

wordt. In de toekomst zal geëvalueerd moeten worden of de handleiding het beoogde effect heeft gehad.

Literatuur

- Mulder S, Schoorl J. Voorstel onderwijs op het gebied van palliatieve zorg. Intern Rapport. Maastricht: Faculteit der Geneeskunde, 1995.
- Buckman R. Communication in palliative care. In: Doyle D, Hanks GWC, MacDonald N. Oxford textbook of palliative medicine. Oxford: Oxford University Press, 1997.
- Kubler-Ross E. Lessen voor levenden, gesprekken met stervenden. Baarn: Ambo, 1978.
- Sporken P. Heb jij aanvaard dat ik sterven moet? Baarn: Ambo, 1981.
- Doyle D. Survey on the current status of professional education in palliative care. Genève: Education Committee of the European Association for Palliative Care (EAPC), 1992.
- Mason C, Howell K. Dealing with dying patients: difficulties and strategies in final year medical students. Death Studies 1988;12:9-20.
- Dozor RB, Addison RB. Towards a good death: an interpretive investigation of family practice residents "practice with dying patients". Family Medicine 1992;24:538-43.
- Dalen J van, Zuidweg J, Collet J. The curriculum communications skills teaching at Maastricht medical school. Med Educ 1989;23:55-61.
- Nieuwenhuijzen EH van den, Campen KWM van, Boot RJ, Leeuwen YD van. Terminale zorg in de medische opleiding: een unieke stage. Medisch Contact 1995;50:1084-5.
- Crebolder HFJM, Meesmakers JFM, op 't Raot JMH, Bartholomeus P, Boshuys P, Boshuizen HPA. Patiëntgebonden onderwijs in de huisartspraktijk binnen de artsopleiding: het programma van de Rijksuniversiteit Limburg. Ned Tijdschr Geneesk 1996;140(25):1320-3.

DE AUTEURS

*F.H.M. van Vollenhoven en D.M.S. Ras zijn studenten geneeskunde aan de Universiteit Maastricht.
Prof. dr. H.F.J.M. Crebolder, hoogleraar Capaciteitsgroep Huisartsgeneeskunde, Universiteit Maastricht.
Dr. Y.D. van Leeuwen, hoofd huisartsopleiding, Universiteit Maastricht.*

Correspondentieadres:

Prof. dr. H.F.J.M. Crebolder, Capaciteitsgroep Huisartsgeneeskunde, Faculteit der Geneeskunde, Universiteit Maastricht, Postbus 616, 6200 MD Maastricht.

Nieuws uit Philadelphia: indrukken van de achtste Ottawa-conferentie

A.J.J.A. Scherpbier, C.P.M. van der Vleuten

Samenvatting

Het Ottawa-congres is vanaf het eerste congres door een delegatie uit Nederland bezocht. Om de lezers van dit tijdschrift te informeren worden in dit artikel de organisatie en de inhoud van het laatste congres in Philadelphia beschreven. Tot slot doen de auteurs op basis van de congresinformatie suggesties voor aandachtsgebieden op het gebied van onderwijsontwikkeling en onderzoek.

Inleiding

Het begrip Ottawa-conferentie is bekend voor degenen die dit tijdschrift al langere tijd lezen.¹⁻⁴ Omdat er ook veel nieuwe lezers zijn, wordt eerst de onstaansgeschiedenis van deze conferentie kort beschreven. De eerste Ottawa-conferentie werd in 1985 georganiseerd in Ottawa (vandaar de naam). Het was een idee van Ian Hart, een endocrinoloog uit Ottawa. Tijdens een "sabbatical leave" in Dundee raakte hij overtuigd van de voordelen van het objectieve structured clinical examination (OSCE) voor het toetsen van klinische competentie. Om de uitwisseling van ervaringen met het stationsexamen, zoals het OSCE in Nederland meestal wordt genoemd, te stimuleren organiseerde hij samen met Ronald Harden uit Dundee de eerste Ottawa-conferentie. Deze conferentie werd bezocht door ongeveer 180 deelnemers uit diverse landen. Het was een succes, mede door de prettige sfeer en het klimaat dat Ian Hart wist te creëren. Er werd dan ook besloten om deze conferentie regelmatig te organiseren. De afgelopen jaren werd de conferentie op verschillende plaatsen in Noord-

Amerika en Europa georganiseerd (1987 Ottawa, 1989 Groningen, 1990 Ottawa, 1992 Dundee, 1994 Toronto, 1996 Maastricht).

