

Performance-criteria in het onderwijs

J.A. Smal

Samenvatting

Er is behoefte aan criteria om de kwaliteit van het onderwijs van individuele docenten te kunnen beoordelen. In gebruik zijn: het volgen van cursussen, het behalen van onderwijskwalificaties, onderwijservaring en de kwaliteit van het gegeven onderwijs, gemeten aan meningen van studenten en collegae of onderwijsprijzen. Het gebruik van objectievere gegevens, zoals studierendement en marktwaarde van studenten en ontwikkeld onderwijsmateriaal, staat nog in de kinderschoenen.

Inleiding

Onderwijs moet aan de medische faculteiten concurreren met onderzoek en patiëntenzorg. Onderwijs verliest gewoonlijk de slag om de tijd en aandacht van de klinische docent.¹ Als een van de verklaringen daarvoor wordt gewezen op de duidelijke criteria die er bestaan om onderzoek te beoordelen, terwijl dergelijke criteria voor onderwijs ontbreken. De vraag dringt zich op of er soortgelijke criteria te ontwikkelen zijn voor het onderwijs in de hoop dat het prestige van het onderwijs daardoor verbetert.

Het lijkt een drogreden om te veronderstellen dat het bestaan van criteria op zichzelf het prestige van onderwijs zou verhogen. Deze gedachte is overigens niet eens zo onzinnig: er zijn talrijke voorbeelden aan te wijzen waarbij meetbaarheid leidt tot onderling vergelijken en ordenen. Neem het verwerven van een plaats in het Guinness Book of Records. Maar ook al is meetbaarheid op zichzelf geen oorzaak van meer prestige, in de praktijk geeft het gelegenheid tot gedifferentieerde beloning en daarmee de aanzet tot een grotere aantrekkingskracht.

Een vergelijking met de historie van de beoordeling van wetenschappelijk onderzoek is hier op haar plaats. Toen de voorwaardelijke financiering van wetenschappelijk onderzoek werd ingevoerd, bestonden er nog geen criteria om de kwaliteit van onderzoek te meten. In de toenmalige wetenschapscommissies stond men zelfs huiverig tegenover het gebruik van kwantitatieve methoden zoals citation indices. Maar er moesten criteria gevonden worden om te bepalen welk onderzoek wel geaccepteerd zou worden en welk niet. De noodzaak tot gedifferentieerde financiering leidde tot het opstellen van criteria.

Het ontwikkelen van criteria hangt nauw samen met gedifferentieerde financiering van het onderwijs. Op twee plaatsen waar op dit moment hard aan de ontwikkeling van beoordelingscriteria wordt gewerkt, is daaraan voorafgaand een besluit genomen om onderwijs mee te laten wegen in een universitaire carrière.^{2 3} Op dat moment wordt het noodzakelijk om aanvaardbare criteria te ontwikkelen voor de beoordeling van de kwaliteit van het onderwijs van een docent, zoals eerder gebeurd is bij de gedifferentieerde financiering van onderzoek.

De plaats waar onderwijsprestaties een rol moeten krijgen, is het beleid rond de bevordering van universitaire docenten naar een hogere salarisschaal. Tot nu toe spelen daarin voornamelijk onderzoeksprestaties een rol. In dit artikel wordt een onderzoek beschreven waarbij de vraag werd gesteld welke onderwijsprestaties op dit moment een plaats hebben in het carrièrebeleid en welke in de toekomst een rol moeten gaan spelen.

Methode

Aan de acht faculteiten geneeskunde in Nederland is gevraagd welke criteria destijds - 1995 -gebruikt werden voor de beoordeling van de onderwijsprestaties en bij bevordering van docenten.⁴⁻⁹ Daarnaast is gebruik gemaakt van de publicaties die over dit onderwerp in 1995 verschenen in *Academic Medicine*, een bekend internationaal tijdschrift over medisch onderwijs.^{3 10 11} Op twee congressen, The 7th Ottawa International Conference on Medical Education, gehouden te Maastricht juni 1996 en de Conferentie Professionalisering in universitair onderwijs te Ede september 1996, is gericht gezocht naar ervaringen op dit gebied.

Resultaten

Bij zes faculteiten bleken reeds initiatieven te bestaan om onderwijsprestaties in enigerlei vorm mee te laten wegen bij het bevorderingsbeleid van universitaire medewerkers. In een aantal gevallen betrof het voorlopige en nog niet goedgekeurde plannen. De verschillende methodes die hier en in de overige genoemde bronnen zijn vermeld, worden hieronder geordend in twee groepen: ingangskenmerken van de docent en beoordeling van de onderwijsprestaties.

