

De invloed van toetsprocedures en examenregels op het studierendement

J. Cohen-Schotanus

Inleiding

Vanaf 1968 zijn, vrijwel zonder onderbreking, van de Groningse medische studenten de studievoortgangsgegevens bijgehouden op een manier die statistische analyse mogelijk maakt. Dit gegevensbestand is de basis voor de volgende beschouwing over de invloed van toetsprocedures en examenregels op het studierendement.

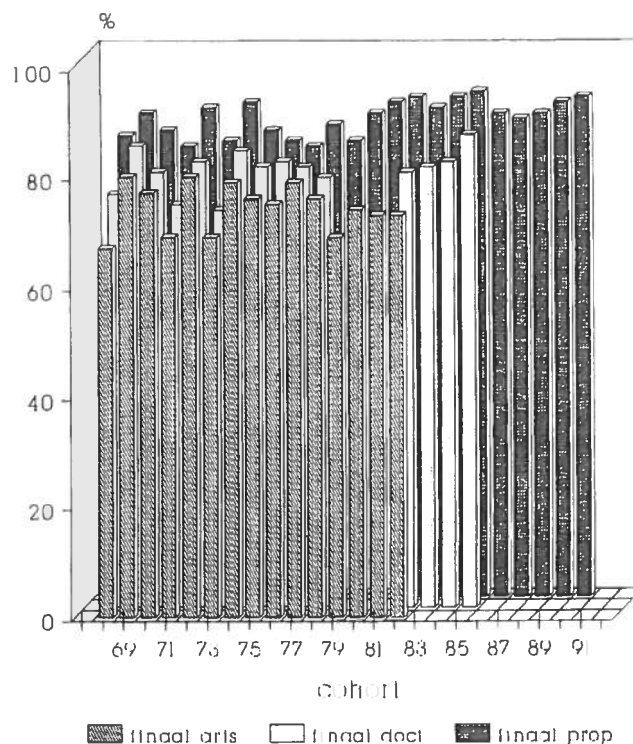
De rendementen

Sinds 1968 zijn ruim 6000 studenten met de studie geneeskunde in Groningen begonnen. Vanaf dat jaar was het studieprogramma zes jaar, daarvoor zeven jaar. De finale rendemen-

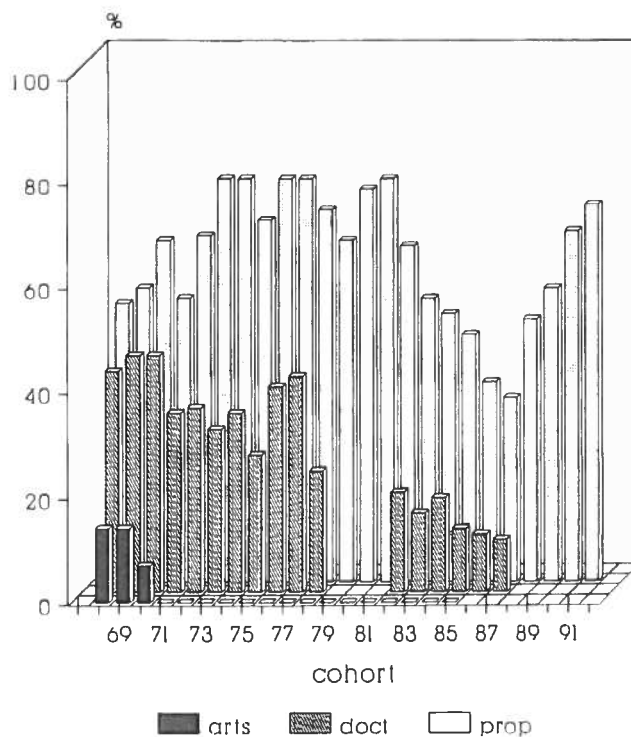
ten (het percentage dat uiteindelijk het diploma haalt, ongeacht de tijd die daarvoor nodig is) van deze studenten zijn weergegeven in figuur 1.

