

Interview met Prof. dr. Lambert (L.W.T.) Schuwirth

Examinering is voor mij een beslismoment. Toetsing is een middel om de ontwikkeling van studenten te stimuleren en studenten aan te raken

■ Harry Molkenboer

Ing. H.F.A.M. Molkenboer MEd is master toetsdeskundige en eigenaar van Bureau voor Toetsen & Beoordelen en is lid van de redactie van EXAMENS. E-mail: harry.molkenboer@toetsen-beoordelen.nl

■ Jan Jacobs

Ir. J.W.J. Jacobs is adviseur onderwijs en examinering binnen Yuverta en redactielid van EXAMENS. E-mail: j.jacobs@yuverta.nl

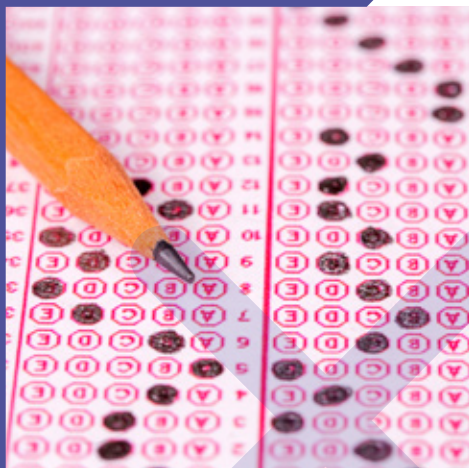
In de periode 2007-2012 zijn in Examens, tijdschrift voor de toetspraktijk twaalf portretten gepubliceerd van 'Kopstukken uit de Nederlandse examenwereld'. Het waren portretten van Nederlanders die een bijzondere plaats innemen in de wereld van toetsing en examinering. Nu ruim tien jaar verder is de redactie gestart met een nieuwe serie portretten van 'Kopstukken uit de Nederlandse toetswereld'. We hebben veel nieuwe Kopstukken!

In deze uitgave van Examens het interview met ons zesde kopstuk Prof. dr. Lambert (L.W.T.) Schuwirth. We spraken Lambert op 3 mei 2023 online; wij vanuit Nederland vanaf 09:00 uur en Lambert vanuit Adelaide in Australië waar het al 16:30 uur was. Afstand hoeft geen probleem meer te zijn.

In de voorbereiding op het interview was Lambert terughoudend om mee te werken aan zijn portret als Kopstuk binnen de Nederlandse toetswereld. Hij vindt zich geen Kopstuk. Verschillende andere examendeskundigen hier in Nederland dachten daar anders over en noemden hem nadrukkelijk wel een kopstuk. Het overzicht van zijn 219 wetenschappelijke publicaties (Elsevier Scopus, geraadpleegd op 3 mei 2023) laat zien dat Lambert veel betekent op het gebied van assessments. Zijn



Lambert Schuwirth tijdens het interview op 3 mei 2023 met op de achtergrond zijn achtertuin.



Zondermeer kun je je afvragen of cijfermatig toetsen beter is dan narratief.

bescheidenheid blijkt bij hem te passen, zo bemerken we tijdens het interview.

Het artikel 'A model for programmatic assessment fit for purpose' (Van der Vleuten et al., 2012) heeft de basis gelegd voor waar in Nederland momenteel veel aandacht voor is, namelijk programmatisch toetsen. In Examen-2011-3 hebben Cees van der Vleuten en Lambert hierover al een artikel geschreven met als titel 'Een model voor programmatisch toetsen, een toetsprogramma van de Universiteit Maastricht.'

Daarnaast schreef Lambert in Examen-2014-4 een gastcolumn met als titel 'Examineren in spagaat.' Hierin beschrijft hij de student als de min of meer commerciële klant van de onderwijsinstelling die betaalt voor het diploma (in Australië betalen studenten nog meer om onderwijs te mogen volgen dan in

Nederland) met als gevolg dat de student dus sterk diplomagericht studeert. Hij plaatst dit in die column tegenover wat hij noemt onze socratische traditie van leren door middel van een kritische dialoog van de student met zijn docent en het gebruiken van zelfreflectie en zelfontwikkeling.

