

I.L.D. Houtman  
Onderzoeker

Correspondentieadres:  
Docentenopleiding,  
Vakgroep Psychologie  
Faculteit der  
Bewegingswetenschappen  
Vrije Universiteit  
van der Boerhorststraat 9  
1081 BT Amsterdam

**A**ls men het over de kwaliteit van de docent heeft, stelt men feitelijk de vraag of iemand een goede docent is. Wanneer is iemand nu een goede docent? Is iemand een goede docent als hij/zij veel kennis van onderwijskundige principes heeft, als de studenten hoge studieprestaties halen, als de studenten de docent goed vinden, als collega's dat vinden of als de docent dat zelf vindt? Welke vaardigheden en kenmerken van de docent bepalen of studentprestaties hoog zijn of beoordelingen gunstig uitvallen? Gaat het hierbij om zaken die te leren zijn of om min of meer persoons- (docent)gebonden kenmerken? Dit zijn zo enkele vragen die gesteld worden wanneer gesproken wordt over de kwaliteit van de docent. In dit artikel zal, aan de hand van literatuur, evaluatiegegevens van het Onderwijs Advies Bureau (OAB) aan de VU en enkele in eigen onderzoek verzamelde gegevens, op bovenstaande vragen worden ingegaan.

Waarom wordt afgemeten of iemand een goed docent is?

In de praktijk van het onderwijs en in het onderzoek naar doceer kwaliteiten blijkt dat op nogal uiteenlopende wijze wordt vastgesteld of iemand goed onderwijs geeft of niet.

Vier methoden die worden gebruikt om iets over de kwaliteit van de docent te zeggen, worden hieronder kort besproken: (1) het oordeel van studenten over de docent; (2) de studieprestaties van studenten; (3) zelfbeoordeling; (4) beoordeling door collega docenten.

#### 1. Het oordeel van studenten over de docent

Deze methode wordt ongetwijfeld het meest toegepast in de praktijk van het hoger onderwijs. Studenten vullen na afloop van een college(reeks) of cursus een evaluatieformulier in waarop zij allerlei vragen beantwoorden over het genoten onderwijs en over de docent die dat gaf. In de literatuur wordt dit als een bruikbare methode gezien wanneer men de bedoeling heeft om de docent of de faculteit informatie te verschaffen over de kwaliteit van de docent en eventueel over het verbeteren daarvan.<sup>1</sup> Het blijkt echter dat studentoordelen over de docent door zowel kenmerken van de student als kenmerken van de docent worden beïnvloed. Bij de interpretatie van studentoordelen moet hier dan ook rekening mee worden gehouden.

Ter verdediging van de bruikbaarheid van studentoordelen wordt onder andere de lange termijn stabiliteit van de studentoordelen aangevoerd. Correlaties tussen studentoordelen blijken na 1 jaar nog  $r = .83$  te zijn (deze waarden worden overigens alleen gevonden wanneer met klas- of groepsgemiddelden wordt gerekend; de stabiliteit van individuele beoordelingen ligt veel lager). Studentoordelen over eenzelfde docent blijken tevens generaliseerbaar voor verschillende cursussen. Marsh vond dat de beoordelingen van verschillende docenten die eenzelfde cursus gaven slechts in geringe mate samenhangen ( $r = -.05$ ), terwijl de studentbeoordelingen van eenzelfde docent in verschillende cursussen wel samenhangen ( $r = .61$ ).<sup>1</sup> Over de factoren die van invloed zijn op studentoordelen verderop meer.

#### 2. De studieprestaties van studenten

De redenering hier is even voor de hand liggend als simpel: als hij/zij goed onderwijs geeft zullen de studenten meer leren dan wanneer hij/zij slecht onderwijs geeft. Berk noemt echter een aantal redenen waarom studentprestaties geen goede indicatie geven van de kwaliteit van de docent.<sup>2</sup> Studentprestaties worden bepaald door tal van factoren waarop de docent niet of nauwelijks invloed heeft. Zo wordt er een sterke samenhang gevonden tussen de prestatie van de



student en diens intelligentie (correlaties variëren van  $r = .40$  tot  $.70$  maar kunnen - mits ze gebaseerd zijn op klassegemiddelden - zelfs stijgen tot  $r = .90$ ). Overigens blijkt de samenhang tussen het oordeel over de docent en de studentprestatie significant te zijn ( $r = .43$ ).<sup>3</sup> De twee tot nu toe besproken methoden om docentkwaliteit te meten hangen weliswaar significant samen, maar de samenhang is dus niet erg hoog. Men moet dan ook concluderen dat beide methoden slechts in beperkte mate eenzelfde aspect van de docentkwaliteit meten.

