reguliere onderwijs. Bij Arabisch, Russisch, Spaans en Turks zijn de gegevens gebaseerd op alle kandidaten uit het reguliere onderwijs. In de presentatie van deze afnamegegevens zijn dus geen kandidaten uit het VAVO opgenomen. Voor het vak Nederlands worden per school de resultaten van een beperkt aantal kandidaten opgevraagd.

Toelichting op de gegevens in tabel 1:

- de p-waarde, deze geeft het gemiddeld percentage correcte antwoorden weeraan (p). De p-waarde geeft aan of het examen voor de populatie die het examen heeft afgelegd al dan niet moeilijk is geweest. De p-waarde is zowel afhankelijk van de moeilijkheid van het examen als van de vaardigheid van de kandidaten. De p-waarde is daarom slechts een indicatie die aangeeft hoe goed de moeilijkheid van de toets is afgestemd op de vaardigheid van de populatie.
- de cesuur is het midden tussen twee opeenvolgende scorepunten dat exact de grens tussen voldoende en onvoldoende aangeeft (ces);
- het percentage kandidaten dat een onvoldoende cijfer heeft gehaald op het centraal examen bij de vastgestelde cesuur (% onv) en
- het gemiddeld cijfer (gem. cijf.). Door versoepeling van de cesuur kan een examen met een relatief lage p-waarde toch een redelijk gemiddeld cijfer te zien geven.

Moeilijkheidsgraad en cesurbepaling

Omdat de precieze moeilijkheidsgraad van het examen vooraf niet bekend is, wordt de cesuur pas na afname definitief vastgesteld. Voor een (te) moeilijk examen kan dan gecompenseerd worden door een lagere cesuur vast te stellen. Voor het bepalen van de

grens tussen voldoende en onvoldoende maakt de Cevo gebruik van de volgende gegevens:

- afnamegegevens van (een steekproef van) kandidaten;
- gegevens uit de normvergelijking voor de vwo-, havo- en D-examens;
- gegevens over de afstand tussen vbo/mavo C- en D-niveau op basis van de overlapvragen;
- reacties van docenten, docentenorganisaties en leerlingen (LAKS).

De aard van de verzamelde gegevens bij de overlapvragen bij C en D en bij de normvergelijking zullen hieronder kort worden toegelicht.

Afstand vbo/mavo C- en D

De C- en D-examens worden zo geconstrueerd dat er een redelijk verschil in moeilijkheidsgraad tussen beide examens is te verwachten. Echter ook hier geldt dat de moeilijkheidsgraad vooraf niet bekend is. Achteraf wordt gecontroleerd of het gewenste niveauverschil in de examens wel gerealiseerd is. Daartoe is in beide examens een overlapdeel opgenomen. Aan de hand van de resultaten die C- en D-kandidaten op het overlapgedeelte behalen kan zowel het verschil in vaardigheid tussen de C- en D-populatie als het verschil in moeilijkheid tussen het C- en D-examen worden vastgesteld. Als uitgangspunt geldt dat de afstand tussen de C- en D-cesuur niet meer dan 1,5 en niet minder dan 0,5 cijferpunt mag bedragen.

Normvergelijking

De afnamegegevens van examens van verschillende jaren kunnen niet zonder meer met elkaar worden vergeleken. Uit de jaarlijkse afnamegegevens wordt bijvoorbeeld niet duidelijk of een examen moeilijker is dan dat van het voorafgaande jaar of dat de populatie van dit jaar over een geringere vaardigheid beschikt. Om examens over jaren heen met elkaar te kunnen vergelijken wordt normvergelijking toegepast.

Elk jaar wordt het nieuwe examen vergeleken met een door de Cevo aangewezen referentie-examen. Het referentie-examen is een in alle opzichten geslaagd eerste tijdvakexamen. De cesuur op dit examen wordt beschouwd als een belangrijk referentiepunt bij het vaststellen van de cesuur op het nieuwe examen.

Bij de normvergelijking wordt zowel het nieuwe examen als het referentie-examen door eenzelfde groep leerlingen (direct na afname van het examen) gemaakt.

