

file: Docentbeoordeling '80
80-9-3_docentbeoordeling.rtf
80-9-3_docentbeoordeling.pdf

beoordeling onderwijsactiviteiten van docenten

het gebruik van empirische resultaten van het onderwijs.

Wanneer het bij de beoordeling van onderwijsactiviteiten gaat om de kwaliteit van die activiteiten, ligt het voor de hand na te gaan of en in hoeverre een dergelijke kwaliteit zich weerspiegelt in objectief vaststelbare resultaten van het onderwijs. De gegevens waaraan in dit verband gedacht kan worden zijn: slaagpercentages, tijdbesteding, opkomttpercentages en de mate waarin blijkens de toetsresultaten de specifieke onderwijsdoelstellingen gerealiseerd zijn.

Hoe voor de hand liggend ook, het is eveneens overduidelijk dat het hier niet

zozeer gaat om de mate waarin de docent een kwalitatief behoorlijke onderwijsinspanning heeft geplaagd, als wel om de mate waarin de studenten zich ingespannen hebben. De gegevens beoordelen niet zozeer de docent, als wel de studenten. Met dit argument zou de zaak al als afgedaan beschouwd kunnen worden. Toch zal een nadere toelichting en summiere uitwerking hieraan toegevoegd worden.

Slaagpercentages. Of voor het tentamen veel of weinig studenten (in één keer) slagen, zegt niets over de juistheid van de inrichting van de beoordelingssituatie. Sterker nog: wanneer het slaagpercentage in vergelijking met eerdere slaagpercentages verkregen door de voorganger van deze docent, gezakt of gestegen zijn, zegt dat op zich niets over de kwaliteit van de beoordelingssituatie of de toets. Slaagpercentages hangen af van de eisen die gesteld worden, en de studieinspanning van de studenten. Vooral de laatste zijn sterk afhankelijk van allerlei andere situationele gegevens die op de

totale onderwijssituatie betrekking hebben, zich ook uitstrekken buiten het onderwijs van de docent waar het om gaat.

Aan de ene kant betekent dit dat slaagpercentages niet zonder meer interpreteerbaar zijn of vertaalbaar naar doceerqualiteiten. Aan de andere kant betekent het dat slaagpercentages eenvoudig te beïnvloeden door bepaalde maatregelen die misschien de kwaliteit van het onderwijs schaden. Zou ondanks de methodologische bezwaren aan slaagpercentages een bepaald gewicht toegekend worden bij de beoordeling van de docent, dan ontstaat het risico dat de docent vooruitlopend op dit soort beoordeling bepaalde ongewenste maatregelen treft.

tijdbesteding. Ook al zouden harde tijdbestedingsgegevens beschikbaar zijn, dan volgt daar nog geenszins op eenvoudige wijze een conclusie uit m.b.t. de kwaliteit van de inrichting van het onderwijs, laat staan met betrekking tot de onderwijsactiviteiten van de docent. Er zijn

zeer grote persoonlijke verschillen in tijdbesteding, waarbij' het bovendien zo is dat er nauwelijks enig systematisch verband bestaat tussen verschillen in tijdbesteding en verschillen in behaald studieresultaat. Dat stelt het niet zo eenvoudige probleem wát precies de onderwijsdoelstellingen in dit opzicht zijn: een bepaalde gemiddelde tijdbesteding realiseren, ongeacht individuele verschillen, of juist proberen om individuele verschillen in tijdbesteding binnen bepaalde grenzen te houden? Het gekozen uitgangspunt zal ook gevolgen hebben voor de resultaten van je onderwijs, onderscheiden naar verschillende deelgroepen van studenten.