### 3. Beoordeling door docenten: zelfbeoordeling

De docent kan ook zelf een oordeel geven over zijn kwaliteiten als docent. In een literatuuroverzicht van Shrauger en Osberg wordt zeker zo veel voorspellende waarde toegekend aan zelfbeoordelingen als aan beoordelingen van anderen, zelfs van experts, wanneer het gaat om het oordeel over te leveren prestaties.<sup>4</sup> De auteurs merken echter op dat zelfbeoordelingen valide kunnen zijn, maar dit niet per se zullen zijn. Indien consequenties aan de beoordelingen verbonden zijn, zou de nauwkeurigheid van zelfbeoordelingen wel eens gevoelig kunnen zijn voor 'zelf-presentatie' effecten. Avi-Itzhak vond ten aanzien van de beoordeling van de docentkwaliteit alleen een overeenstemming tussen de zelfbeoordeling en het oordeel van studenten wanneer het ging om zeer concrete doceeractiviteiten, maar niet wanneer het om abstracte zaken ging (bijvoorbeeld het motiveren van studenten).<sup>5</sup> In zijn overzicht geeft Marsh aan dat correlaties tussen zelf- en studentoordelen variëren van  $r = .20$  tot  $.65$ .<sup>1</sup>

### 4. Beoordeling door collega-docenten

Marsh concludeert op grond van in de literatuur gerapporteerde onderzoeksresultaten dat het oordeel van collega's, gebaseerd op een bezoek aan een les, geen goede indicatie geeft van de kwaliteit van de docent. In methodologisch goed opgezette onderzoeken werd slechts een geringe overeenkomst tussen drie observerende collega's gevonden ( $r = .26$ ) en bleken collega-oordelen slecht te correleren

met andere maten voor docentkwaliteit. Gedragsobservaties van getrainde, externe observatoren, waarbij het dan gaat om het beoordelen van de helderheid van het betoog, enthousiasme, structuur van het college en dergelijke, bleken echter wel positief samen te hangen met studentoordelen.<sup>1</sup>

Tot slot kan worden opgemerkt dat de docentkwaliteit niet goed kan worden afgemeten aan de hand van de kennis van de docent over onderwijskundige principes, aangezien het bij doceren juist om het toepassen van deze kennis gaat. Overigens wordt kennis over een onderwerp wel vaak als een voorwaarde gezien voor het toepassen ervan.

## WANNEER IS IEMAND EEN GOEDE DOCENT?

Hierboven zijn vier meetmethoden besproken die worden gebruikt om iets over de kwaliteit van een docent te zeggen. Wanneer is iemand eigenlijk een goede docent? Het startpunt van de ontwikkeling van meetinstrumenten voor een bepaald gedrag of bepaalde kwaliteit komt in de regel voort uit een idee over of (theoretisch) model van het gedrag dat men wil meten. Wat wordt er nu van een docent verwacht? Een docent moet in ieder geval over gedegen inhoudelijke kennis van een vak beschikken om het te kunnen doceren. Het algemene idee is echter dat dit slechts één voorwaarde is om goed onderwijs te kunnen geven. Een docent moet het proces van het zich eigen maken van kennis en vaardigheden door studenten vergemakkelijken via verschillende onderwijsvormen zoals het geven van colleges, begeleiden van werkgroepen, practica of scripties, en het leerproces toetsen en beoordelen. Bij deze activiteiten doet de docent in verschillende mate een beroep op zijn/haar vaardigheden als het structureren van kennis, het presenteren van kennis, het motiveren van studenten, het formuleren van opdrachten, het werk van studenten beoordelen en hierover terugkoppeling geven aan die studenten (vooral bij het begeleiden van werkgroepen en practica en in mindere mate bij het begeleiden van

scripties), het herkennen en verhelderen van problemen (vooral bij het begeleiden van scripties), het maken van toetsen en het omgaan met studenten (tijdens en buiten de contact-uren).