Ingangskenmerken van de docent

Een aantal benaderingen richt zich op kwaliteiten die de docent moet bezitten voordat hij/zij in aanmerking komt voor een bevordering naar een hogere (salaris)schaal. Men abstraheert daarbij van de resultaten die in het onderwijs daadwerkelijk bereikt worden.

Docententrainingen

Bij twee faculteiten wordt het volgen van cursussen verplicht gesteld om toegelaten te worden tot een hogere salarisschaal. Wie in Maas-

tricht bijvoorbeeld tutor wil worden moet een tutortraining volgen; wie in schaal 13 wil komen moet een "kadercursus" volgen.⁹ En om een taak te krijgen in het curriculum 2000 in Groningen moet elke docent een daarop gerichte training volgen.⁷

Het is echter nog niet uit onderzoek gebleken dat door docententrainingen de kwaliteit van het onderwijs verbetert. Korte cursussen van minder dan twee dagen hebben geen aantoonbaar effect, en ook bij langere cursussen blijkt het moeilijk om aan te tonen dat de cursussen leiden tot ander gedrag en tot effectiever onderwijs.¹² Dit probleem is overigens niet alleen gesignaleerd bij nascholing van docenten, ook bij nascholing van artsen in het algemeen blijkt nauwelijks effect aantoonbaar te zijn.¹³

Behaalde onderwijskwalificaties

In Utrecht moet elke docent een zogeheten "basiskwalificatie onderwijs" behaald hebben en voor hogere functies is een "seniorkwalificatie" vereist. Bij de invoering van deze kwalificaties moesten eerst criteria geformuleerd worden waaraan de te kwalificeren docent moet voldoen. Deze criteria bevatten vakinhoudelijke, didactische en organisatorische bekwaamheden. Bij de formulering is begonnen met een algemene omschrijving op instellingsniveau; deze wordt per faculteit geconcretiseerd. Een algemene omschrijving luidt bijvoorbeeld: "De docent kan een student individueel begeleiden". In de faculteit geneeskunde is deze omschrijving verbijzonderd door de eis dat een docent "bedside teaching moet kunnen geven". In de praktijk bleken de omschrijvingen nog te vaag en moesten zij geoperationaliseerd worden in zichtbaar gedrag. Er is daartoe een lijst met voorbeelden opgesteld waaruit afgeleid kan worden of een docent over een gewenste bekwaamheid beschikt.¹⁴ Een van de vereiste bekwaamheden is als volgt geformuleerd: "de docent incorpo-

reert nieuwe wetenschappelijke ontwikkelingen in zijn/haar onderwijs". Deze bekwaamheid zou zich bijvoorbeeld kunnen uiten in het doen verschijnen van een "companion to Sterns", een tekst met aanvullingen op een gebruikt leerboek chirurgie. Tenslotte is het noodzakelijk om een beoordelingsprocedure te ontwerpen. In Utrecht geschiedt dit door een hiertoe per faculteit te benoemen beoordelingscommissie, die een kandidaat-docent beoordeelt aan de hand van een onderwijsportfolio. De onderwijsportfolio is een verzameling van voorbeelden van gegeven onderwijs, bijvoorbeeld de doelstellingen van onderwijsonderdelen, beargumenteerde aanpak, gebruikte toetsvormen, studentenoordelen, en eventueel video-opnamen van het onderwijs.¹⁵ Belangrijk is ook een reflectie, waarin de docent het onderwijs analyseert en eventueel nieuwe plannen aangeeft.

Ervaringsjaren

Bij minstens twee faculteiten hangt een bevordering mede af van het aantal jaren ervaring.^{5 9} Voor een bevordering naar schaal 10 wordt vier jaar onderrichtservaring geëist, voor schaal 11 is zes jaar ervaring nodig en voor schaal 12 ten minste twaalf jaar. Soms wordt dit aantal genuanceerd. Men spreekt over "toe te rekenen ervaringsjaren", en men kan met minder jaren ervaring toe als de beoordeling van het gegeven onderwijs "zeer goed" of "uitmuntend" is geweest. Op andere plaatsen - bijvoorbeeld Utrecht - wordt echter expliciet verboden om docenten op grond van ervaring te bevorderen. Dit zou namelijk strijdig zijn met het uitgangspunt dat bevordering moet geschieden op grond van onderwijskwalificaties.