De meeste voordrachten op de eerste twee conferenties in Ottawa gingen over toetsing. Bij de derde conferentie in 1989 in Groningen werd meer aandacht aan onderwijs besteed. Deze verbreding van de thematiek is bij de volgende conferenties gehandhaafd. De volledige titel van het congres in Philadelphia was dan ook "The Eighth International Ottawa Conference on Medical Education and Assessment" met als ondertitel "Evolving Assessment: Protecting the Human Dimension". Deze ietwat plechtige ondertitel benadrukt het belang van toetsing voor de kwaliteit van de gezondheidszorg. Hoewel er wat ons betreft eigenlijk had moeten staan "Education and Assessment", is het terecht dat de organisatoren wijzen op het belang van investeringen in de kwaliteit van de dokters van de toekomst.

In dit artikel wordt de organisatie van het congres beschreven en wordt aandacht besteed aan de op het congres gesignaleerde ontwikkelingen. Op basis van de subjectieve (de auteurs) impressies van deze ontwikkelingen wordt een aantal aandachtsgebieden voor de komende jaren geformuleerd.

Organisatie

De conferentie werd georganiseerd door de National Board of Medical Examiners van 12-15 juli in Philadelphia. Er waren 561 deelnemers, waarvan 28 uit Nederland. De organisatie had een Ottawa-website geopend en abstracts konden uitsluitend via deze website verkregen worden. Helias was er vaak filevor-

ning op de elektronische snelweg. Wel konden in het Cyber Café tijdens de conferentie abstracts geprint worden. Het niet uitbrengen van een abstractboek is voor de organisatie natuurlijk een aardige bezuiniging. Hoewel een abstractboek slechts een tijdelijke functie heeft, ging dit ons, als deelnemers, toch net iets te ver. De conferentie werd georganiseerd in een hotel met conferentieruimten, zoals gebruikelijk in Noord-Amerika. De indeling van de conferentie was over het algemeen traditioneel: pre-conferentieworkshops, plenaire sprekers en parallelle sessies. Bijzonder was dat er veel gelegenheid was om experts te ontmoeten in "meet-the-experts sessions" tijdens ontbijt en lunch. De parallelle sessies waren onderverdeeld in large-group presentations (15 minuten spreektijd) en small-group presentations (5 minuten spreektijd). Er waren in totaal 71 large-group presentations in vier parallelle sessies, en 371 small-group presentations in vijftien parallelle sessies. Het ontbrak aan een centrale tijdsbewaking, zodat wisselen gedurende de parallelle sessies vrijwel onmogelijk was. Gelukkig was er van alle small-group presentations ook een poster, zodat het rustig bekijken van de informatie en eventueel praten met de auteur een goed alternatief voor deze sessies was. Naast deze traditionele congressvormen waren er vier "expert panels". Hierbij gaven steeds drie experts hun mening over een bepaald onderwerp, waarna discussie met de aanwezigen in de zaal kon plaatsvinden.

- De onderwerpen waren:
- Toepassing van toetsing op basis van casus in de beroepsopleiding;
- Waarde van toetssystemen voor de controle of sturing van de internationale mobiliteit van professionals in de gezondheidszorg;
- Effectiviteit van programma-evaluatie met betrekking tot verbetering van de onderwijskwaliteit;
- Gevolgen van nationale examens voor het onderwijs in de faculteiten.

Door de grootte van de groep, dominantie van "experts" in de zaal, de inrichting van de ruimtes, et cetera kwam de bedoeling van deze panels niet altijd even goed uit de verf.

Voor de lezers die het Gezond Onderwijs Congres (GOC) wel eens bezocht hebben: ons GOC wint de vergelijking met het Ottawa-congres wat betreft organisatie.⁵ Het niveau van de presentaties was gevarieerd, net als op het GOC. U hoeft dus niet ver te reizen voor een goed congres. Het gemiddelde niveau van de presentaties was op het Ottawa-congres echter hoger dan op het GOC. Met name was de onderzoeksvraag vaak beter ingebed in de bestaande literatuur.