Inmiddels zal duidelijk zijn waar de toets-technische drive, of moeten we zeggen de ontwikkelingsgerichte drive voor studenten, bij Lambert ligt.

Wie ben je en wat is je specialisme?

"Kort samengevat ben ik onderzoeker en dat vind ik ongelooflijk leuk. Bij mijn werkwijze als onderzoeker heeft het voordelen dat ik geneeskunde heb gestudeerd, omdat geneeskunde qua onderzoek meer een kwantitatieve, logisch-positivistische kant heeft, terwijl een clinicus eigenlijk meer kwalitatief, subjectivistisch of constructivistisch bezig is. Die verschillende perspectieven zijn ook in ons onderzoek belangrijk. Daarbij houd ik mij inmiddels dertig jaar met onderwijskunde en met name met toetsing bezig. Ik bevind mij eigenlijk precies op dat snijvlak. Mijn specialisme, en dus veel van mijn onderzoeken, richt zich op het zoeken naar onderwijskundige en toetstechnische verbeteringen. Dat gaat gepaard met innovatie. Die innovatie is voor mij geen doel op zich. Verbeteringen moeten waardevol zijn. Ze moeten aansluiten bij de actualiteit om waardevol te kunnen zijn. Ik zal jullie uitleggen hoe ik via geneeskunde in de medische onderwijskunde en de toetsing ben beland en nog steeds met geneeskunde te maken heb (zie ook het cv in Kader 1).

Van de studie geneeskunde naar onderwijskunde

Ik ben geneeskunde gaan studeren omdat ik dat interessant vond. Vooral hoe het brein en het geheugen werken boeide mij, en hoe ik mijn belangstelling voor psychologie en geneeskunde zou kunnen combineren. Neurochirurg worden leek mij wel wat. Tijdens

het laatste deel van mijn studie richtte ik mij als anios (arts niet in opleiding tot specialist, red.) op de psychiatrie, maar dat was toch niet wat ik wilde en ben ik wat anders gaan doen. Er deed zich aan de Universiteit Maastricht een vacature voor bij 'Onderwijsontwikkeling en Research' van de universiteit. Vanaf de eerste dag voelde ik mij daar als een vis in het water en kreeg ik alle mogelijkheden om me te ontwikkelen. Achteraf gezien was dat de beste beslissing die ik ooit genomen heb, trouwens net zoals het verhuizen naar Australië.

Promotie bij Cees van der Vleuten

Ik ging in Maastricht samenwerken met Cees van der Vleuten. Hij was mijn baas en mentor en werd mijn vriend. Wat ik zo bewonder aan Cees waren zijn oorspronkelijke gedachten

en dat hij altijd bezig is anderen te stimuleren. Het ging nooit over hemzelf, altijd over anderen. Ik deed een vrij simpel klusje en ik zei tegen Cees dat ik mij verveelde. Cees opperde om mij te verdiepen in 'Computergestuurd casusgericht toetsing binnen geneeskunde'. Hij zei: 'Dat is een grote klus en daar ga je je in verliezen'. Het betrof toetsing met vele casussen over veel verschillende onderwerpen met gesloten en open vragen. Het leidde tot mijn promotieonderzoek over onder meer het cueing-effect (het herkennen van de goede antwoorden in toetsen met gesloten vragen). Het bleek dat cueing leidt tot hogere scores in vergelijking met open vragen. Er is dus sprake van herkenning. Maar er is ook sprake van hoge correlaties tussen de scores op de gesloten en de open vragen. Als je hetzelfde

Werkzaamheden en opleiding

Werkzaamheden

Juni 2017 – heden:

- Schuwirth Consulting (Australië)

Oktober 2015 – heden:

- Professor of medicine (education) Uniformed Services University of the Health Sciences Bethesda Maryland, USA

Oktober 2014 – heden:

- Distinguished professor of Medical Education Chang Gung University, Taiwan

Augustus 2011 – heden:

- Professor of medical education / Chair Prideaux Discipline of Clinical Education, Flinders University, Adelaide
- Director Flinders University Prideaux Centre for Research in Health Professions Education. Conduct research, support researchers, management