Kwaliteit van prestaties

Andere benaderingen gaan niet uit van de a priori bekwaamheden van de docent, maar van

bewezen prestaties. Men spreekt in dit verband ook wel van "performance-criteria".

Het onderwijsrendement

In geen van de geraadpleegde nota's wordt reeds het behaalde onderwijsrendement als criterium voor goed onderwijs genoemd. Men is hierin waarschijnlijk terughoudend, omdat het rendement door veel andere factoren bepaald wordt: intelligentie, inzet van studenten, begin-niveau en externe factoren zoals studiefinanciering. Het rendement lijkt zeker niet afhankelijk van de inzet van één docent.

Omgekeerd valt niet te ontkennen dat in de hogere echelons van kwaliteitszorg, zoals in de politiek en bij de onderwijsvisitaties, veel belang wordt gehecht aan het studierendement. Op een congres van de Vereniging van Samenwerkende Nederlandse Universiteiten over professionalisering van docenten was opvallend hoe vaak -vooral door managers van onderwijsinstellingen - dit criterium het meest vanzelfsprekend wordt genoemd: men moet de kwaliteit van het onderwijs beoordelen door te kijken of het doel bereikt wordt. Op dit moment is het bereiken van het onderwijsdoel niet direct terug te voeren op de inspanningen van één docent. Maar het is niet ondenkbaar dat een docent of een groep docenten ooit als taak op zich neemt om een bepaald onderdeel van het Raamplan 1994 te realiseren en daarop beoordeeld wil worden.¹⁶

Het oordeel van studenten

Het blijkt een omstreden kwestie te zijn in hoeverre het oordeel van studenten betrokken moet worden bij de beoordeling van docenten. Bij minstens één faculteit wordt dit uitdrukkelijk afgewezen: "Evaluaties door studenten van docent-activiteiten zijn door hun subjectieve aard in het kader van een loopbaanbeleid van secundaire betekenis....".⁷ Ook elders acht men alleen de beoordeling door peer-groepen

aanvaardbaar.⁶ In Utrecht daarentegen geldt het omgekeerde. In de portfolio moeten studentenevaluaties opgenomen zijn. Voorstellen voor procedures waarin dit niet is opgenomen, werden niet goedgekeurd.

In de discussie over de waarde van studentenoordelen spelen twee vragen een rol. Hoe betrouwbaar zijn studentenoordelen en hoe valide zijn zij?

De *betrouwbaarheid* - ook wel *reproduceerbaarheid* genoemd - heeft vooral te maken met de omvang van de groep studenten. Niet iedere student heeft dezelfde opvattingen over wat goed onderwijs is. Uit onderzoek is komen vast te staan dat het oordeel van een groep van 35 studenten voldoende betrouwbaar is.¹⁷ Dit betekent dat een andere steekproef van 35 studenten uit hetzelfde studiejaar dezelfde uitkomst zou geven. Ook is voor een aantal docenten nagegaan in hoeverre hun beoordeling over de jaren schommelt. Studenten blijken in hun beoordeling uiterst stabiel. De rangorde verandert niet en ook de absolute scores tonen weinig variatie.

Validiteit betreft de vraag of de student wel in staat is te bepalen wat goed onderwijs is. Laat de student zich niet te veel leiden door irrelevante factoren zoals de zwaarte van het vak of het acteertalent van de docent? Er wordt wel beweerd dat de student de hoogste score geeft aan de meest populaire docent, die de minste eisen stelt en waar de student het minst bij leert. Minstens zestig jaar wordt al over deze kwestie gediscussieerd.¹⁸ In een meta-analyse kwam Marsh tot de uitspraak dat er in grote lijn een positieve correlatie is tussen het oordeel van studenten over docenten en hun leerprestaties.¹⁹ Docenten waarbij de studenten het meest leren, krijgen de hoogste waardering. Hieruit kan men afleiden dat studenten een goede bron van informatie zijn.

Beoordeling door peers

In verschillende nota's wordt voorgesteld do-

centen door collega-docenten (peers) te laten beoordelen. Men denkt daarbij aan een systeem waar enige ervaren docenten, bij voorkeur leerstoelhouders, het college van een collega bezoeken en daar een oordeel over geven.⁶ Ervaringen elders hiermee zijn niet onverdeeld gunstig. Een docent die op deze wijze werd gevisiteerd vertelt: "Opeens kwamen er drie grijze heren de collegezaal binnen. Alle studenten keken om. Ik was zelf opeens ook zenuwachtig. De sfeer werd wat gespannen, de studenten wisten niet goed meer hoe ze moesten reageren." Het is de vraag wat de waarde is van zo'n incidentele invasieve observatie.