Uiteindelijk kom je dan terecht bij afspraken aangaande deze zaken zoals die voor álle docenten moeten gelden, waarna vervolgens de mogelijkheid in zicht komt om verschillende studieonderdelen met elkaar te vergelijken wat tijdbesteding/studieresultaten betreft. Maar ook dan worden niet de doceerkwaliteiten van docenten vergeleken, zodat voor de

beoordeling van die kwaliteiten al dit werk niet direct nodig is. Wat je vergelijkt is immers de inspanning die studenten zich voor verschillende vakken willen getroosten, en een grotere motivatie voor bepaalde vakken zal veeleer van de aard en inhoud van die vakken afhangen dan van de docenten die het vak verzorgen.

collegebezoek. Een opkomst van gemiddeld 100 % is meestal prachtig, maar zal door zijn absurde hoogte ook een aanwijzing zijn dat er misschien in dat college dingen gebeuren die oneigenlijk zijn, wanneer je ze tegen het licht van de onderwijsdoelstellingen houdt. Met andere woorden: ook hier zeggen opkom, 5tpercentages op zich heel weinig.

mate waarin doelstellingen bereikt worden. Ook hier geen cijfers die een vanzelfsprekendheid van zichzelf hebben. Het is mogelijk heel 'goede' resultaten te bereiken door de doelstellingen heel bescheiden te formuleren, of door bepaalde maatregelen te nemen die op zich de kwaliteit van het onderwijs

schaden.

In het algemeen kan aan het voorgaande nog toegevoegd worden dat effecten van onderwijsverbeteringen maar zelden aan de hand van empirische gegevens aantoonbaar zijn. Een zee van onderwijsresearch zou te noemen zijn als illustratie bij deze opmerking. Wat daarbij de achterliggende redenen of mechanismen zijn is niet altijd even duidelijk. Dat het om spectaculaire verschillen kan gaan, die desondanks empirisch verduisterd kunnen worden, wordt aardig geïllustreerd door onderzoek waarbij het gemiddelde niveau van de toegelaten studenten drastisch gemanipuleerd wordt (door selectie), en dan toch gemiddelde studieresultaten **op hetzelfde niveau** blijven. Het kan zijn dat docenten niet in staat zijn om dit soort verschillen ook op reële wijze in hun beoordelingen en cijfergewoonten te verwerken (denk aan de Wet van Posthumus), het kan zijn dat studenten zich in hun inspanningen aanpassen aan het niveau dat docenten impliciet van hen

verwachten.

Concluderend kan alleen maar vastgesteld worden dat empirische resultaten van het onderwijs geen vanzelfsprekende rol kunnen en mogen vervullen bij de beoordeling van de onderwijskwaliteiten van individuele docenten.

evaluatieve meningen van studenten.

In de Verenigde Staten worden docenten zwaar beoordeeld op de wijze waarop zij hun onderwijstaken uitvoeren. Daarbij valt een sterk accent op het gezamenlijk oordeel van zijn studenten, zoals dat blijkt uit vragenlijst onderzoek dat routinematig wordt uitgevoerd. Hoe zwaar er aan deze wijze van evalueren getild wordt, mag blijken uit het verschijnsel dat er vrijwel geen aflevering van Amerikaanse onderwijsresearch journals afkomen waarin niet op de één of andere wijze een onderzoek naar deze wijze van evalueren gerapporteerd wordt. Het is dan ook niet verwonderlijk dat van die Amerikaanse

traditie een sterke druk uitgaat op bijvoorbeeld de Europese landen om dezelfde methoden te gebruiken. Ook in Nederland werd, eind van de zestiger jaren, wel gewerkt met 'collegevragenlijsten', zij het dat de bedoeling daarbij niet was gelegen in de beoordeling van de docent, maar in het zoeken naar mogelijkheden (of knelpunten) om zijn onderwijs te verbeteren. De ervaringen met die collegevragenlijst zijn niet erg gunstig geweest. Er waren ruw gezegd twee categorieën 'afnemers': docenten die intensief bezig waren met hun eigen onderwijs zo goed mogelijk gestalte te geven, en docenten die aan de hand van de vragenlijstuitkomsten wilden laten zien dat hun college uitstekend was. Aan de ene kant dus docenten die de lijst niet nodig hadden, aan de andere kant docenten die de lijst oneigenlijk gebruikten. Deze ervaringen, gekoppeld aan de problematiek dat er geen noodzakelijke identiteit is tussen de wensen van studenten en de beste richting waarin naar onderwijsverbetering gezocht moet worden, hebben geleid tot het niet meer

hanteren van deze methode in ons land.