Maart 2005 - September 2019:

- Professor for innovative assessment, Maastricht University

Opleiding

1998:

- Phd: An approach to the assessment of medical problem solving: Computerised Case-based Testing

1979 – 1988:

- University Maastricht; grade: Medical Doctorate

bevraagt, maakt het niet uit of je open of gesloten vragen stelt. Maar dat wil niet zeggen dat iedere vraagvorm bij iedere vraaginhoud past. Dus het bleek beter om eerst de essentie per onderwerp vast te stellen en dan pas de vraagvorm te kiezen. Het heeft te maken met de stimulus van de vraag, meer dan met de responsvorm.

Daarnaast deed ik onderzoek naar verschillen in probleemoplossen bij ervaren huisartsen en medisch studenten, waarbij ik uitging van de scripttheorie (experts versus non-experts). Hierbij vergeleek ik casustoetsen met feitenkennisvragen. Het bleek dat de proefpersonen (experts nog meer als non-experts) probabilistisch te werk gingen bij de casusvragen. Ze deden een weging van alle aspecten uit de casus. Bij de feitenkennisvragen gebeurde dit nauwelijks en werden de feiten simpelweg aan elkaar geregen. Ook bleek uit dit onderzoek dat de experts de casussen anders lazen dan non-experts. Non-experts hebben nog onvolgende scripts ontwikkeld en lezen de casus sequentieel, maar experts hebben een script en gaan op basis daarvan door scripts heen. Experts scoorden goed, non-experts scoorden minder goed en hadden meer woorden nodig om te beschrijven hoe ze tot de diagnose kwamen. Totaal heb ik elf van dit soort onderzoeken gedaan. Cees van der Vleuten was hierbij mijn promotor.

Mede grondlegger van de interuniversitaire voortgangstoets en programmatisch toetsen

Na mijn promotie werd ik voorzitter van de toetsgroep binnen de afdeling Onderwijsontwikkeling en Research van de Universiteit Maastricht. Samen met onder andere Cees hebben wij de Interuniversitaire Voortgangs Toets Geneeskunde (IVTG) opgezet. De voortgangstoets bestond al langer, als geesteskind van Wynand Wijnen, maar werd lange tijd alleen in Maastricht gebruikt. Voortgangstoetsing is een instrument om de kennisprogressie van studenten Geneeskunde gedurende hun studie te meten. Deze voortgangstoets wordt

nu door alle geneeskunde-universiteiten in Nederland viermaal per jaar afgenomen bij ongeveer 12.000 studenten.

Verder hebben we met een groot team een vierjarige Geneeskundeopleiding (plus een opleiding tot Klinisch onderzoeker) opgezet voor studenten die een Bachelor hadden afgerond binnen het gezondheidsdomein. Voor deze opleiding hebben we een toetsprogramma ontwikkeld op basis van de concepten van programmatisch toetsen (zie het eerder genoemde artikel van Van der Vleuten et al., 2012). Prof. dr. Sylvia Heeneman speelde hierbij een grote rol, zoals ze ook een grote rol speelt bij de operationalisatie van programmatisch toetsen binnen Geneeskunde Maastricht.

Programmatische toetsen uitgediept

Ook elders in de wereld en in Amerika (Cleveland Ohio) in het bijzonder werd rond dezelfde periode (2012) gewerkt aan programmatisch toetsen. Dat was ongeveer volgens dezelfde principes opgezet zoals wij die zagen. Het bevestigde ons dat we met programmatisch toetsen op de goede weg zaten. We hebben toen geprobeerd om de nieuwe geneeskundeopleiding (zie hiervoor, red.) geheel op basis van de principes van programmatisch toetsen te ontwikkelen, wat ons goed lukte. De studenten van de nieuwe opleiding hadden al een voorsprong door de al aanwezige kennis op ongeveer het niveau van het tweede jaar geneeskunde. Na de opleiding bleek dat door het programmatisch toetsen de studenten hogere bekwaamheden hadden dan bij de traditioneel opgeleide geneeskundestudenten. Voor programmatisch toetsen is het essentieel dat informatie over de student van verschillende bronnen wordt gecombineerd en op een narratieve wijze wordt ontsloten. Dat kan omdat de bronnen inhoudelijk bij elkaar passen. Bij traditioneel toetsen tellen we cijfers op en nemen we daar het gemiddelde van (de stationstoets¹ is bijvoorbeeld zo opgezet). Dat is vreemd als je het nader beschouwt. Zondermeer kun je je afvragen of cijfermatig toetsen