Door deze vergelijking kan vrij nauwkeurig een eventueel verschil in moeilijkheidsgraad tussen beide exa-

Tabel 1. Overzicht p-waarden, cesuren, % onvoldoenden, gemiddeld cijfer

VWO	1995				1996				1997		
	p	ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	p	ces	% onv
Nederl.samenv.	*	54,5	33	5,8	*	54,5	19	6,4	*	54,5	26
Nederl. stelopdr.	*	54,5	18	6,2	*	54,5	12	6,5	*	54,5	9
Frans	68	28,5	22	6,6	66	27,5	23	6,5	68	30,5	29
Duits	68	29,5	23	6,4	65	28,5	25	6,4	70	29,5	20
Engels	65	27,5	26	6,3	68	29,5	26	6,4	65	28,5	30
Spaans	77	30,5	18	7,3	76	30,5	13	7,2	75	31,5	16
Russisch	68	27,5	26	6,7	75	29,5	17	7,3	73	29,5	22

* gemiddeld cijfer

HAVO	1995				1996				1997		
	p	ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	p	ces	% onv
Nederl.samenv.	52	51,5	30	6,0	58	54,5	23	6,2	53	54,5	40
Nederl. stelopdr.	*	54,5	10	6,6	*	54,5	18	6,2	*	54,5	18
Frans	68	28,5	23	6,6	67	28,5	24	6,6	62	25,5	25
Duits	64	28,5	28	6,2	64	29,5	33	6,0	66	30,5	33
Engels	69	29,5	25	6,5	62	26,5	28	6,3	70	30,5	27
Spaans	73	26,5	13	7,4	76	26,5	10	7,7	72	27,5	20
Russisch	66	27,5	56	5,5	55	26,5	50	5,5	74	27,5	35

* gemiddeld cijfer

C-NIVEAU	1995				1996				1997		
	p	ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	p	ces	% onv
Nederl. leesv.	66	54/5	8	6,9	51	49/5	24	6,1	7	54/5	25
Nederl. funct.schr.	*	54/5	11	7,0	*	54/5	21	6,5	*	54/5	28
Nederl.stelopdr.	*	54/5	14	6,4	*	54/5	24	6,4	*	54/5	12
Frans	65	26/5	23	6,5	62	27/5	33	6,1	62	27/5	32
Duits	68	30/5	28	6,3	60	26/5	31	6,1	58	26/5	34
Engels	70	28/5	19	6,8	66	26/5	14	6,7	68	27/5	16
Spaans (jeugd)	73	30/5	28	6,9	72	30/5	30	6,7	64	28/5	44
Arabisch	65	24/5	24	6,8	70	25/5	21	7,2	74	26/5	18
Turks	77	30/5	13	7,2	73	28/5	18	7,25	69	26/5	23

D-NIVEAU	1995				1996				1997		
	p	ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	p	ces	% onv
Nederl. leesv.	63	54/5	16	6,7	55	52/5	25	6,2	55	54/5	29
Nederl. funct.schr.	*	54/5	19	6,5	*	54/5	26	6,2	*	54/5	15
Nederl.stelopdr.	*	54/5	25	5,9	*	54/5	13	6,0	*	54/5	13
Frans	56	24/5	26	6,0	53	24/5	39	5,7	55	24/5	32
Duits	67	31/5	35	6,1	60	27/5	37	5,9	55	26/5	43
Engels	65	30/5	37	5,9	66	29/5	31	6,2	65	28/5	26
Spaans (jeugd)	53	23/5	42	5,5	53	25/5	55	5,7	60	25/5	33
Arabisch											
Turks											

1998							
1999							
ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	
54,5	14	6,7	*	52,5	33	5,9	
54,5	10	6,6	*	54,5	11	6,6	
29,5	34	6,2	62	28,5	37	5,9	
28,5	26	6,1	66	30,5	34	6,0	
29,5	31	6,2	72	29,5	18	6,9	
29,5	28	6,0	65	28,5	27	6,1	
29,5	23	6,9	67	27,5	20	6,5	

1998							
1999							
ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	
54,5	34	5,9	*	54,5	29	6,1	
54,5	19	6,2	*	54,5	14	6,4	
27,5	29	6,1	61	26,5	29	6,3	
30,5	39	5,7	65	29,5	32	6,0	
29,5	25	6,4	68	28,5	22	6,6	
24,5	49	5,9	56	24,5	38	6,0	
27,5	0	6,5	82	27,5	25	8,2	

1998							
1999							
ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	
54,5	27	6,0	61	54,5	15	6,4	
54,5	18	6,6	*	54,5	24	6,4	
54,5	8	6,8	*	54,5	5	6,9	
25,5	36	5,9	60	27,5	38	5,9	
27,5	36	5,9	59	27,5	36	5,9	
28,5	21	6,8	69	27,5	18	6,7	
27,5	34	6,7	69	27,5	35	6,7	
31,5	14	7,8	82	27,5	9	8,2	
27,5	15	7,3	64	24,5	19	6,8	

