

Het artsexamen: een theoretisch overzicht*

In deze bijdrage komt aan de orde het artsexamen in onderwijskundig perspectief. Dat wil zeggen: het artsexamen opgevat als een meet- en beoordelings-instrument. Dit is noodgedwongen een wat theoretische benadering - vooral ook omdat onderzoeksgegevens over dit onderwerp zo uitermate schaars zijn. Op de vraag waarom dat zo is, terwijl er bijvoorbeeld wel een overmaat aan literatuur, ook onderzoeksliteratuur bestaat, over kennismeting (multiple choice examens bijvoorbeeld), wordt niet ingegaan.

Het artsexamen bestaat meestal uit meerdere onderdelen. Doorgaans worden de co-assistentenschappen in de meeste klinische disciplines afgerond - al dan niet direct aansluitend - met een praktisch klinisch examen in het betreffende vakgebied. Het artsexamen is dan de combinatie van deze pke's. Er komen afwijkingen van het stramien voor (Maastricht kent een eigen evaluatie-systeem ook elders komen andere vormen dan het pke voor: min of meer theoretische examens, varianten van het GKE, computersimulaties e.d.). Deze worden hier buiten beschouwing gelaten. De aandacht is hier beperkt tot het traditionele praktisch klinisch examen. Men zou dit kunnen omschrijven als een mondeling examen, waarbij, naar aanleiding van het onderzoek van een patiënt, wordt beoordeeld of iemand in staat is dit onderzoek volledig en correct uit te voeren, in een schriftelijk verslag vast te leggen, waarbij tevens differentiaal-diagnostische mogelijkheden, voorstellen voor verder onderzoek en therapeutische en prognostische overwegingen aan de orde komen.

Wat valt op bij een nadere beschouwing van deze examenvorm?

1. Het is een mondeling examen

Deze examenvorm staat in onderwijskundige kring in een wat dubieus daglicht, zeker als het gaat om de beoordeling van andere dan mondelinge uitdrukkingsvaardigheden en zeker als de exameneisen niet gespecie-

ficeerd zijn en het examen slechts door één examinator wordt afgenomen.

5

Vaak wordt als argument voor deze wijze van examineren genoemd: de mogelijkheid tot het verkrijgen van een "algemene indruk", vergelijkbaar met de "klinische blik". Het is niet eenvoudig precies aan te geven wat daaronder verstaan moet worden. Vermoedelijk betreft dit kwaliteiten, zoals persoonlijk functioneren. Twee opmerkingen: 1) de Wet schrijft voor het artsexamen een onderzoek voor naar de bekwaamheid en niet naar de geschiktheid en 2) beoordeling van het functioneren kan beter over een wat langere periode ("continuous assessment") plaatsvinden en niet in één stressvolle momentopname.

2. Naar aanleiding van een patiëntenonderzoek

De dagelijkse praktijk is, dat het patiënten-onderzoek buiten de aanwezigheid van de examinator plaatsvindt. De bevindingen bij dit onderzoek worden in een schriftelijke status weergegeven en deze status vormt de inzet van het examen. Klinische vaardigheden worden dus hoogstens indirect beoordeeld en dan nog in termen van onderzoeksbevindingen. Uitkomsten van het onderzoek dus, en niet het onderzoek zelf. Vraagtekens moeten voorts gezet worden 1) bij de waarde van de patiëntenstatus als instrument voor het meten van medische competentie (een validiteitsprobleem) en 2) bij de waarde van de beoordeling van de status als een schriftelijk werkstuk (een betrouwbaarheidsprobleem).

3. Een derde opmerking betreft de onderwerpen, die in een dergelijk examen wel en niet aan de orde komen. De kans, dat het examen louter een theoretisch vragen antwoordspel wordt is groot. Dit gaat ten detrimente van het praktisch karakter van het examen.

De kernvraag luidt: welke eisen moeten aan het examen als onderwijskundig instrument gesteld worden? Met andere woorden:

1. wat wordt geëxamineerd?
2. hoe wordt geëxamineerd?

WAT

De wettelijke omschrijving voor het artsexamen luidt: Het artsexamen onderzoekt de bekwaamheid tot het uitoefenen van de geneeskunst". Sleutelwoord in deze beschrijving is: bekwaamheid. In het verlengde daarvan stelt zich de vraag naar het niveau, dat ten tijde van het basisartsexamen zou dienen te zijn bereikt.