Nu er behoefte ontstaat aan een 'beoordelingsinstrument' voor het vaststellen van de kwaliteit van de wijze waarop de individuele docent zijn onderwijs verzorgt, wordt wederom de vraag gesteld of dergelijke vragenlijstmethodes daarbij bruikbaar zouden zijn, of gebruikt zouden mogen worden. Het antwoord daarop moet ondubbelzinnig ontkennend luiden.

Allereerst moet er uitdrukkelijk op gewezen worden dat wat in de Verenigde Staten gebruikelijk is, niet zonder meer naar ons eigen onderwijs overgeplant kan worden. Instellingen van hoger onderwijs in de Verenigde Staten opereren veelal als een soort winkelbedrijf:

zij moeten financieel rond zien te draaien uit betaalde collegegelden, en zijn dan ook afhankelijk van het aanbod van studenten, en van de mate waarin aangekomen studenten hun studie aan dit instituut blijven voortzetten. 'Daar past dan ook de gedachte bij dat het onderwijs zo ingericht moet zijn dat het voor de studenten aantrekkelijk is, en wel in die mate dat de

instelling daarin concurrerend is met andere instellingen. Eén van de middelen in die concurrentiestrijd is er voor te zorgen dat het docentencorps bestaat uit personen die hoog scoren in de opinion polls die onder de consumenten gehouden worden. Het hoeft geen verder betoog dat vragenlijsten die hun oorsprong vinden in deze nogal typisch Amerikaanse setting, in het Nederlandse onderwijs volstrekt onbruikbaar zijn.

De literatuurstudie van Blom en Langerak (1979) over Amerikaans onderzoek van studenten oordelen over docenten zet bovenstaande conclusie nog eens extra scherp aan door ook de methodologische tekortkomingen van deze vragenlijstmethoden te evalueren als te groot om de methoden bruikbaar te doen zijn. Hun uiteindelijke conclusie is dat "*de oordelen van studenten van onvoldoende kwaliteit zijn om een belangrijke rol te spelen in beslissingen rond aanstelling en bevordering*".

Een aantal van de methodologische

bezwaren die Blom en Langerak signaleren, zijn ook van belang bij het overwegen van andere methoden om doceerkwaliteiten beoordeelbaar te krijgen. Bijvoorbeeld zal het ook bij andere beoordelingsmethoden moeilijk of onmogelijk zijn om de kwaliteiten van de docent te scheiden van typische kenmerken van het vak dat hij of zij doceert. De mate waarin zo'n scheiding niet goed gemaakt kan worden, maakt een d e r g e l i j k e b e o o r d e l i n g o o k onaanvaardbaar.

Het oordeel van studenten kan en mag in normale omstandigheden geen rol spelen bij de beoordeling van de medewerker waar het om zijn onderwijstaken gaat. Studenten zijn niet bijzonder deskundig waar het er om gaat de wijze waarop de docent zijn onderwijs heeft ingericht en uitvoert, te beoordelen. Daarover hoeft geen verdere discussie gevoerd te worden. Er moeten echter twee dingen aan toe gevoegd worden.

Ten eerste: opvattingen van studenten over

de inrichting van het onderwijs zijn op zich natuurlijk w&l relevant. Daar zitten ófwel aanwijzingen in over de richting waarin je aan je onderwijs zou kunnen gaan sleutelen wanneer de docent zich in die opvattingen ook kan vinden, ófwel er zit de noodzaak in om in geval de docent een andere opvatting heeft, tot onderling« overleg en uitwisseling van argumenten te geraken.

Ten tweede: in bijzondere gevallen kan het oordeel van studenten ook voor de beoordeling van belang zijn. Het is duidelijk dat een docent die op onbehoorlijke wijze zijn of haar onderwijs verzorgt, daarvoor in eerste aanleg door de eigen studenten op de vingers getikt zal worden. Het gaat dan echter om gedragingen die zich laten kenmerken als *onbehoorlijk bestuur': het niet nakomen van afspraken, het niet uitvoeren van vastgelegde onderwijstaken, willekeurig handelen waar het beoordeling van studenten betreft. Voor deze extreme situaties kun je natuurlijk niet een vragenlijstmethode routinematig gaan hanteren; het volstaat om binnen de

subfaculteit de communicatiekanalen van studenten naar bestuur open te houden.