Het gemiddelde rendement voor de propaedeutische is 90%, voor het doctoraal ongeveer 80% en voor het artsexamen ongeveer 75%. Er bestaan geen significante verschillen tussen de verschillende cohorten. Door administratieve oorzaken ontbreken de doctoraalgegevens van de cohorten 1979, 1980 en 1981.

Er is zowel binnen als buiten de faculteit nogal wat veranderd gedurende al die jaren: examenregels veranderden, vakken werden toegevoegd, volgordes van vakken werden al dan niet opnieuw vastgesteld, er kwam een numerus fixus, beperkte inschrijfduur, studie-financiering en sinds 1983 zijn de vrouwelijke



Figuur 1. Finale redementen geneeskunde vanaf 1968 (Groningse studenten)



Figuur 2. Nominaal rendement vanaf 1968 propaedeuse-, doctoraal- en artsexamen

studenten in de meerderheid. Niets van dit alles lijkt van invloed te zijn op de finale rendementen. Voor de nominale rendementen (het percentage studenten dat binnen de officiële programmaduur slaagt) ziet het plaatje er anders uit, zoals is weergegeven in figuur 2.

Bij het artsexamen zien we dat alleen van de eerste drie cohorten een aantal studenten nominaal afstudeert. Dat kon omdat de instroom in het vijfde studiejaar voor die cohorten in parallelgroepen was. Daarna werd overgestapt op een sequentiële instroom met een vaste volgorde van programmaonderdelen en co-assistentenschappen. Dit leidde tot wachttijden, waardoor het behalen van het artsexamen binnen de nominale studieduur niet meer mogelijk was. De nominale doctoraalrendementen laten wel een significant verschil zien tussen de cohorten ($\chi^2 = 85.32$, $df=16$, $p<.001$). Op een paar uitzonderingen na is er sprake van een dalende tendens. In het begin was het nominale doctoraalrendement ruim 40%. Op dit moment is het ongeveer 10%. Ook bij het nominale propaedeuserendement be-

staat een significant verschil tussen de cohorten ($\chi^2 = 58.50$, $df=22$, $p<.001$).

Mogelijke oorzaken van de verschillen in nominale rendementen

De vraag is wat de oorzaak is van deze verschillen in nominale rendementen tussen de cohorten. Mijns inziens zijn er vier mogelijkheden:

Het ligt aan de *studenten*. Bijvoorbeeld: de samenstelling van de groep instromende studenten is sinds 1968 aantoonbaar veranderd. Zo werd in 1972 de numerus fixus ingevoerd. Als gevolg van de gehanteerde toelatingsmethode was het gemiddelde eindexamencijfer van de groep eerstejaars studenten na de invoering van de numerus fixus hoger dan daarvoor. Met andere woorden, de cohorten werden beter in termen van eindexamencijfers. Gezien de positieve correlatie die meestal gevonden wordt tussen eindexamencijfers en de latere studieresultaten, zou vanaf de invoering van de numerus fixus het rendement omhoog moeten

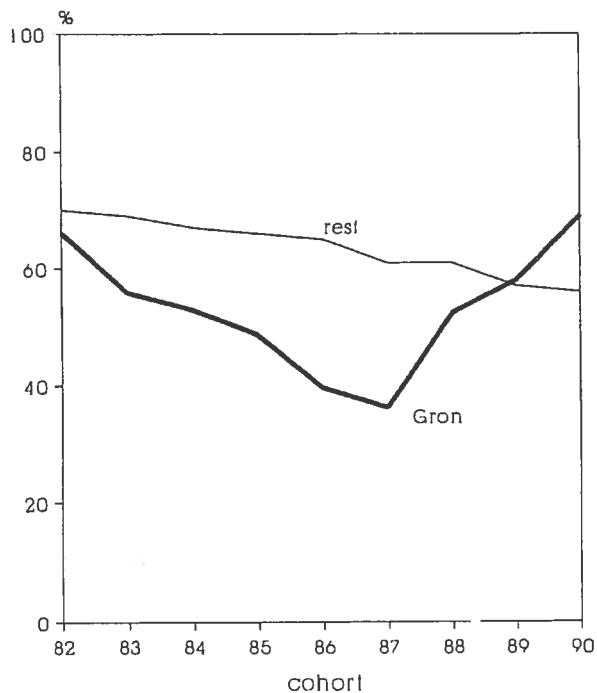
gaan.^{1 2} Bij de propaedeuseresultaten is die tendens inderdaad te zien, maar bij dezelfde cohorten werd het nominale doctoraalrendement steeds lager.