Wat betreft de keuzes van de presentatievormen was het Ottawa-congres iets te veel "lecture-based". Aan de ene kant begrijpelijk, omdat ervoor gekozen is om veel abstracts aan te nemen, aan de andere kant verwacht je op een onderwijscongres iets meer aandacht voor details die de kwaliteit van informatieoverdracht beïnvloeden en meer ruimte voor interactie. Een eenvoudige te realiseren verbetering is het vervangen van de small-group presentations door thematische postersessies. De posters van een dergelijke sessie hebben een gemeenschappelijk thema en hangen in de ruimte waar de sessie plaatsvindt. Nadat de deelnemers de posters hebben bekeken, wordt iedere poster door de auteur kort toegelicht. Tenslotte vindt een discussie over het thema plaats. Deze thematische postersessies zijn bedacht door Henk Schmidt en op verschillende congressen is gebleken dat deze presentatievorm goed functioneert.

Ontwikkelingen

Gezien de omvang van het congres, wordt hier slechts een aantal ontwikkelingen besproken. Als er voorbeelden worden genoemd, wordt de naam van de eerste auteur genoemd. De lezers die meer willen weten over deze voorbeelden

kunnen dan te zijner tijd de proceedinge raadplegen.

Het gebruik van computers

Op veel plaatsen worden zowel voor onderwijs als voor toetsing casus ontwikkeld die studenterend werd over de resultaten van een samenwerkingproject waarin een simulator, Harvey, en een aantal interactieve modules op cd-rom, U-medic, werden ontwikkeld (Gordon, Issenberg). Deze programma's worden inmiddels in 75 centra gebruikt en de effectiviteit is aangetoond. Het is ons niet bekend of Harvey en U-medic ook in Nederland gebruikt worden. Zo niet dan is het wellicht een idee om contact op te nemen met deze collega's. Er wordt ook onderzoek verricht naar de mogelijkheden om studenten te toetsen met behulp van deze programma's (Issenberg).

Vanaf 1999 zullen computercasus gebruikt worden in het landelijk artsexamen in de Verenigde Staten (Orr). De computercasus zijn complementair aan andere onderdelen van het landelijk artsexamen, zoals multiple-choice-toetsen en toetsing met behulp van simulatiepatiënten. Het programma is zodanig opgezet dat de prestaties van de kandidaat automatisch gescorwd kunnen worden. Uit een experiment met een aantal van deze casus bleek helaas dat een aantal praktiserende artsen die in de praktijk niet goed functioneerden niet significant lager scoorden dan een groep goed functionerende artsen (La Duca). Naar de oorzaak van deze teleurstellende bevinding wordt verder onderzoek gedaan.

Een ander voorbeeld van casus werd door Schuwirth gepresenteerd. Hij heeft computer-casus ontwikkeld op basis van de "key feature approach". Deze aanpak betekent dat experts overleggen over een patiëntenprobleem en gezamenlijk zoeken naar de essentiële besluiten bij dit specifieke probleem. Deze besluiten worden in een toets verwerkt. Op de conferen-

tie presenteerde hij een onderzoek waarbij met hardop-denkp protocollen (een onderzoekstechniek uit de cognitieve psychologie) nagegaan werd hoe ervaren artsen casus toetsen en feitenvragen oplossen. Het bleek dat met casus-toetsen inderdaad hogere cognitieve vaardigheden worden getoetst dan met feitenvragen.

Er waren ook presentaties over het gebruik van de computer als itembank (De Jong). De tijd dat een itembank alleen een geautomatiseerde Kaartenbak was, is voorbij. Zo beschreef Daming Zhu een techniek om met behulp van de computer gelijkwaardige toetsen te genereren.

Een van de grootste problemen bij het toetsen is het probleem van de casusspecificiteit, waarmee bedoeld wordt dat de prestatie van een examenkandidaat op het ene probleem geen voorspellende waarde heeft voor de prestatie op een ander probleem. De computer wordt daarom steeds vaker gebruikt als hulpmiddel om in korte tijd veel casus te kunnen toetsen. Dankzij de toepassing van multimedia kan de realiteit steeds beter benaderd worden. Als er ook nog een key feature approach en automatische scoring worden toegepast, zijn de voordelen van de computer - als hulpmiddel - evident. De kosten van het ontwikkelen van de casus zijn echter hoog. Interfaciltaire samenwerking ligt dan ook voor de hand.