Op andere plaatsen denkt men aan de blokcoördinator die beoordeelt hoe een docent zich van zijn/haar taak kwijt.⁷ Internationaal zie je vaak een combinatie van beoordeling door studenten en collegae. Peers laten zich in hun oordeel door andere factoren beïnvloeden dan alleen het onderwijs.²⁰ De klinische expertise, het prestige in onderzoek beïnvloeden de beoordeling sterk. Studentenoordelen zouden daarom veel correcter zijn. Bovendien is het in de praktijk veel veiliger om op studentenoordelen af te gaan. In de Verenigde Staten volgen nogal eens beroepszaken tegen beslissingen op loopbaangebied. Het oordeel van (een groep) studenten over een docent is veel sterker voor de rechtbank dan het oordeel van één onderwijsdecaan of enkele collega's.

Prijzen en nominaties

In veel faculteiten wordt een docent van het jaar gekozen. Op een faculteitsdag wordt die docent toegesproken en in het dankwoord mag hij of zij een bewijs van onderwijskwaliteit geven. Soms is er een geldbedrag aan verbonden, variërend van 500-5000 gulden.⁸ Elders krijgt de docent van het jaar geen geld, maar een sjerp of een kokarde.⁷ Op de Brown University School of Medicine in de Verenigde Staten is een vergelijkbaar systeem al jaren in gebruik. Alle docenten die bij de evaluatie op

een vijfpuntsschaal een 4,5 of hoger hebben, krijgen een button met het logo van het ziekenhuis en het jaartal.²¹ Er zijn docenten die elk jaar in de prijzen vallen en een reeks van deze buttons opspelden. Het doet denken aan de onderscheidingstekens die militairen dragen. Deze overeenkomst is niet toevallig. Bij militairen zijn deze distinctieven zo belangrijk omdat de uitgevoerde acties zelf onzichtbaar zijn. De eretekens zijn symbolen, waaraan elke kenner kan zien welke heldendaden iemand heeft verricht. Het gebruik van dergelijke onderscheidingen en paraferalia in het onderwijs is derhalve te verklaren uit het feit dat ook onderwijsprestaties weinig zichtbaar zijn.

Onderwijsmateriaal

Aan de Harvard Medical School is getracht objectieve criteria voor bevorderingen te ontleen aan onderwijsmateriaal.² De kwaliteit van dit materiaal zou, analoog aan de beoordeling van onderzoeksverslagen, beoordeeld kunnen worden door referenten of door het gebruik ervan aan andere medische opleidingen. Zo wordt er voor een benoeming tot assistant professor verwacht: "the development of teaching materials that significantly improve the quality of teaching". Van een associate professor wordt verwacht: "the *continued* development of teaching materials that make a *unique contribution* to the quality of teaching" en van een full professor: "the development of *original* teaching materials that are *widely used throughout the country*".

Discussie

De situatie op het gebied van onderwijs nu is vergelijkbaar met de situatie van dertig jaar geleden in het wetenschappelijk onderzoek. Het aanzien van onderzoek is sindsdien sterk gestegen, niet zozeer door extra financiering - de voorwaardelijke financiering was eerder een bezuiniging - maar mede dankzij de zicht-

baarheid van onderzoeksprestaties en de garantie van kwaliteit door peer-beoordeling. Dit lijkt ook de meest productieve weg op het gebied van onderwijs.

Onderwijs speelt zich op dit moment veelal af in het verborgene. Hoorcolleges zijn eenmalige prestaties, werkgroepbegeleiding is onzichtbaar, onderwijsontwerp blijft impliciet en onbekend. Wat zou er op tegen zijn om elk onderwijsprogramma eerst te laten beoordelen door experts, eventueel zelfs door collega's van buiten de faculteit? Een begin van intervisie is al gemaakt tussen enkele faculteiten. Een onderwijsprogramma kan uitgegeven worden, wellicht in een eigen tijdschrift. Vergelijk de tijdschriften op het gebied van docenten in het vwo. Nu weet bijna niemand wat er precies elders gebeurt.