Blom, S.J.M. en Langerak, W.F. Beoordelen van docenten door studenten. Pedagogische Studiën, 1979, 56, 308318.

objectieve kwaliteitsmaatstaven voor verzorgd onderwijs.

Aan welke maatstaven denk je dan, wanneer je spreekt over objectieve kwaliteitsmaatstaven voor verzorgd onderwijs? Zijn die maatstaven objectief? En wat hebben die maatstaven met de docent te maken?

De **maatstaven**. Je kunt daarbij aan twee dingen denken: minimum condities waaraan voldaan moet zijn, en kwaliteiten die boven de minimum condities uitgaan.

Wat de **minimum condities** betreft: je kunt daarbij stelselmatig alle onderwijsactiviteiten (ook de voorbereidende) langs gaan, en minimumcondities daarbij formuleren. Een

soort superminimum conditie is dan dat het onderwijs inderdaad **voorbereid** is, dat de docent niet een houding heeft van 'ach, we zien het wel wanneer we met het praktikum bezig zijn', of 'ik bedenk in de koffiepauze wel welk onderwerp op het college te behandelen'. In vrijwel alle situaties kan de voorbereiding ook uit schriftelijke stukken blijken.

Voor specifieke activiteiten kunnen specifieke minimum condities wel genoemd worden, bijv, voor toetsen dat er een minimum aan doorzichtigheid geboden wordt, dat de zakslaag grens die gekozen is ook beargumenteerd wordt, e.d. Maar wat schiet je met dergelijke minimum condities op m.b.t. de beoordeling van de docent? Daarover straks dus.

En dan **kwaliteit**, wat je zou kunnen omschrijven als alles dat boven de minimum condities uitgaat. Je kunt daar prachtige verhalen over aan het papier toevertrouwen, op welke manieren het allemaal beter kan dan minimaal, met eventueel daar nog rangordeilingen bij, etcetera. Om bij het toetsvoorbeeld te blijven: h e t g e v e n v a n

proeftoetsgelegenheid, het publiceren van een uitgebreid toetsvragenboek, het tevoren bekend maken van de cesuur, het informeren over slaagkansen die de studenten hebben wanneer ze niets meer doen aan het niveau dat ze bij de proeftoets al hebben, etcetera.

Zijn die maatstaven **objectief**? Voor een belangrijk deel denk ik dat je die maatstaven objectief kunt maken, wanneer ze dat van zichzelf al niet zijn. Wat veel moeilijker objectief te maken is, is de inspanning die de docent zich heeft te getroosten om bepaalde kwaliteiten te realiseren. En daarmee komen we op het derde topic:

Wat hebben die maatstaven met de **docent** te maken? Om te beginnen met de minimum condities; deze zijn allerminst vanzelfsprekend, in veel gevallen zal blijken bij het formuleren van minimum condities dat een groot deel van het studieprogramma niet aan die condities voldoet. Om te beginnen moeten er dan ook afspraken over die minimum condities

worden gemaakt, veelal op faculteitsraad niveau. Daarbij moet ook worden vastgelegd waaruit zal blijken of aan de minimum conditie voldaan is. En dan zitten we al midden in het afspraken stelsel dat in dit rapport ook voorgesteld wordt voor de beoordeling van de onderwijsactiviteiten van de docent. Bij alles wat boven de minimum kwaliteiten, sorry, condities uitgaat kom je al helemaal in de problemen. Wat moet je immers met een stuk onderwijs dat hier en daar superkwaliteiten heeft, maar waarvan duidelijk is dat die kwaliteiten niet op de relevante punten aanwezig zijn. M.a.w.: 'kwaliteit' is mooi, maar niet altijd relevant, en 'kwaliteit' kost geld, inspanning, middelen, die ook anders aangewend hadden kunnen worden.

Daaruit volgt dat je de docent er niet aan mag wagen de beoordeling louter achteraf op 'kwaliteit' te laten plaats vinden. Misschien heeft hij of zij dan jarenlang hard gewerkt aan het bereiken van 'kwaliteit' op punten die de beoordelings commissie achteraf minder relevant blijkt te vinden, of

zelfs veroordeelt als verkeerd geïnvesteerde energie.