Vanaf 1983 zijn vrouwelijke studenten in de meerderheid. Dit zou van invloed kunnen zijn op het rendement. Omdat vrouwelijke studenten betere studieprestaties leveren dan hun mannelijke collega's - ze halen hogere cijfers en studeren sneller - zou van deze veranderde samenstelling een positief effect verwacht mogen worden.³ Toch daalde vanaf 1983 zowel het propaedeuse- als het doctoraalrendement. Dus, ook hier gebeurde het tegengestelde van wat verwacht kon worden. Ik concludeer dan ook dat de schommelingen in de nominale rendementen niet te maken hebben met de samenstelling van de groep studenten.

Het ligt aan de *docenten*. Zij zijn immers degenen die de studenten voor tentamens laten zakken. In hoeverre het zakken voor tentamens alleen aan studenten zou liggen is de vraag. Er zijn immers nog altijd docenten die stof toetsen uit een klapper die een week voor het tentamen pas te verkrijgen is. Of een boek opgeven dat uitverkocht is en waarvan de herdruk nog niet te verkrijgen is. Of te gedetailleerde vragen stellen, of slechte vragen stellen. Of de cesuur te hoog stellen. En het komt nog steeds voor dat, als een essaytentamen door een andere docent wordt nagekeken, ineens 35% in plaats van de gebruikelijke 15% van de studenten zakt, terwijl het collegebezoek en de gemiddelde tentamenvorbereidingstijd van de studenten niet veranderd is. Natuurlijk gedragen niet alle docenten zich als bovenbeschreven (het gaat in Groningen om zo'n 120 docenten), en waarschijnlijk wordt dan ook slechts een klein deel van de schommelende rendementen hierdoor verklaard. Toch is het een niet onbelangrijk aspect, omdat dit docentgedrag onderdeel is van de onderwijs- en examencultuur. Hiermee kom ik bij de derde factor die van invloed kan zijn op het nominale rendement.

Het ligt aan de *organisatie van het onderwijs en de examenregelgeving*. Bijvoorbeeld,

het nominale artsexamenrendement daalde tot nul op het moment dat de parallelle instroom in het vijfde studiejaar een sequentiële instroom werd. Een veranderde organisatie van het onderwijs, die tot gevolg had dat de doorstroom vertraagde. In datzelfde kalenderjaar werd voor het vierde studiejaar het jaarsysteem afgeschaft. In het jaarsysteem kregen studenten tweemaal per jaar de mogelijkheid tentamen te doen: eenmaal regulier tentamen en een herkansing in augustus. Dan werd de balans opgemaakt. Als het puntenaantal niet voldoende was en/of het aantal onvoldoendes te hoog, moest de student het hele jaar overdoen, ook de tentamens die wel gehaald waren. Na de afschaffing van het jaarsysteem kon elk tentamen viermaal per jaar worden afgelegd. Om te slagen voor het vierde studiejaar moesten nu wel alle tentamens voldoende zijn. Het gevolg was dat studenten tentamens gingen spreiden (voor zich uit schuiven) met als gevolg dat de studie vertraagd werd. Ook studenten met hoge gemiddelde eindexamencijfers doen dat. Hiermee was het te verwachten hogere rendement, vanwege de hogere eindexamencijfers, van de baan. De faculteit was ervan uitgegaan dat het jammer is dat studenten opnieuw tijd moeten besteden aan vakken die ze al gehaald hebben. Door alleen onvoldoende gemaakte vakken te herhalen zou de gemiddelde vertraging van de groep studenten kleiner moeten worden. We zien hier echter dat een voor studenten minder 'streng' examenreglement leidt tot meer studievertraging.⁴ Voor de cohorten 1976 en 1977 ging het even iets beter (figuur 2). Mijn verklaring daarvoor is dat, toen die studenten in de doctoraalfase verbleven, op landelijk niveau een discussie gevoerd werd over de twee-fasenstructuur. Met name was er in die tijd sprake van dat niet elke doctoraalstudent naar de tweede fase door kon stromen. Een dergelijk schrikbeeld blijkt een positieve invloed op de studiedoorstroom te hebben. Dit brengt mij bij de vierde factor die van invloed is op de doorstroomsnelheid van studenten.