beter is dan narratief. Toetsen is altijd een evaluatieve actie; het leidt tot een waarde. 'Goed' is bijvoorbeeld beter dan 'slecht'. Dat kun je ook weergeven als '8' en '4'. Het is prima om dit als ordinale variabelen te zien, maar als je gaat middelen, maak je het toch ineens een intervalvariabele. Of anders gezegd, je middelt appels met peren. In de patiëntenzorg zeg je ook niet tegen een patiënt: "uw bloedsuiker is te hoog maar uw natriumgehalte is te laag, dus gemiddeld gaat het wel goed met de labwaarden". Het narratieve verhaal is: 'moeheid, dorst, veel plassen, slecht genezende wondjes, slechte pulsaties in de kleine slagaders van de voet', met als conclusie suikerziekte. Daarbij combineren we verbale, visuele, tactiele en numerieke informatie. Dat is een essentieel principe van programmatisch toetsen, het als het ware diagnosticeren van een competentie. Daarom hechten we binnen programmatische toetsen veel waarde aan de longitudinale narratieve onderbouwing van informatie uit verschillende bronnen en hechten we juist niet alleen maar aan cijfers. Een eindbeslissing mag en kan hierbij voor de student nooit een verrassing zijn.

Binnen geneeskunde is dit door de aard van de opleiding goed te doen. Daar sluit alles wat wij de studenten leren bij elkaar aan rondom een patiënt. Er is integratie van kennis, inzichten en vaardigheden rond de patiënt. Het is dus logisch dat je niet op elk moment een beslissing neemt met een toets en een cijfer, maar dat je continu evalueert; meetmomenten zijn niet altijd automatisch ook beslismomenten. De beslissing wordt over een geheel genomen. Je kunt het misschien vergelijken met een opleiding zoals engineering, waarbij een oplossing voor een probleem moet komen in samenhang met de opdrachtgever/klant.

Met de voortgangstoets toetsen we alle studenten geneeskunde viermaal per jaar op

eindniveau van de opleiding over ongeveer vijftien vaste geneeskundige categorieën. Het doel van voortgangstoetsing is het stimuleren van continu leren, in plaats van tentamengericht leren. De student krijgt geen cijfer, maar kan aan de hand van grafieken en scores zien op welke categorieën de student goed scoorde en op welke minder, en hoe hij scoorde ten opzichte van zijn jaargenoten. De student wordt dus gefaciliteerd met relevante informatie gericht op zijn ontwikkeling. Er wordt gedurende de opleiding geen beslissing genomen alleen op basis van de voortgangstoets. Maar aan het einde van de opleiding verwachten we wel dat de student zich heeft ontwikkeld op het niveau dat van een beginnend arts verwacht mag worden. We kunnen ook per toets terugzien hoe de student scoorde en hoe hij heeft gewerkt aan zijn ontwikkeling. Deze data levert veel informatie op over de student. Beslissingen passen meer binnen het leven-lang-leren-concept, dan binnen een opleidingsjaar.

Zelfverantwoordelijkheid van de student en hem hierbij helpen

Met programmatisch toetsen bieden we de student veel autonomie. De student is zelf verantwoordelijk om de aangereikte informatie te gebruiken voor zijn eigen ontwikkeling. Niet alle studenten beschikken vanaf het begin van de opleiding in voldoende mate over zelfverantwoordelijkheid. Dit moet toegevoegd worden aan de student. Dit doen wij op verschillende manieren, zo moet de student bij een onvoldoende gemaakte opdracht zelf een remediëring ontwikkelen; hij is immers in eerste instantie probleemeigenaar. De student moet zich afvragen waar de slechte score aan lag en voor een oplossing zorgen. Een docent kan een student bij dit denkproces helpen en de student activeren. Dit hoort ook bij het probleemoplossend kijken van de docent. Dus meer van voor de klas naar in de klas.