De groepen uit Texas (Papa) en Calgary (Haraşym) zoeken naar andere oplossingen van het probleem van de casusspecificiteit. Met behulp van kunstmatige-intelligentiesystemen zoeken zij naar elementen in casus die voorspellende waarde hebben voor de diagnostische competentie. Zij hopen daardoor met minder casus betrouwbare uitspraken te kunnen doen. Via dezelfde systemen werken zij ook aan de mogelijkheid om met behulp van de computer van één casus veel verschillende casus te maken.

Bepalen van de slaag/zakgrens

Het bepalen van de grens tussen voldoende en onvoldoende is een lastig probleem. In Nederland wordt ook vaak gediscussieerd over de normen, maar uiteindelijk heeft hier het examenreglement waarschijnlijk een grotere invloed op de voortgang van de student dan het slaag/zakbesluit voor één tentamen.⁶ In Noord-Amerika zijn de landelijke examens van grote invloed op de carrière van de studenten. Het exact (met zo min mogelijk fouten) bepalen van de slaag/zakgrens is daarom uitermate belangrijk. Dillon rapporteerde op de conferentie over het systeem dat is opgezet om het bepalen van de grens te bewaken. Populair voor het bepalen van slaag/zakgrenzen is de methode van Angoff of een van de zeer vele varianten hierop (Downing, Grosso, Van der Steeg). Bij de Angoff-methode wordt aan een aantal experts gevraagd om van elke toetsvraag aan te geven welk deel van de "borderline" studenten geacht wordt de vraag goed te scoren. De methode kost erg veel tijd en dat is een van de redenen dat er veel varianten op bedacht zijn. Een van de mogelijkheden om de kosten te beperken is dat de toetsproducenten de grens bepalen. Uit onderzoek van Grosso blijkt dat de grens van de toetsproducenten overeenkomt met de grens van onafhankelijke experts. In stationsexamens blijken de observatoren redelijke overeenstemming over de slaag/zakgrens te kunnen bereiken (Margolis, Klass, Reznick).

Stationsexamens

Sinds de eerste Ottawa-conferentie is het stationsexamen een wereldwijd verspreide "examenvorm". De inhoud van stationsexamens is echter heel gevarieerd. Het stationsexamen is dan ook eigenlijk geen examenvorm, maar een organisatievorm. Er kan bijvoorbeeld een onderscheid gemaakt worden tussen OSCE's, veel korte stations waarbij deelaspecten van een probleem worden getoetst, en SP-based

testen, waarbij in elk station een simulerende patiënt aanwezig is en het oplossen van het patiëntprobleem als geheel wordt beoordeeld. De grote inhoudelijke variatie betekent dat in feite bij elke publicatie over stationsexamens het betreffende examen zodanig beschreven moet worden dat de lezer zich een goed beeld van kan vormen. Met het organiseren van stationsexamens voor grote aantallen deelnemers hebben de Canadezen de meeste ervaring. Sinds 1993 is het stationsexamen een onderdeel van het landelijk artsexamen in Canada. Voor ongeveer 1800 kandidaten wordt op veertien plaatsen tegelijk een stationsexamen (in twee talen) georganiseerd. Het examen duurt twee dagen en bestaat uit twintig stations van tien minuten met een simulatiepatiënt. De beoordeling geschiedt door artsen. Uit onderzoek bleek dat 70% van de examenkandidaten zo goed presteerde dat er geen twintig stations nodig waren maar dat ook met tien stations betrouwbare beslissingen over hen genomen konden worden. Deze onderzoeksbevinding betekent dat sequentiële toetsing overvloedig kan worden. Bij sequentiële toetsing wordt eerst met een relatief korte toets een besluit genomen over de zeer goede en eventueel de zeer slechte studenten. De resterende "middenmoot" krijgt een verlengde toets, omdat het moeilijker is om betrouwbare besluiten over deze groep te nemen. Op de conferentie rapporteerden de Canadezen (Reznick, Dale Dauphinee, Smeel) over hun eerste ervaringen met sequentiële toetsing bij het landelijke stationsexamen. Logistiek was het een ingewikkelde organisatie, omdat op de eerste dag bepaald moest worden welke kandidaten nog te horen dat er veel tijd besteed moest worden aan voorlichting van studenten en docenten, omdat er veel onrust was over dit experiment (zoals bij elke verandering). Het is nog onduidelijk of dit experiment uiteindelijk de standaardprocedure wordt. Het grote voordeel is uiteraard dat het een gigantische kostenbespa-