Daarmee komen we op het terrein van de eindtermen van de basisartsopleiding. Deze eindtermen zijn doorgaans niet of gebrekkig geformuleerd, althans niet in specifieke en concrete bewoordingen. Men kan van mening verschillen over de vraag hoe erg dat is. Het ene uiterste is het "weg-bagatelliseren" van het probleem. Omschrijving van exameneisen hoeft helemaal niet. We nemen toch als sinds jaar en dag aan de lopende band examen af, door onze ervaring is een hoge mate van onuitgesproken consensus tussen de examinatoren ontstaan, e.d.

Het andere uiterste: het prediken van het heil van "Gegenstandskataloge": rigide, uiterst gedetailleerde opsommingen van exameneisen.

Component van bekwaamheid

<u>Perceptuele vaardigheden</u>	proces product
<u>Intellectuele vaardigheden</u>	proces product
<u>Motorische vaardigheden</u>	proces product
<u>Sociale vaardigheden</u>	proces product

Wat hier wordt bepleit is de erkenning van de legitimiteit van het stellen en beantwoord krijgen van enkele kritische vragen aan het artsexamen, dat toch ook het wolmerk is dat de faculteiten ten behoeve van de maatschappij aan hun afstuderende artsen meegeven.

Zo zou men willen weten:

Vraag 1: Hoe representatief zijn de examenopdrachten?

Vraag 2: Hoe systematisch vindt beoordeling plaats?

Vraag 3: Hoe eerlijk en billijk is het examen?

De vragen kunnen ook worden geformuleerd in de vorm van enkele onderwijskundige desiderata. Bij het praktisch klinisch examen zou men moeten beschikken over:

1. Een, tenminste operationeel gedefiniëerde, omschrijving van wat geëxamineerd wordt, bij voorkeur in termen van te beheersen vaardigheden.
2. Een omschrijving, die als richtsnoer dient bij een systematische samenstelling van examen-opdrachten, zodat er
3. een representatief en eerlijk examen ontstaat. Dat schone theorie en dagelijkse praktijk elkaar niet hoeven te bijten, toont het zgn. PIMS-model.

Examen-opdracht bijvoorbeeld

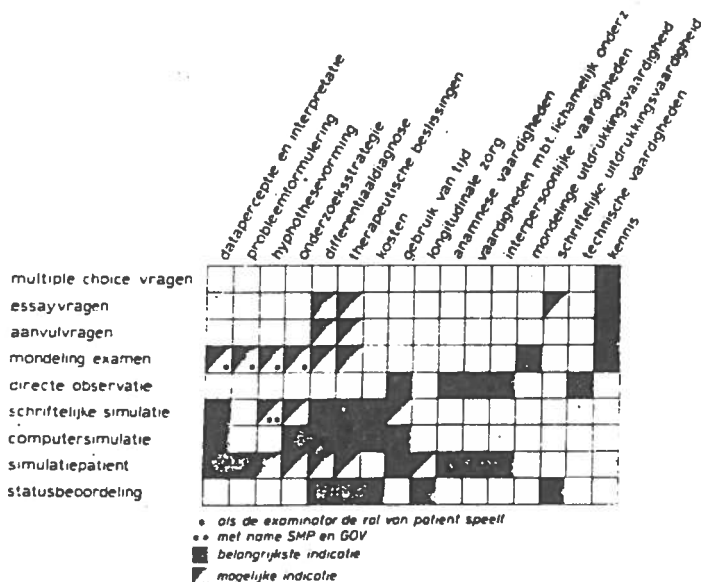
Analyseren van een soufflé.
Herkennen van een erythema nodosum.
Selecteren van samenhangende gegevens uit een ziektegeschiedenis.
Voorschrijven van een geneesmiddel bij een gegeven ziekte.
Palperen van de milt.
Aangeven van de plaats en de omvang van een demping.
Leefregel voor een patiënt toelichten. p.m. te meten aan satisfactie van een patiënt.

In de praktijk zal het eenvoudiger zijn een dergelijke serie examenopdrachten te realiseren in een systeem van "short cases", dan in ons systeem waarbij slechts sprake is van één "long case".

HOE

De zojuist besproken analytische benadering van het examenobject, maakt het ook mogelijk een gefundeerde keuze te maken uit het tegenwoordig beschikbare arsenaal aan meetmethoden.

Het diagram, ontleend aan Barrows, waar van links naar rechts de belangrijkste vaardigheden staan opgesomd en van boven naar beneden verschillende methoden. De bijbehorende, meest geëigende meetmethoden zijn in de zwarte cellen samengebracht.



Dit overzicht geeft allereerst een indruk van het genoemde arsenaal aan meetmethoden. Ook duidelijk wordt, dat bepaalde vaardigheden met een bepaalde methode niet adequaat gemeten kunnen worden; en omgekeerd, dat bepaalde methoden, bijvoorbeeld het mondeling examen, ongeschikt zijn voor het meten van bepaalde vaardigheden. Met andere woorden: competentie is een té complex en té heterogeen geheel om met één, al dan niet simpele methode gemeten te kunnen worden. Mogelijk een argument vóór meerdere "short-cases".