Figuur 3. Finale rendementen propaedeutische Groningen versus rest

De invloed van *externe factoren*, met name de beslissingen die in Den Haag genomen worden. Vanaf 1982 treedt een sterke daling van het propaedeutiserendement op en ook het doctoraalrendement daalt steeds verder. De meest voor de hand liggende verklaring is dat dat te maken heeft met de invoering van de twee-fasenstructuur in 1982. Studenten kregen twee jaar tijd voor hun propaedeutische en in totaal zes jaar tijd voor het doctoraal. Die tijd wordt ook gebruikt, en, is het argument, daar kun je als faculteit niets aan doen. Stel dat deze redenering juist is, afgezien van het feit dat dat nu juist niet de bedoeling van de wetgever was, dan moet bij de andere medische faculteiten hetzelfde te vinden zijn. Het visitatierapport biedt de mogelijkheid dit nader te onderzoeken.⁵ Wanneer de Groningse doctoraalrendementen vergeleken worden met de gemiddelde rendementen van de zusterfaculteiten, dan blijkt dat de tendens dezelfde is: ook het totaal gemiddelde doctoraalrendement bij de zusterfaculteiten wordt in de jaren tachtig steeds lager. Per faculteit is het echter moeilijker de neerwaartse tendens terug te vinden. Vrijwel

elke faculteit laat schommelingen zien. Dit is bij het nominale propaedeutiserendement ook zo. Het totaal gemiddelde over de zusterfaculteiten laat weer een neerwaartse tendens zien, maar per faculteit zijn de schommelingen tussen de jaren groot. De conclusie kan zijn dat het lijkt dat de wettelijke maatregelen wel van (negatieve) invloed zijn geweest op de nominale rendementen van de medische faculteiten, maar dat er meer aan de hand moet zijn.

Het Groningse propaedeutiserendement

Wanneer het Groningse propaedeutiserendement vergeleken wordt met het rendement van de andere faculteiten, dan blijkt dat het Groningse rendement, na jaren een van de laagste geweest te zijn, de laatste jaren een van de hoogste is. Met de in Groningen genomen maatregelen is het gelukt de landelijke (medische) tendens te doorbreken (figuur 3).

Wat was er in Groningen aan de hand? In de zomer van 1987 waren zo weinig eerstejaars studenten geslaagd voor de propaedeutische dat de examencommissie zich zorgen ging maken. Als oplossing besloot de propaedeutische examencommissie voor het volgende cursusjaar (1987/1988) twee extra toetsen in te voeren. Ondertussen waren diverse facultaire gremia zich met het rendementsprobleem gaan bemoeien. Uiteindelijk werd een extern adviseur benoemd, en zijn voorstellen werden overgenomen en ingevoerd voor het cursusjaar 1988/1989. Ondertussen hadden de maatregelen van de examencommissie niet geleid tot een hoger rendement. De maatregelen van de extern adviseur daarentegen leidden wel tot een beter rendement. Deze maatregelen waren de volgende. De te toetsen stof werd regelmatig over het jaar verdeeld door aan het eind van elk trimester alle behandelde stof te toetsen in de vorm van deeltaetsen. Hiermee steeg het aantal toetsen van negen naar zestien. Voor elk vak werden de prestaties op de deeltaetsen verrekend tot één eindcijfer (tentamencijfer). De deeltaetsen van een vak waren onderling vol-