ring oplevert. In 2001 wordt in de Verenigde Staten gestart met een stationsexamen als onderdeel van het artsexamen (Klass). Men verwacht minimaal 18000 examenkandidaten per jaar. Er wordt intensief samen gewerkt door de National Board en de faculteiten om de organisatie van dit examen te kunnen realiseren. Wij zijn van mening dat we in Nederland een voorbeeld moeten nemen aan de moeite en de energie die de Noord-Amerikaanse collegie besteden aan het toetsen van competentie.

De vraag of kandidaten beoordeeld moeten worden met gedetailleerde "checklists" of met "global ratings" kwam in een aantal presentaties aan de orde (Weinreb, Hodges, Ziv, Boulet). Eerdere berichten uit de literatuur dat globale beoordelingen de voorkeur verdienen, werden in deze presentaties bevestigd. Hodges toonde zelfs aan dat ervaren artsen beter scoorden dan co-assistenten als er globale lijsten werden gebruikt en slechter als er gedetailleerde lijsten werden gebruikt.

Hoewel in veel publicaties geconcludeerd wordt dat met behulp van een stationsexamen (van voldoende lengte) betrouwbare uitspraken over kandidaten gedaan kunnen worden, zijn er nog vele vragen over de betrouwbaarheid onopgehelderd. Als bijvoorbeeld stationsexamens op grotere schaal worden afgenomen, is het gebruikelijk om gedurende het examen de casus te veranderen om te voorkomen dat de kandidaten aan elkaar vertellen wat er getoetst wordt. Martz en Battles presenteren onderzoek waarin nagegaan werd hoe de moeilijkheidsgraad van casus gecorrigeerd kan worden. Andere vragen die op de conferentie aan de orde kwamen: Welke invloed heeft de subjectieve indruk over de examenkandidaat op de score (De Champlain)? Is het ras van simulatiepatiënt en examenkandidaat van invloed op de prestatie (Collier)? Een kanttekening bij deze laatste twee vragen is dat het in de echte praktijk natuurlijk ook zo is dat de relatie tussen arts en patiënt invloed heeft op het arts-patiëntcontact. Ons inziens behoort de

blauwdruk van de toets (inclusief de samenstelling van de simulatiepatiënten) representatief te zijn voor de echte praktijk en dan was het genoemde onderzoek eigenlijk niet nodig geweest. Over een dergelijke blauwdruk was er een presentatie van Ziv. Zij gebruiken statistische gegevens over de gezondheidszorg om representatieve casus voor het stationsexamen te selecteren.

Kwaliteit in de praktijk

Gedurende de afgelopen jaren is er geleidelijk meer aandacht gekomen voor onderzoek naar het functioneren van artsen in de praktijk. Er waren bijvoorbeeld twee voordrachten (Epstein, Gorter) over studies waarbij simulatiepatiënten "undercover" artsen bezochten en daarbij de kwaliteit van het feitelijk handelen beoordeelden. De uitkomsten van dergelijk onderzoek zijn van belang voor de kwaliteit van de zorg, maar zij kunnen ook vertaald worden naar de nascholing, de beroepsopleiding en de basisopleiding.

Over een andere methode om de kwaliteit te beoordelen rapporteerden de Australische collegae (Page) Een van de voorwaarden om Fellow van het Royal Australian College of Physicians te blijven is het zogenaamde Practice Quality Review. Een dergelijk review betekent dat er collegae op bezoek komen in de praktijk, consulten bijwonen, et cetera. Tot nu toe zijn 45 van deze reviews gedaan. Hoewel het veel werk is, zien de betrokkenen de reviews als een belangrijke bijdrage aan het handhaven van de kwaliteit.