^{1*} Een stationstoets is een praktijkgerichte toets die bestaat uit verschillende stations. Een station bestaat meestal uit een patiënt met een probleem. De student moet kort onderzoek naar het probleem doen en tot een onderbouwde diagnose en een behandelvoorstel komen. Zo kunnen in bijvoorbeeld twee uur 12 stations worden doorlopen, waarna per station tot een beoordeling van de bekwaamheid van de student wordt gekomen en uiteindelijk tot een eindoordeel.

Competentie, zelf ontwikkelen, zelf sturen, zelf leren oplossen, zelf verantwoording nemen staat haaks op de toetstechnische testtheorie die uitgaat van ranking of onderscheid maken tussen de getoetste studenten. We vinden nu dat toetsing moet leiden tot effectief leren, met name gericht op langdurige retentie (vasthouden van kennis en inzichten, red.) en ontwikkeling van competentie in plaats van op een eenmalige prestatie op een toetsmoment. Dit is volgens mij de grootste uitdaging voor het komende decennium: toetsing weghalen van het idee van een 'objectieve' cijfermatige meting. We moeten ons meer richten op het stimuleren van langdurige retentie. Als je continu leert in plaats van op het laatste moment blokken voor een toets en als je kennis steeds activeert, wordt de retrieval (zelf kennis/informatie oproepen of proberen terug te halen uit het geheugen, red.) steeds hoger. Dit moet je dus zeker ook gebruiken. Assessment for learning is niet gericht op het behalen van een vak en vervolgens snel de inhoud van het vak vergeten.

Programmatische toetsing kent geen herkansingen, maar continue remediëring. Achterstanden moeten opgelost worden, maar de student kan wel door. Het is ook belangrijk bewust te zijn van complexiteit van processen. Dat is moeilijk in een strak management keurslijf te plaatsen met gestandaardiseerde toetsen. Dit probleem staat bekend als 'McNamara's fallacy'. Hij was de minister van defensie in de VS ten tijde van de Vietnamoorlog, maar daarvoor een succesvolle manager in de auto-industrie (wat meer voorspelbare processen bevat) en managede de oorlog op dezelfde manier (nu zouden we dat spreadsheetmanagement noemen). Dat ging niet goed, omdat oorlog een meer complex proces is (we zien het nu dagelijks in het nieuws). Onderwijs en geneeskunde zijn ook complexe processen.

Een ander probleem is dat als je een reductionistische maat² gebruikt (bijvoorbeeld

een cijfer), je al gauw het signaal afgeeft dat wat niet meegenomen wordt in het cijfer niet belangrijk is. Ook zullen studenten soms proberen hoge cijfers te halen op andere wijzen dan door gewoon goed te leren. Dit zijn twee voorbeelden van risico's als je kwaliteit kwantificeert. Ik ben niet tegen getallen, wel tegen het misbruiken van cijfers om kwaliteit te beschrijven. Ik noem in dezen Goodhart's law: 'When a measure becomes a target, it ceases to be a good measure'.

Als je minder cijfermatig kijkt, maar meer 'diagnostisch', kun je ook proberen te begrijpen waarom een student het niet goed doet. Zo kun je met de student het gesprek aan gaan over slecht leren, zijn familie of een gezondheidsprobleem en kunnen docenten ook meer probleemoplossend gaan kijken, wat leidt tot een heel andere wijze van denken over de rol van toetsing. Ik geef als voorbeeld de voortgangstoets hier in Flinders (Universiteit in Adelaide) waarbij studenten in het voorlaatste jaar soms nog problemen hebben. Door nadrukkelijk verbinding te realiseren, door als docent vanuit de onderwijsorganisatie zorgplicht te erkennen en hulpvaardig contact te leggen, komt uit die coaching vaak het echte studieprobleem van de student voor de dag. Dan reageren studenten toch nog verbaasd dat ze geholpen worden. Dus maak toetsing niet puur selecterend, raak de studenten aan. Studenten zijn bijvoorbeeld onhandig met hun IT gebruik voor betekenisvol leren (ze kunnen bijvoorbeeld slecht informatie en bronnen op het internet vinden) en worden vaak verleid tot hulpjes bij korte-termijn leren (althans dat vinden we hier in Australië), ze organiseren hun leren en kennis soms inefficiënt en vinden het dan moeilijk om just-in-time informatie te vinden – iets dat ze als arts vaak zullen moeten kunnen. Hiermee kunnen wij de student helpen. Wat de student zelf moet doen is zich goed voorbereiden, op het juiste niveau willen presteren, werken aan motivatie en hard wil-