1998							
1999							
ces	% onv	gem cijf.	p	ces	% onv	gem cijf.	
46,5	30	6,0	59	54,5	22	6,3	
54,5	18	6,6	*	54,5	8	7,1	
54,5	11	6,5	*	54,5	13	6,5	
24,5	38	5,8	51	23,5	41	5,7	
27,5	37	5,8	55	25,5	39	5,7	
28,5	37	6,0	62	28,5	37	5,9	
25,5	35	6,3	60	25,5	31	6,3	
23,5	24	6,6	67	28,5	28	6,4	
27,5	15	6,8	74	31,5	17	6,8	

mens worden vastgesteld. Tevens is het mogelijk om op het nieuwe examen een cesuur aan te geven die overeenkomt met hetzelfde prestatieniveau als de cesuur op het referentie-examen.

De normvergelijkingsprocedure levert naast informatie over verschillen in moeilijkheidsgraad ook informatie op over de vaardigheid van de kandidaten. Daar zal in het tweede deel van dit verslag op worden ingegaan.

Resultaten

Voor de volgende examens zijn cesuren vastgesteld die gelijkwaardig zijn aan de door de Cevo beoogde referentienorm:

- Nederlands D;
- Duits en Engels vwo;
- Duits en Engels havo;
- Engels D.

De examenkandidaten van 1999 worden hier dus met eenzelfde maatstaf beoordeeld als bijvoorbeeld de kandidaten uit 1996, 1997 en 1998. Een hoger of lager gemiddeld cijfer weerspiegelt bij deze vakken dan ook een hoger of lager prestatieniveau.

De cesuren voor de examens Frans en Duits vbo/mavo D zijn iets onder de beoogde referentienorm vastgesteld. Het gemiddelde cijfer voor deze vakken kan daardoor niet zonder meer met dat van andere jaren vergeleken worden. Ondanks deze geringe versoepeling van de gehanteerde norm halen voor Frans-D en Duits-D nog steeds 38 respectievelijk 36% van de kandidaten een onvoldoende.

In tabel 2 worden de C- en D-examens en de C- en D-populaties met elkaar vergeleken. Het verschil in vaardigheid tussen de populaties wordt uitgedrukt in het verschil in p-waarde op het overlapdeel. Het verschil in de moeilijkheid tussen beide examens wordt op basis van het overlapdeel geschat en uitgedrukt in het verschil in p-waarde dat de D-populatie op beide examens zou hebben gehaald. Tenslotte wordt de afstand tussen de cesuur op beide examens uitgedrukt in cijferpunten.

Vak:	verschil tussen	verschil tussen	afstand tussen
	C- en D-populatie	C- en D-examen	C- en D-cesuur
Nederlands leesvaardigh.	12	10	0,9
Frans	19	5	1,2
Duits	13	8	1,1
Engels	23	14	1,3
Spaans	21	7	1,3
Arabisch	26	8	1,5
Turks	11	20	1,3

Tabel 2. Vergelijking van de C- en D- populaties en examens

Met uitzondering van Turks is bij alle vakken het verschil in moeilijkheid tussen beide examens (soms zelfs aanzienlijk) kleiner is dan het verschil in vaardigheid tussen beide populaties. Bij vaststelling van de cesuur van het D-examen maakt de Cevo gebruik

	Moelijkheid examen:	Vaardigheid populatie:
	Het examen '99 is in vgl. met het referentie-examen...	De populatie '99 is in vgl. met de referentie-populatie ...
Engels vwo	2 pnt makkelijker	1,5 pnt vaardiger
Frans vwo	even moeilijk	3 pnt minder vaardig
Duits vwo	1 pnt makkelijker	2 pnt minder vaardig
Engels havo	even moeilijk	1 pnt vaardiger
Frans havo	1 pnt makkelijker	4 pnt minder vaardig
Duits havo	even moeilijk	2,5 pnt minder vaardig
Nederlands D	3 pnt moeilijker	2 pnt vaardiger
Engels D	1 pnt moeilijker	even vaardig
Frans D	2 pnt makkelijker	4 punt minder vaardig
Duits D	2 pnt moeilijker	2,5 pnt minder vaardig

Tabel 3. Overzicht van verschillen met referentie examens en referentiepopulatie

van de gegevens van de normvergelijking. Bij vaststelling van de cesuur voor het C-examen wordt rekening gehouden met de gewenste afstand tussen C- en D-examens.

Bij Frans is bijvoorbeeld het verschil in moeilijkheid tussen het C- en D-examen gering. Het C-examen is naar verhouding voor de C-kandidaten te moeilijk ($P=.51$). Door voor het C-niveau een lagere cesuur vast te stellen is de afstand tussen beide cesuren uiteindelijk toch identiek aan die van vorig jaar.