Op eerdere conferenties is uit Noord-Amerika gerapporteerd over signalering en scholing van slechtfunctionerende artsen. Gesignaleerd worden bijvoorbeeld artsen die obsoliete medicamenten of verkeerde doseringen voorschrijven. Op deze conferentie werd gerapporteerd over een systeem dat in Engeland is opgezet om slechtfunctionerende dokters op te sporen, te beoordelen en bij te scholen (Jolly, Southga-

te, McAvoy). De eerste indruk is dat het heel grondig is opgezet. Er is bijvoorbeeld goed gebruik gemaakt van ervaringen in het buitenland, klachten van patiënten worden geanalyseerd, et cetera. Onze conclusie is dat wij in Nederland van deze aanpak wel wat kunnen leren.

Onderwijs- en toetsontwikkeling

Zelfbeoordeling is bij elke vorm van scholing van groot belang. Als iemand goed in staat is om zelf zijn sterke en zwakke kanten te beoordelen, kan hij de zwakke kanten verbeteren, gericht bijlessholen, meer lezen, et cetera. Helaas is uit de literatuur bekend dat de individuele zelfbeoordeling meestal niet erg accuraat is. In twee voordrachten (Gupta, Lujck) werden deze bevindingen bevestigd. Harder bepleitte, ons inziens terecht, meer aandacht voor zelfbeoordeling in het onderwijs. Zelfbeoordeling moet eigenlijk gewoon regelmatig geoefend worden.

In de literatuur wordt de laatste jaren meer aandacht besteed aan leren en werken in de praktijk. Raghoebar-Krieger doet onderzoek naar de invulling van de co-assistentenschappen met behulp van logboekjes. Deze logboekjes zijn een richtlijn voor de studenten en zij geven de docenten inzicht in wat de co-assistenten doen. Op deze conferentie rapporteerde zij over een studie waaruit bleek dat de manier waarop het logboekje wordt toegepast verschillen oplevert in het aantal door de co-assistenten gerapporteerde ervaringen. Uit onderwijsvoogpunt is het praktisch als het logboekje gebruikt wordt om tijdens het co-assistent-schap bij te sturen. Een ander probleem bij leren in de praktijk is dat studenten zelden geobserveerd worden tijdens het arts-patient-contact. Lane rapporteerde over structureerde klinische observaties tijdens het co-assistent-schap kindergeneeskunde. Voor deze observaties zijn formulieren ontworpen en de docenten zijn van tevoren getraind in het observeren van studenten en het geven van feedback. De eerste evaluatieresultaten zijn zeer

positief. Heel interessant was dat de docenten van mening waren dat deze observaties de voortgang van hun klinisch werk niet te veel belemmerden. Een interessante ontwikkeling met betrekking tot toetsing zijn observaties van het arts-patientcontact. Abouna rapporteerde over een geobserveerd klinisch examen waarbij studenten een aantal echte patiënten zien van verschillende disciplines. De beoordeling geschiedt door interne en externe examinatoren. De eerste ervaringen gedurende twee jaar waren positief. McKinley presenteerde de eerste ervaringen met de Leicester Assessment Package (LAP) bij het beoordelen van studenten in de huisartsenpraktijk. Elke student ziet zes ongelecteerde patiënten en wordt daarbij door twee examinatoren met de LAP beoordeeld. Tot nu toe hebben 106 kandidaten dit examen gedaan en het blijkt dat de observaties logistiek haalbaar zijn en dat er betrouwbare uitspraken gedaan kunnen worden over de prestaties van studenten.

Aandachtsgebieden onderwijs en onderzoek

De verschuiving van docentgecentreerd onderwijs naar studentgecentreerd onderwijs betekent in grote lijnen dat de docentnaak verschuift van informatie-overdracht naar het creëren van een optimale leeromgeving.⁷ De consequentie van deze verschuiving voor onderwijsontwikkeling en onderzoek van onderwijs is dat we ons zouden moeten concentreren op deze leeromgeving. Deze uitspraak komt wellicht over als een "open deur", maar is het niet, want we weten nog maar heel weinig over de leeromgeving en al helemaal weinig over een optimale leeromgeving. Omdat leeromgeving echter zo'n ruim begrip is dat er nauwelijks sprake kan zijn van concentreren, wordt in het onderstaande een aantal specifieke aandachtsgebieden beschreven.