In ons examensysteem ligt de wijze van examineren, het HOE, vrijwel vast: een mondeling examen na een onderzoek. Het gevolg is, dat het WAT daarvan afhankelijk

is. De gekozen methode bepaalt in feite, ⁷ wat de mogelijkheden en beperkingen zijn van hetgeen in het examen aan de orde komt. Kenmerk van ons huidige systeem is, dat de vorm constant is en de inhoud variabel. De casuïstiek en daarmee de beoordeelde vaardigheden variëren per kandidaat. De les, die het scherm van Barrows ons leert, is de wenselijkheid van de omgekeerde situatie. Vastleggen van het WAT - zoals geïllustreerd is aan het PIMS-model - maakt het mogelijk de vorm te variëren en dus voor iedere component de meest geïndiceerde vorm te kies; en daarbinnen de meest valide en betrouwbare.

Validiteit: d.w.z.: wordt "bekwaamheid" gemeten? De lange traditie van onze wijze van examineren lijkt een argument voor een bevestigend antwoord op deze vraag. Deze wijze van examineren zou zich niet tot op de dag van vandaag hebben kunnen handhaven, als er niet tenminste sprake zou zijn van "face-validity" - validiteit op het le gezicht; niet ongeestig ook wel - "faith validity" genaamd.

Dit grapje leidt tot het voornaamste probleem: de validiteit van ons artsexamen is onderzoeksmatig nooit vastgesteld - in de huidige setting kan dat ook moeilijk. Op theoretische gronden, die in het voorafgaande aan de orde zijn geweest, moet deze validiteit echter serieus betwijfeld worden.

Betrouwbaarheid: hiervoor geldt mutatis mutandis hetzelfde. Een niet gestandaardiseerd mondeling examen, dat telkens door één examinerator wordt afgenomen is vrijwel ontoegankelijk voor betrouwbaarheidsonderzoek. Op theoretische gronden wil ik voorlopig als hypothese formuleren, dat het huidige artsexamen - in meettechnisch opzicht - onvoldoende betrouwbaar is, tot het tegendeel is vastgesteld. Zolang er sprake is van 3 variabelen: student, patiënt en examinerator, waarvan ik alleen geïnteresseerd ben in de eerstgenoemde, maar de beide andere niet onder controle heb, verwerp ik deze hypothese vooralsnog niet.

Het bovenstaande is nog eens samengevat in 5 discussiestellingen.

Discussiestellingen in de vorm van aanbevelingen

1. Sinds uit de wettelijke omschrijving voor het artsexamen de imponderabele term "geschiktheid" geschrapt is en uitsluitend een onderzoek wordt ingesteld naar de "bekwaamheid tot het uitoefenen van de geneeskunst", is de oude aanbeveling van onderwijskundigen om exameneisen naar inhoud en niveau te specificeren een stuk realistischer geworden en zou dus uitgevoerd kunnen worden.

2. De complexiteit van de term "bekwaamheid" maakt het noodzakelijk, dat bij de samenstelling van een examen waarin die gemeten wordt, systematisch te werk wordt gegaan en gestreefd wordt naar representativiteit. Dit wordt vergemakkelijkt door gebruik te maken van een analytische specificatietabel, waarin alle componenten van bekwaamheid staan opgenomen.

3. Als bij het huidige praktisch klinisch examen student A een 5 haalt en student B een 6, dan kan met voldoende zekerheid worden vastgesteld of B bekwaamer is dan A, dan wel A een moeilijker patiënt had dan B, dan wel B een mildere examinerator had dan A. Deze situatie kan o.a. verbeterd worden door meerdere studenten aan dezelfde examensituatie te onderwerpen en door meerdere, getrainde examineratoren te laten beoordelen.

4. In de huidige wijze van examineren is de methode constant en de inhoud variabel. Voordeel van een omkering daarvan is, behalve realisering van het in stelling 1 genoemde, dat voor ieder examenonderdeel de meest geëigende (valide, betrouwbare en hanteerbare) methode kan worden geselecteerd.

5. De complexiteit van de geëxamineerde materie rechtvaardigt een wijziging in het gebruik om de beoordelingsresultaten uit te drukken in één cijfer, dat theoretisch loopt van 1-10 en in de praktijk van 5-7. Dit geldt zowel voor het gebruik van cijfers ter evaluatie van studieresultaten als van onderwijsprogramma's.