ledig compenseerbaar. In het voorgaande systeem moest het cijfer voor een deeltoets minimaal een vijf zijn. Er kwamen minder herkansingsmogelijkheden: alleen onvoldoende gemaakte tentamens konden worden overgedaan. Het herkansen voor deeltoetsen werd afgeschaft. Hiermee werd het aantal 'beslispunten' waarop het al dan niet geslaagd zijn voor de propaedeuse gebaseerd werd, verlaagd van negen naar zes tentamens. Vanaf 1990 werd het mogelijk één vak te compenseren (er mocht één vijf blijven staan, mits gecompenseerd door een zeven). En voor de reguliere tentamens werd de zogenaamde 70%-regeling ingevoerd. Deze regel houdt in dat, wanneer meer dan 30% van de studenten voor het reguliere tentamen dreigt te zakken, een onderzoek wordt gedaan naar de mogelijke oorzaak van dat resultaat (wat heel iets anders is dan de norm aanpassen). Aan de inhoud van het propaedeuseprogramma werd niets veranderd. Voor het cohort 1991 was het rendement na één studiejaar 72%. Bovengenoemde maatregelen zijn als volgt samen te vatten: het onderwijs werd beter georganiseerd en de tentamenregelgeving werd minder vrijblijvend.

Wanneer het erom gaat een verklaring te zoeken voor de schommelende nominale rendementen dan is mijn conclusie dat facultaire maatregelen op het terrein van de examenregels en de onderwijsprogrammering de meeste invloed uitoefenen op het nominale studierendement. Externe, landelijke maatregelen spelen wel een (negatieve) rol, maar maatregelen op facultair niveau zijn 'sterker'. Deze uitspraak wordt ondersteund door de bevinding dat, hoewel de finale rendementen van de verschillende faculteiten geneeskunde veel op elkaar lijken, er toch sprake is van statistische verschillen tussen de faculteiten. De oorzaak van deze verschillen moet ook gezocht worden in de organisatie rondom de examenregelgeving.⁶

Conclusie en discussie

In het rapport betreffende de studeerbaarheid van onderwijsprogramma's worden 83 punten genoemd die van invloed zijn op de studeerbaarheid en dus op het nominale rendement.⁷ In het rapport wordt geen melding gemaakt van de 'belangrijkheid' van deze punten voor de studievoortgang (nominale studierendement). Uit bovenstaande analyse van de nominale rendementen van de studenten geneeskunde te Groningen gedurende een periode van meer dan twintig jaar, blijkt dat de organisatie van het onderwijs en de examenregelgeving de belangrijkste factoren zijn waarmee het nominale rendement te sturen is. De invloed van zowel de onderwijsorganisatie als de examenregels op het rendement is al vaker aangetoond.⁸⁻¹⁰ Toch blijkt het vaak moeilijk hier in de praktijk iets mee te doen. Dit komt mijns inziens mede doordat vrijwel bij geen enkele studierichting een 'streefpercentage' voor het nominale (propaedeuse) rendement vastgesteld is. Mijn voorstel is dat wel te doen en bij geneeskunde ervan uit te gaan dat een nominaal rendement van 80% haalbaar moet zijn. De redenering voor deze 80% is de volgende. Het finale propaedeuserendement voor geneeskunde is bekend: 90% van de ingestroomde studenten slaagt voor de propaedeuse. Er is geen reden te bedenken waarom dat de komende jaren zal veranderen. In een curriculum dat jaarsgewijs van opzet is, en dat is bij de meeste medische faculteiten het geval, is het van belang dat zoveel mogelijk studenten volgens dit programma studeren. Met andere woorden, bij de propaedeuse moet het nominale rendement zo dicht mogelijk bij het finale rendement liggen. Het is niet reëel te verwachten dat het nominale rendement gelijk wordt aan het finale rendement. Immers, er zullen altijd studenten zijn die ziek worden, aanpassingsproblemen krijgen etcetera. Hier kan een faculteit niets aan doen. Faculteiten moeten er mijns inziens wel voor zorgen dat dit de enige vertragende factor blijft. Studenten behoren geen vertraging op te lopen door een onzorgvuldige planning van het

programma. Of door het zakken voor toetsen van inferieure kwaliteit, of door het feit dat er te veel beslispunten zijn waarop de studievoortgang wordt gebaseerd, of door hun tijd vooral te besteden aan het doen van de herkansingstentamens, etcetera. In de praktijk betekent dit dat in professioneel opgezet, goed functionerend onderwijs:

- de toetsen van voldoende kwaliteit zijn;
 - de toetsmomenten goed over het jaar verspreid zijn;
 - er voldoende toetsmomenten zijn die worden verrekend naar:
 - weinig beslismomenten (4-6 tentamens);
 - er niet te veel herkansingsmogelijkheden zijn.
- Wanneer het onderwijs met betrekking tot de examens op bovenstaande manier georganiseerd is, zullen de misstanden die eerder gesignaleerd zijn (te laat uitgebrachte klappers, etcetera), ook minder snel voorkomen.