Wanneer iemand over al deze vaardigheden beschikt en dit gepaard gaat met een grondige kennis van het vak dat de docent doceert, dan is de docent ongetwijfeld een goede docent. Het is de vraag of de genoemde methoden, waarmee wordt gemeten of iemand een goed docent is, een goede indicatie geven van deze kwaliteiten, en dus een goede operationalisatie zijn van het begrip 'goede docent'. Marsh houdt het er in zijn overzicht op dat er niet één methode is waarmee de kwaliteit van de docent afdoende wordt geoperationaliseerd.<sup>1</sup>

Hieronder ga ik in op de meest gebruikte methode, namelijk het oordeel van studenten over de docent. Daarna zal ik nog kort stil staan bij mogelijkheden om de vaardigheden die iemand tot een goede docent maken te trainen.

#### STUDENTOORDELEN OVER DE DOCENT

Het blijkt dat diverse factoren van invloed zijn op het oordeel dat studenten geven over het onderwijs dat zij krijgen en de docent die dat geeft. Een zo'n factor is de *interesse in het vak*. Van Os rapporteert enkele gegevens over de invloed van studentinteresse op het oordeel over de docent.<sup>6</sup> De intrinsieke motivatie of interesse van studenten is hoger voor kernvakken dan voor steunvakken en uit zich onder andere in hogere beoordelingen voor de docent en de cursus bij de kernvakken. Een kernvak is een zeer belangrijk/essentieel vak voor de betreffende studie dat dicht bij de 'kern' van de studie staat. Een steunvak daarentegen staat wat verder weg van deze 'kern' en betreft onderwerpen die als 'achtergrond' of 'oriënterende' informatie bij de 'kern' dienen. Een voorbeeld van een kern- en steunvak is het vak 'embryologie' in de propaedeuse bij de opleidingen Geneeskunde en Tandheelkunde. Embryologie is een kernvak in de propaedeuse van de Geneeskunde opleiding, maar een steunvak in de propaedeuse van Tandheelkunde (hier zou hooguit embryologie van het kaakgewricht als een 'kernonderdeel' kunnen worden beschouwd). Bij systematische evaluatie van medisch onderwijs aan de VU bleek dat de interesse

van de student minder samenhang met het oordeel over de docent wanneer het ging om een kernvak, dan wanneer het ging om een steunvak. Dit effect kon niet worden verklaard met het idee dat het een bepaald type docent was (of is) die de kern- of juist de steunvakken doceerde. De correlaties veranderden namelijk niet als alleen gegevens van docenten die een vak doceerden dat voor de ene studie een kernvak en voor de andere een steunvak was, werden geanalyseerd.

Kenmerken van de docent als *sexe*, *positie (status)* en *persoonlijkheidseigenschappen* blijken eveneens samen te hangen met het oordeel dat studenten over de docent(e) geven. De invloed van de sexe van de docent op het studentoordeel is regelmatig onderzocht. Hoewel het voor wat betreft de studentoordelen over de docent vaak niet uitmaakt of de docent (dan wel de student) een man of vrouw is, vonden Kierstead, D'Agostino en Dill in enkele experimentele situaties wel degelijk significante interactie effecten van de sexe van de docent met andere parameters: bij twintig mannelijke en twintig vrouwelijke studenten bleek de informatie over het feit dat de docent(e) veel of weinig sociale contacten had met studenten buiten de les om niet van invloed op het oordeel over de mannelijke docent, maar wel op dat over de vrouwelijke docent.<sup>17</sup> Ook het feit dat de docent(e) wel of niet glimlachte tijdens het college bleek uitsluitend bij vrouwelijke docenten van invloed op het oordeel van de studenten. Indien de vrouwelijke docent veel contacten had of veel glimlachte werd ze even goed beoordeeld als haar mannelijke collega. Indien ze dit gedrag echter niet vertoonde werd ze significant slechter beoordeeld dan haar mannelijke collega, ook al werd exact dezelfde les ten gehore gebracht. De oordelen van de mannelijke en vrouwelijke studenten bleken in het geheel niet van elkaar te verschillen. De onderzoekers concluderen dan ook dat vrouwelijke docenten niet alleen competent moeten zijn, maar zich ook moeten gedragen volgens stereotyped sexerolverwachtingen, als ze door hun studenten gunstig beoordeeld willen worden.