Wil je met kwaliteitsmaatstaven werken, wat geen onzinnige zaak is, dan moet dat in een kader gebeuren dat los staat van de beoordeling van de docent. Kwaliteitsverbeteringen moeten in lijn lopen met het vakgroeps beleid en het fakulteits beleid, indien enigszins mogelijk. Omdat de beoordeling van de docent echter ook aan de orde is, kun je daar best 'kwaliteitsmaatstaven' in betrekken, maar dan alleen via tevoren vastgelegde afspraken waarin activiteiten en producten zijn vastgelegd met wederzijdse instemming. De kwaliteitsmaatstaven die in afspraken worden vastgelegd moeten uiteraard objectief zijn, maar in speciale gevallen kun je ook wel subjectieve maatstaven hanteren wanneer er voor tussentijdse bijstellingen ruimte gegeven wordt.

Het oordeel van collega's.

Wordt er met een beoordeling door collega's gewerkt, dan wordt daarmee bedoeld dat de kwaliteiten van de onderwijsactiviteiten van de docent' (achteraf door die collega's beoordeeld worden. In de voorgaande paragraaf werd al vastgesteld dat het achteraf aanleggen/hanteren van kwaliteitsmaatstaven niet behoorlijk mogelijk is. In veel sterkere mate geldt dat bezwaar natuurlijk waar het gaat om oordelen van **subjectieve** aard waar het hier om gaat bij het collegaoordeel.

Daar komen dan nog andere bezwaren bij, zoals de ondekundigheid van docenten op onderwijskundig gebied, en de mate van subjectiviteit van de oordelen waar het dan om gaat.

Ontbreken van deskundigheid. Bij het beoordelen van onderwijsactiviteiten hebben we met een heel andere situatie te maken dan bij het beoordelen van onderzoekactiviteiten, Bij het laatste zijn de collega's uit de vakgroep immers bij uitstek deskundig, bij het laatste zijn ze (bijna) even grote semi professionals als de

beoordeelde. Bedenk dat het juist bij de beoordeling van kwaliteit zo ontzettend belangrijk is daarbij de totale onderwijssituatie te betrekken, inclusief de gedragingen van de studenten, het karakter van de leerstof, e.d.. Welnu, doe je dat, dan moet je van heel goede onderwijskundige huize komen om daarbij geen onzorgvuldigheden te begaan.

Onontkoombare subjectiviteit. Hoe je het ook wendt of keert, altijd zullen dit soort oordelen subjectief zijn, verschillen al naar de persoon van de beoordelaar en zijn of haar specifieke opvattingen. Je kunt daar maar heel weinig ten goede aan veranderen door de beoordelaars uitgebreide beoordelingschalen ter beschikking te stellen (gesteld al dat je dat verantwoord zou vinden en zou willen). Dat kan niet andere betekenen dan dat je op deze wijze niet kunt en niet mag beoordelen.

Het oordeel van collega's is wel van belang, maar op andere wijze. Het oordeel van collega's is van belang waar

het gaat om het beleid van de vakgroep. In die zin moet dat oordeel ook enigszins tot uiting komen in de onderwijsactiviteiten van de docent. De juiste weg daartoe is via afspraken vooraf, waarbij concrete producten, objectieve (en doorzichtige) maatstaven worden afgesproken.

Het werken met inventarisatielijsten á la Crombag.

De lijst van Crombag behelst onderwijsactiviteiten die de docent verricht zou kunnen hebben. De vragen daarover zijn van een zoveel mogelijk feitelijk karakter.

Zo'n lijst is vanzelfsprekend een prima, en waarschijnlijk ook minimaal noodzakelijk, hulpmiddel bij de beoordeling. Maar dan niet in de vorm van invulling en beoordeling achteraf, maar **vooraf**. Gebruik dezelfde checklist dan ook als jaarlijkse of tweejaarlijkse **afspraa]ak** (nadat voor ieder item zo ondubbelzinnig mogelijk is ingevuld wat de docent

daarvoor aan activiteiten zal verrichten).