De vraag is nu hoe ver het nominale rendement af mag liggen van het finale rendement. De ervaringen van de (Groningse) studieadviseurs zijn dusdanig dat het een redelijke schatting lijkt dat niet meer dan 10% van de studenten die de propaedeuse zullen halen, om persoonlijke redenen vertraging oploopt (ziekte en dergelijke). Dit betekent dat een propaedeuse goed functioneert als 90% van het uiteindelijke finale rendement binnen de nominale studieduur gehaald wordt. Voor de propaedeuse van geneeskunde betekent dit dat gestreefd moet worden naar een nominaal rendement van goed 80%. Elke medische faculteit die daar niet aan toekomt, moet zoeken naar mogelijkheden om de organisatie van het onderwijs en de examenregelgeving te optimaliseren. Voorwaarde is wel, dat dergelijke pogingen niet bemoeilijkt worden door politieke beslissingen.

Literatuur

1. Snijders JT, Wijnen WHFW, Henny CM. Het propedeutisch examen in de geneeskunde in relatie tot V.H.M.O.-resultaten. Groningen: Rijksuniversiteit Groningen, Psychologisch Instituut Heymans, 1965.
2. Van Noord F, Cohen-Schotanus J, Starren J. De effecten van selectie op eindexamencijfers en afschaffing van het jaarklassensysteem op het studierendement in de faculteit geneeskunde te Groningen. Universiteit en Hogeschool 1984; 30: 221-38.
3. Cohen-Schotanus J. Enkele rendementsgegevens vrouwen-mannen. Groningen: Rijksuniversiteit Groningen, Bureau Onderzoek van Onderwijs Geneeskunde, 1993. Interne notitie.
4. Cohen-Schotanus J. Admission procedures, Examination Rules and Study Success. In: Bender W, ed. Collaborative research in medical education. Groningen: Rijksuniversiteit Groningen, Bureau Onderzoek van Onderwijs Geneeskunde, 1993.
5. Visitatierapport Geneeskunde en Gezondheidswetenschappen. Utrecht: Vereniging van Samenwerkende Nederlandse Universiteiten, 1992.
6. Cohen-Schotanus J. De optimale verblijfsduur: studierendement bij acht medische faculteiten. In: Van der Vleuten CPM, Scherpier AJJA, Pollemans MC, eds. Gezond Onderwijs-I. Houten/Zaventem: Bohn Stafleu Van Loghum, 1992; 18-22.
7. 'Te doen of niet te doen?' Advies over de studeerbaarheid van onderwijsprogramma's in het hoger onderwijs. Zoetermeer: Ministerie van Onderwijs en Wetenschappen, 1992.
8. Vos P. Het ritme van het rooster. Onderzoek van Onderwijs 1992; 4: 51-3.
9. Jansen EPWA. Kan onderwijsprogrammering studiesnelheid beïnvloeden? Onderzoek van Onderwijs 1992; 4: 54-5.
10. De Gruyter DNM. Een propaedeuse zonder compensatie? Onderzoek van Onderwijs 1989; 4: 51-2.
11. Starren J. De beoordeling als hefboom voor onderwijsverbetering. De Psycholoog 1989; 25(3): 109-13.

DE AUTEUR

J. Cohen-Schotanus is als psycholoog werkzaam bij het Bureau Onderzoek van Onderwijs Geneeskunde, Faculteit der Geneeskunde, Rijksuniversiteit Groningen.

Correspondentie-adres:

J. Cohen-Schotanus, Bureau Onderzoek van Onderwijs Geneeskunde (Boog), Ant. Deusinglaan 1, 9713 AV Groningen.