De positie van de docent in de opleiding blijkt eveneens van invloed op studentbeoordelingen. Assistenten worden in het algemeen lager beoordeeld dan de vast aangestelde docenten.<sup>13</sup> Ook persoonlijkheidseigenschappen blijken van invloed. Het gaat hierbij om

(in principe beïnvloedbare) docentkenmerken als expressiviteit, enthousiasme en zelfwaardering. Hoe hoger men op deze kenmerken scoort, des te gunstiger het studentoordeel. Een anekdotisch voorbeeld in dit verband is het Dr. Fox effect. Hiermee wordt bedoeld op de anekdote dat een acteur werd ingehuurd om een college te geven dat inhoudelijk niet zinvol was, maar dat zeer levendig, met veel anekdotes, gebracht werd. De beoordeling van deze acteur was zeer gunstig. Ook in methodologisch goed onderzoek blijken verschillen in expressiviteit of enthousiasme van de docent meer variantie te verklaren dan verschillen in inhoud. Het Dr. Fox effect werd echter niet ondersteund wanneer studenten van tevoren gezegd werd hun aandacht vooral op de inhoud van de les te richten, omdat zeer vervolgens op getest zouden worden en beloond zouden worden naargelang het aantal examenvragen dat ze goed hadden beantwoord.<sup>1</sup>

In eigen onderzoek bleek de geobserveerde gespannenheid van de docent van invloed op zowel het studentoordeel als op het oordeel van een deskundige. Aankomende docenten werden aan het eind van hun stagelessen op een HBO-opleiding door de aanwezige HBO-studenten beoordeeld. Tevens werden oefenlessen die zij op de universiteit voor mede-aankomende docenten en twee stafleden verzorgden door een deskundige beoordeeld op basis van een video-opname die van deze les was gemaakt. Tenslotte werden ook hun schoolpracticumbeggeleiders (ervaren docenten) beoordeeld door hun (HBO-)studenten na afloop van een door deze ervaren docenten gegeven les op de opleiding. Alle gevonden correlaties zijn negatief: hoe meer gespannen men de docent vindt des te slechter vindt men de docent en zijn of haar college. Het verband is weliswaar niet zeer sterk maar wel zeer consistent (circa  $r = -.25$ ). Zelfbeoordelingen van de docenten (zowel van de aankomende als van de ervaren docenten) over zijn of haar gespannenheid bleken niet samen te hangen met de beoordeelde gespannenheid. Uit dit laatste mag worden geconcludeerd dat het hier om een waarnemings- en beoordelingsfout gaat. Tenslotte blijkt ook het aantal studenten van invloed op het studentoordeel over de docent. Naarmate er meer studenten zijn (het publiek groter is) is het studentoordeel minder gunstig. De relatie lijkt echter niet geheel

lineair: als het aantal studenten de 200 overschrijdt lijkt het studentoordeel bij toenemende groepsgrootte weer gunstiger te worden.<sup>1</sup>

## EFFECTEN VAN DOCENTENTRAININGEN

Voor het (beter) leren beheersen van de vaardigheden die iemand tot een goed docent maken, worden door onderwijs-researchcentra en onderwijs-adviesbureaus trainingen georganiseerd. De vraag doet zich voor of deze vaardigheden ook werkelijk trainbaar zijn, met andere woorden of de in het algemeen kortdurende docententrainingen effecten hebben op docentgedrag.

De mate waarin deze docententrainingen effecten hebben op docentgedrag of op de beoordeelde kwaliteit van de docent is in Nederland echter weinig onderzocht. Een van de weinige Nederlandse onderzoeken op dit gebied is dat van Van der Sijde.<sup>8</sup> Hij bekeek de effectiviteit van een training van slechts 1 dag op een - voor de onderzochte training relevante - effectmaat: de tijd besteed aan een bepaalde lesfase in een standaard wiskundeles op de middelbare school. Hij vond in zijn, methodologisch goed opgezette, onderzoek inderdaad bewijs voor het beoogde effect. De trainingsgroep besteedde significant meer tijd aan een bepaalde lesfase dan de controlegroep, die niet de training had gevolgd maar wel het trainingsmateriaal toegestuurd had gekregen. In Amerikaans onderzoek is de vraag naar de effectiviteit van kortdurende docententrainingen beter uitgezocht.<sup>3, 9, 10</sup>