In het werken met zo'n vragenlijst zit ook een element van **tijdbesteding**, niet alleen in de zin van over welke activiteiten de docent zijn tijd verdeeld heeft, maar ook of de docent een redelijke hoeveelheid tijd **in totaal**, en in vergelijking tot onderzoek en bestuur, besteed heeft (zal besteden) aan onderwijsactiviteiten. Op het punt van tijdbesteding zul je echter met onzekerheden moeten werken, omdat tijdbesteding zowel moeilijk als niet controleerbaar is, en het veelal ook niet eenvoudig zal zijn redelijke precieze schattingen van benodigde tijd te maken (bijv. tijd nodig voor voorbereiden colleges, voor het schrijven van een syllabus, voor het verzorgen van de toetsing van itemschrijven tot en met feedback). Je zult wat tijdbesteding betreft dan tot afspraken in redelijkheid moeten komen, met de mogelijkheid van tussentijdse bijstelling wanneer de omstandigheden en ontwikkelingen daar aanleiding toe geven.

De lijst die Crombag geeft is maar een

voorlopig probeersel. We zullen in een meer volledig advies een betere lijst moeten geven. Waarschijnlijk zul je ook moeten benadrukken dat de afspraken niet in dusdanige termen gesteld hoeven te zijn dat ook voor buitenstaanders onmiddellijk duidelijk is hoe het met de aard en de zwaarte van de afgesproken activiteiten staat. Bijv.: het herzien van een syllabus zal al naar gelang de omstandigheden een zeer kort, dan wel uitgesproken lang tijdsbeslag vergen, en daar hebben de collega's in de vakgroep uiteraard goed zicht op, maar in het algemeen de buitenstaander niet. Dat betekent dat sommige kwantificeringen naar buiten toe eigenlijk slechts als pseudokwantificeringen zullen werken, terwijl zij intern heel goed zullen functioneren.

bijdrage aan kwaliteitsverbetering van het onderwijs

Het onderwijs is nog altijd een beetje een ondergeschoven kindje voor de wetenschappelijk medewerker. Dat is voor

die medewerker vervelend, en komt het onderwijs ook niet ten goede. De waardering voor de medewerker is immers vnl. waardering voor zijn werk als onderzoeker, niet als onderwijzer.

Een goede beoordelingsprocedure bewerkstelligt dan ook dat het onderwijs een meer eigen plaats krijgt in het denken en doen van de medewerker. Creativiteit en inspanning die de medewerker in zijn onderwijsactiviteiten steekt moeten ook als zodanig erkenning kunnen vinden. Het is echter niet zonder meer duidelijk hoe je daarbij tewerk zou kunnen gaan, omdat het zonder meer in de beoordeling een groter gewicht aan de onderwijsactiviteiten toekennen niet als vanzelfsprekend tot gevolg zal hebben dat aan dat onderwijs dan ook meer en betere aandacht gegeven zal worden. Zeker een beoordelingachteraf zal in dit opzicht alleen maar teleurstellend werken.

Wil je de beoordeling laten bijdragen tot de kwaliteitsverbetering van het onderwijs, dan lijkt dat eigenlijk alleen maar goed

mogelijk door die kwaliteitsverbetering van tevoren af te spreken in het werkplan. Daarbij kan de betrokken docent natuurlijk een grote inbreng leveren door zijn suggesties op welke punten het onderwijs verbeterd zou kunnen worden. Hoe je dat dan concreet uitwerkt is een zaak van alle betrokkenen tesamen, waarbij eventuele advisering door het COWO op de zakelijk onderwijskundige punten zijn goede diensten zou kunnen bewijzen.

Het COWO kan die advisering voor een deel automatiseren, door het aanleveren van wederom een mooie checklist waarop de meest voorkomende en meest bruikbare suggesties kort vermeld staan, met een eventuele bronnenvermelding (COWO rapporten en boeken bij voorkeur) waar meer gedetailleerde informatie gewenst zou kunnen zijn. Een voorlopige uitwerking van zo'n checklist zou bij ons advies (aan de CPD) gevoegd kunnen worden. Een punt dat daarbij onze bijzondere aandacht zou moeten hebben is het volgende:

Z o v e e l m o g e l i j k m o e t e n d i e

kwaliteitsverbeteringen die door de docent worden bewerkstelligd, ook in producten resulteren, enigszins analoog aan zijn onderzoeksrapportage.