Leren en werken in de praktijk

Van de beschikbare tijd voor de opleiding tot arts wordt ongeveer 30% besteed aan de praktijk, de co-assistentenschappen. Deze fase van de opleiding is van groot belang voor de competentie. Er is al veel onderzoek gedaan naar het meten van de competentie, zoals onder andere blijkt op de Ottawa-conferenties. Op een aantal faculteiten wordt nu ook onderzoek gedaan naar de leeromgeving waar een groot deel van de competentie wordt geleerd. Dit onderzoek is erop gericht om te achterhalen wat co-assistenten doen, wat zij leerzaam vinden, welke ziektebeelden zij tegenkomen, et cetera.⁸ Samenattend, gaat het dus om een inventarisatie van de leeromgeving. Voor de verdieping van dit aandachtsgebied zijn de resultaten van cognitief-psychologisch onderzoek van belang, met name onderzoek naar expertiseworming. De uitdaging voor de komende jaren is ons inziens dan ook om de expertise en onderzoekstechnieken van artsen, onderwijskundigen en psychologen te combineren. Informatie over de "black box" van het praktisch onderwijs kan dan gezamenlijk worden ingewonnen en vanuit verschillende invalshoeken worden geïnterpreteerd. Op grond van wat tot nu toe bekend is, speculeren wij (de auteurs) dat een en ander leidt tot een herziening van de traditionele inroostering en invulling van de co-assistent-schappen en een nieuwe plaats voor de praktijk in de gehele opleiding. Hoe verantwoord deze speculatie is, zal de toekomst leren.

Relatie tussen onderwijs en toetsing

Het effect van examens en toetsen is dat studenten gaan studeren. Over het algemeen is dat ook de bedoeling en op zich is daar niets mis mee. Toetsing heeft een belangrijk sturend effect op het leren van de student. Over het toets- en examenstelsel wordt zelfs wel gesteld: "It's the tail that wags the dog", waarbij "the dog" het curriculum is. Er is echter nog weinig bekend

over de precieze invloed van toetsing op de leeromgeving. Bijvoorbeeld, als de voortgang van studenten alleen wordt beoordeeld met toetsen waarin kennis van geïsoleerde feiten getoetst wordt, is het waarschijnlijk dat studenten hun studeergedrag hierop afstemmen en alleen feiten gaan leren. De vraag is of dat wel effectief studeergedrag is. Het is namelijk bekend dat de halfwaardetijd van geïsoleerde feiten in het geheugen kort is, zeker als de feitenkennis na de toets niet gebruikt wordt. Terecht wordt er tot nu toe veel aandacht besteed aan betrouwbare en valide slaag/zakbesluiten, maar tegelijkertijd is er, ten onrechte, (te) weinig aandacht voor de onderwijseffecten van toetsing op het leren en de leeromgeving. De uitdaging voor de komende jaren is ons inziens dan ook om de gevolgen van inhoud, vorm en planning van toetsing voor de leeromgeving te bestuderen. Ook is het belangrijk dat studenten leren om zichzelf goed te beoordelen. De bedoeling van dit aandachtsgebied is dat de diagnostische waarde van toetsing niet meer beperkt blijft tot slaag/zakbesluiten, maar dat de feedback van nut wordt voor het individu en voor het curriculum, en toetsing dus - zo nodig - gaat leiden tot een therapeutisch effect voor de incompetentente studenten en voor constructiefouten van het curriculum.

De praktijk van de gezondheidszorg

Er is de afgelopen jaren met grote inzet gewerkt aan de ontwikkeling van toetsen om de competentie van studenten en artsen zo goed mogelijk te beoordelen. Bij deze ontwikkeling is er een duidelijke verschuiving herkenbaar van het geïsoleerd toetsen van feiten en (deel)vaardigheden naar realistische, geïntegreerde toets-situaties die zo goed mogelijk overeenkomen met de praktijk. Ook in het onderwijs vindt een vergelijkbare verschuiving plaats. Bij deze ontwikkelingen is de praktijk van de gezondheidszorg een soort jipspunt. Helaas is dit jipspunt nog niet zo goed gedefinieerd. We weten

Road observatory 7925, Cape Town, South Africa, e-mail: sally@medline.uct.ac.za.