Dit onderzoek toont onder andere aan:

- dat 62% van het methodologisch goed uitgevoerde onderzoek hogere scores op de effectmaten voor docentkwaliteit (studentoordeel, studentprestatie en dergelijke) meldt na deelname aan een docentencursus;
- dat bij methodologisch iets minder vlekkeloos onderzoek (slecht onderzoek bleef buiten beschouwing) dit percentage stijgt naar 78%;
- dat de effecten verschillend zijn voor de verschillende meetmethoden. Zo vindt Wade bijvoorbeeld een tweemaal zo groot trainingseffect voor de methode van het observeren van docenten in vergelijking tot het effect op de studieprestatie;<sup>9</sup>
- dat cursussen meer effect hebben als de inhoudelijke voorbeelden die tijdens de

- cursus gebruikt worden, afkomstig zijn uit het vakgebied van de docent;
- dat docententraining effectiever is, als er goede directe feedback wordt gegeven, bijvoorbeeld op de in de cursus zelf gemaakte toets of op de video-opname van het proefcollege.

## CONCLUSIES

Uit de bovenstaande bespreking kan worden geconcludeerd dat docentkwaliteit vanuit verschillende aspecten, verschillende doceer-kwaliteiten, kan worden beoordeeld.

Studentoordelen kunnen, aangevuld met andere effectmaten (bijvoorbeeld in combinatie met studieprestaties, zelfbeoordelingen) een bruikbare effectmaat vormen voor doceerkwaliteiten. Er zijn echter zowel student- als docentkenmerken van invloed op de studentoordelen. Men kan zich overigens afvragen of dit laatste daadwerkelijk een verzwakking van de bruikbaarheid, eventueel zelfs validiteit van deze effectmaten betekent. Studentkenmerken (bijvoorbeeld interesse en intelligentie) en docentkenmerken (bijvoorbeeld enthousiasme) kunnen elkaar zeer goed wederzijds beïnvloeden: misschien is iemand inderdaad een betere docent en geeft hij beter les wanneer zijn studenten intelligent en geïnteresseerd zijn (of omgekeerd) en worden studenten geïnteresseerder als de docent expressiever, enthousiaster is en zichzelf waardeert, zelfvertrouwen heeft. De beschreven samenhang tussen de student- en docentkenmerken pleit wellicht enkel tegen het eenmalig gebruik van een enkele meetmethode of effectmaat om zicht te krijgen op de kwaliteit van een docent.

## LITERATUUR

1. Marsh HW. Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. *International Journal of Educational Research* 1987; 11 (3): 255-388.
2. Berk RA. Fifty reasons why student achievement gain does not mean teacher effectiveness. *Journal of Personnel Evaluation in Education* 1988; 1 (4): 345-64.
3. Kulik JA, Kulik C-LL. Meta-analysis in education. *International Journal of Educational Research* 1989; 13 (3): 223-340.
4. Shrauger JS, Osberg TM. The relative accuracy of self-predictions and judgment by others in psychological assessment. *Psychological Bulletin* 1981; 90 (2): 322-51.
5. Avi-Itzhak T. Teaching effectiveness as measured by student ratings and instructor self-evaluation. *Higher Education* 1982; 11: 629-34.
6. Van Os W. Evaluatie in het hoger onderwijs: controle op en verbetering van de kwaliteit van hoger onderwijs. Groningen: Wolters Noordhoff (Hoger Onderwijs Reeks), 1987.
7. Kierstead D, D'Agostino P, Dilt H. Sex role stereotyping of college professors: bias in students' ratings of instructors. *Journal of Educational Psychology* 1988; 80 (3): 342-4.
8. Van der Sijde PC. Training the teaching script. Almere: Versluys, 1987.
9. Wade RK. What makes a difference in in-service teacher education? a meta-analysis of research. *Educational Leadership* 1984; 42: 48-54.
10. Smit W. Effecten van docentencursussen, *Onderwijsinfo*, VU: Amsterdam 1987; 12 (3): 2.