Zonder veel inspanning is dit al mogelijk bij examenvragen. Bij de voorgaande visitatie is geadviseerd om vragen eerst te laten controleren door een commissie ter toetsing.²² In sommige steden is dat al gewoonte, in andere niet omdat het te veel tijd kost. Waarom kunnen dergelijke vragen niet uitgegeven worden, zoals in de Verenigde Staten jaarlijks gebeurt? Door ook de auteurs bij examenvragen te vermelden wordt een deel van de onderwijsinspanning zichtbaar.

Literatuur

1. Caellegh AS et al., redactie. Educating medical students. Assessing change in medical education -the road to implementation. Acad Med 1993; Suppl June.
2. Lovejoy FF, Clark MB. A promotion ladder for teachers at Harvard Medical School: experience and challenges. Acad Med 1995;70:1079-86.
3. Doceren aan de universiteit. Basiskwalificatie; universitair kader. Utrecht: Universiteit Utrecht, 1996.
4. Doceren aan de universiteit. Seniorkwali-

- ficatie; universitair kader. Utrecht: Universiteit Utrecht, 1996.
5. Functietypering onderwijsfuncties. Nijmegen: Besluit college van bestuur van 6 januari 1993, nr. CP612.342.
6. Docententraining in het AMC - de docent in opleiding. Amsterdam: AMC, Bestuursnotitie Faculteitsvergadering 13-11-95.
7. Loopbaanbeleid voor docenten. Groningen: Rijksuniversiteit, Faculteit der Geneeskunde, (zonder datum).
8. Nota Taakgroep Professionalisering. Amsterdam: Vrije Universiteit, mei 1995.
9. Docentprofessionalisering en personeelsbeleid onderwijs. Maastricht: Faculteit der Geneeskunde, Universiteit Maastricht, april 1997.
10. Bardes CL, Hayes JG. Are the teachers teaching? Measuring the educational activities of clinical faculty. *Acad Med* 1995;70:111-4.
11. Reiser SJ. Linking excellence in teaching to departments' budgets. *Acad Med* 1995;70: 272-5.
12. Nasmith L, Steinert Y, Franco ED, Saroyan A, Lawn N. Long-term impact of faculty development workshop: a pilot study. In: The 7th Ottawa international conference on medical education and assessment. Abstractboek. Maastricht: Universiteit Maastricht, 1996, 294.
13. Davis DA, Thomson MA, Oxman AD, Haynes RB. Evidence for the effectiveness of CME. *JAMA* 1992;268:1111-7.
14. Werkgroep Concretisering Criteria. Basis-kwalificatie Onderwijs; criteria nader belicht. Utrecht: Universiteit Utrecht UBP, 1996.
15. Tartwijk J van, Hoornweg J, Vries O de. Verslag van een literatuurstudie naar het gebruik van portfolio's als instrument voor de beoordeling van onderwijsgeven. Utrecht: IVLOS, 1994.
16. Metz JCM, Pels Rijcken-Van Erp Taalman Kip EH, Brand Valkenburg BWM van den. Raamplan 1994 artsopleiding: eindtermen van de artsopleiding. Nijmegen: Universitair Publikatiebureau, 1994.
17. Gijselaars WH. Kwaliteit van het onderwijs gemeten. Studies naar de betrouwbaarheid, validiteit en bruikbaarheid van studentenoordelen [proefschrift]. Maastricht: Rijksuniversiteit Limburg, 1988.
18. Rodin M, Rodin B. Student evaluations of teachers. *Science* 1972;171:1164-6.
19. Marsh HW. Multidimensional students evaluation of teaching effectiveness: a profile analysis. *J Higher Education* 1993;64;1-19.
20. Miller R. The improvement of teaching qualities in international perspective. In: Congresboek het onderwijs meester 1995. Utrecht: Universiteit Utrecht, 1995.
21. Metheny WP. A system for documentation and rewarding teaching contributions. In: The 7th Ottawa international conference on medical education and assessment. Abstractboek. Maastricht: Universiteit Maastricht, 1996:274.
22. VSNU-Visitatiecommissie Geneeskunde en Gezondheidswetenschappen. Onderwijsvisitatie geneeskunde en gezondheids-wetenschappen. Utrecht: VSNU, 1992.

DE AUTEUR

Dr. J.A. Smal is hoofd van de stafafdeling Onderwijs en Onderzoek van de Faculteit der Geneeskunde van de Universiteit Utrecht.

Correspondentieadres

Dr. J.A. Smal, Stafafdeling Onderwijs en Onderzoek, Faculteit der Geneeskunde, Postbus 80030, 3508 TA Utrecht.