Ook bij Arabisch en Spaans valt op dat het verschil in moeilijkheidsgraad tussen het C- en D-examen in verhouding tot het verschil tussen de C- en D-populatie gering is. Voor Arabisch is de veronderstelling dat er onder de D-kandidaten veel kandidaten zitten die dit vak gemakkelijk op havo-niveau zouden kunnen afsluiten (indien er een havo-niveau zou worden aangeboden).

Vaardigheid kandidaten en moeilijkheidsgraad van het examen

Over het algemeen lag de beoogde moeilijkheidsgraad van de examens redelijk in de buurt van de beoogde moeilijkheid. Dit wil zeggen in de buurt van de moeilijkheid van de aangewezen referentie-examens. Alleen voor Nederlands leesvaardigheid werd expliciet een iets moeilijker examen beoogd dan het referentie-examen, hetgeen blijkens de resultaten ook gelukt is (zie tabel 3).

Het prestatieniveau van de kandidaten van dit jaar blijkt bij sommige vakken gedaald te zijn ten opzichte van dat van de referentiepopulatie. Deze daling is van dien aard dat – ondanks een geringe versoepeling van de norm bij enkele vakken – dit toch nog tot aanzienlijke percentages onvoldoendes leidde.

Voor de vakken Spaans, Russisch, Arabisch en Turks zijn geen normvergelijkingsgegevens beschikbaar. De cesuren werden vastgesteld op basis van de afnamegegevens, de overlapgegevens tussen de C- en D-examens en op basis van informatie van docenten, vakverenigingen en leerlingen. Opvallend is de hoge p-waarde voor Russisch havo. Het betreft hier echter een zeer gering aantal kandidaten waarbij bovendien door een enkele kandidaat van Russische afkomst het resultaat in positieve zin is beïnvloed.

Door de gegevens die de normvergelijkingsprocedure oplevert kunnen ook uitspraken worden gedaan over

het verschil in vaardigheid tussen de kandidaten van dit jaar en uit het jaar van het referentie-examen. Tabel 3 geeft o.a. dit verschil in vaardigheid aan.

Uit de gegevens over de vaardigheid van de kandidaten blijkt dat de vaardigheid van de kandidaten voor Frans en Duits even als vorig jaar minder is dan de vaardigheid van de referentiepopulatie. Deze constatering bevestigt opnieuw de eerder gesignaleerde daling van het prestatieniveau voor deze vakken bij vbo/mavo en havo. Voor Engels blijft het prestatieniveau relatief constant. Sinds 1997 vindt er ook voor het vwo normvergelijking plaats. Hier is nog niet zo duidelijk sprake van een trend alhoewel de gegevens passen in het bij havo en vbo/mavo gesignaleerde patroon in die zin dat alleen bij Frans en Duits sprake lijkt van een daling van het prestatieniveau. Vorig jaar werd door Van Schooten (1998) gewezen op eenzelfde ontwikkeling bij luistervaardigheid. De scores die de leerlingen halen op de luistertoetsen van het Cito laten een duidelijke daling zien. 'Vooral bij Frans en Duits scoren de leerlingen lager. Komt dit doordat de vaardigheden van de leerlingen afnemen? Worden de toetsen steeds moeilijker? Of zijn de normen aan herziening toe?' De centrale examens worden verondersteld leesvaardigheid te meten en dat ieder jaar op een vergelijkbare wijze te doen. Op basis van die veronderstelling en aan de hand van de verzamelde gegevens zijn deze vragen voor leesvaardigheid te beantwoorden. De resultaten suggereren steeds duidelijker voor Frans en Duits een achteruitgang in prestatieniveau. De examens worden niet structureel moeilijker of makkelijker. Voor incidentele afwijkingen van de beoogde moeilijkheidsgraad wordt bij de vaststelling van de cesuur gecompenseerd. De vraag 'of de normen aan herziening toe zijn' is een politieke keuze. Die vraag kan uitdrukkelijk niet op basis van deze gegevens worden beantwoord.

Literatuur

- Alberts, R., *Examens vbo, mavo, havo, vwo 1999*, Arnhem: Cito, 1999
 Robben, H., Waar ligt de grens, in *Levende Talen* 524, 578-581, 1998
 Schooten, van A., De Kwesitie, in *Levende Talen* 531, 372, 1998

René Alberts en Erna van Hest

zijn werkzaam bij het Cito in Arnhem.