Literatuur

1. Rossum HM van, Schepbier AJA, Hoeks TH, Hiemstra RJ. Verslag Conferentie Ottawa en werkboek McMaster University. Bulletin Medisch Onderwijs 1986;5(1):25-32.
2. Lunsen HW van, Schepbier AJA. Toetsing in bewering: congressverslag van The Second Ottawa Conference on Assessing Clinical Competence, 27-30 juni 1987. Bulletin Medisch Onderwijs 1988;7(1):7-9.
3. Beeck JA van, Leeuwen Y van, Reihans JJ. Indrukken van TICTAC. Bulletin Medisch Onderwijs 1989;8(4):126-30.
4. Leeuwen Y van. Vaart in vaardigheden. Bulletin Medisch Onderwijs 1990;9(3):82-3.
5. Moll J. Gezond Onderwijs 1994: een volwasen congres. Bulletin Medisch Onderwijs 1994;13(4):184-6.
6. Cohen-Schotanus J. Effecten van curriculumveranderingen. [proefschrift]. Groningen: Rijksuniversiteit Groningen, 1994.
7. Dolmans DHJM, Wolthagen HAP, Vleuten CPM van der, Wijnen C. Wanneer is onderwijs studentengereed? Medisch Contact 1997;52(29):1045-7.
8. Vleuten CPM van der, Schepbier AJA, Dolmans DHJM, Schuwirth LWT, Verwijnen GM, Wolthagen HAP. Het artsexamen getoetst. Bulletin Medisch Onderwijs 1998;17(2):78-93.
9. Reihans JJ. Does competence predict performance? [proefschrift]. Amsterdam: Thesis publishers, 1991.

DE AUTEURS

Dr. A.J.J.A. Schepbier, arts, hoofd Skillslab, Faculteit der Geneeskunde, Universiteit Maastricht.
Prof. dr. C.P.M. van der Vleuten, psycholoog, hoogleraar Vakgroep Onderwijsontwikkeling en Onderwijsresearch, Universiteit Maastricht.

Correspondentieadres:

Dr. A.J.J.A. Schepbier, Skillslab, Universiteit Maastricht, Postbus 616, 6200 MD Maastricht.

bijvoorbeeld dat wat dokters zeggen te doen bij een bepaald probleem niet altijd overeenkomt met wat ze in het echt doen.⁹ Wat dokters echt doen, is uiteraard ook niet zonder meer het jekpunt. Binnen de beroepsgroep is op basis van dit soort gegevens overleg nodig om de juiste handelwijze teijken. Het is ons inziens van belang om het onderzoek naar de echte praktijk (hoe dokters echt werken, hoe dokters problemen oplossen, et cetera) te intensiveren en dan vooral ook de bevindingen te vertalen naar onderwijs- en toetsontwikkeling. Dit onderzoek is uiteraard ook van belang voor de kwaliteit van de zorg, de beroepsopleiding en de nascholing. Het ligt dan ook voor de hand dat deze belanghebbenden samenwerken.

Tot slot

In dit artikel worden slechts subjectieve indrukken beschreven. Gezien de omvang van het congres zou een volledige samenvatting vele nummers van dit tijdschrift vullen. Bovendien werkt de organisatie hard aan de procedure, zowel op cd-rom als in gedrukte vorm. Voor lezers die de proceedings willen bestellen, het adres van de organisatoren is: Ottawa Conference Office, National Board of Medical Examiners, 3750 Market Street, Philadelphia, PA 19104, USA, tel.: 2155909870, fax: 2155909755, e-mail: ottawa@mail.nbme.org.

Voor lezers die ook eens deze interessante conferentie willen bezoeken: de volgende Ottawa Conference is van 1-3 maart 2000 in Cape Town, Zuid-Afrika. Voor meer informatie kunt u contact opnemen met Sally Elliot, University of Cape Town Medical School, Anzio