²opvatting in de wetenschapsfilosofie die stelt dat de natuur van complexe entiteiten steeds herleid kan worden tot meer fundamentele entiteiten (bron: Algemeen Nederlands Woordenboek)

len werken. Uiteindelijk willen de meeste studenten toch de best-mogelijke dokter worden, merken we. Na een hulpgesprek komt een student zelden terug; toetsen als aanraken.”

Hoe kwam je in Australië terecht?

“Soms bestaat er geen toeval. Er kwam een verzoek van Flinders University in Adelaide om daar een onderzoeksgroep op te zetten. Tegelijkertijd was ik ook aan een andere uitdaging toe. Onderwijs en geneeskunde staan in Australië op een redelijk hoog niveau, hoewel ik denk dat medisch onderwijskundig onderzoek en ontwikkeling in Nederland nog vaak voor ligt. Maar, Australië biedt veel ruimte en levenskwaliteit.”

Dit laatste zagen we, de interviewers, toen we vroegen naar de achtergrond van het beeld tijdens het Teams-interview. Dat bleek de achtertuin te zijn van de woning van de familie Schuwirth. En aan de voorzijde heeft Lambert ook nog eens uitzicht op zee!

Lambert is samen met zijn vrouw geëmigreerd naar Australië en heeft het daar enorm naar zijn zin. Ruimte, levenskwaliteit en plezier in en met zijn werk zijn de basis.

Hoe zouden we jouw brede expertise op toetsgebied mogen samenvatten?

“Het zal duidelijk zijn dat mijn studie geneeskunde, mijn promotie met als titel ‘An approach to the assessment of medical problem solving: Computerised Case-based Testing’ (1998) en mijn werk bij ‘Onderwijsontwikkeling en Research’ mij veel expertise heeft gebracht op het kruisvlak van geneeskunde, onderwijskunde en toetsing. Mijn expertise varieert hierbij van puur kwantitatief onderzoek (onderzoek met veel data-analyse) tot puur kwalitatief onderzoek (onderwijskundige concepten). Dit betreft bijvoorbeeld redeneren vanuit de casus van de patiënt en de rol van de arts daarbinnen en hoe kunnen we de arts daar optimaal in ontwikkelen? Wat heeft de lerende nodig om de complexe praktijk aan te

kunnen en hoe word je een professional? Wat draag ik via mijn vak bij aan de lerende, zodat de student zo competent mogelijk wordt? Dat soort concrete en praktijkgerelateerde vraagstukken drijven mij.”

Op wie of waarop richt je je vooral?

“Op de lerende. Als de praktijk complex en onvoorspelbaar is, is het niet te voorspellen wat het effect van een behaalde toets is. Wat je moet doen is kennis aanreiken en de student leren om een repertoire van gereedschappen te gebruiken om tot passende oplossingen te komen. De student grenzen leren kennen tussen acceptabele en niet-acceptabele strategieën of oplossingen.”

Wat versta je onder toetsing en/of examinering?

“Examinering is voor mij een beslismoment. Het resultaat ervan wordt in een geregisseerde en gereduceerde vorm aangereikt, zoals een cijfer. Toetsing regisseer ik niet tot een cijfer, maar tot een middel om de ontwikkeling van studenten te stimuleren en studenten aan te raken. Het moet leiden tot rijke informatie over een student in de vorm van een narratief om uiteindelijk als onderwijs- en examencommissie goede beslissingen te kunnen nemen.”

Hoe en wanneer ben je geraakt door het onderwerp toetsing?

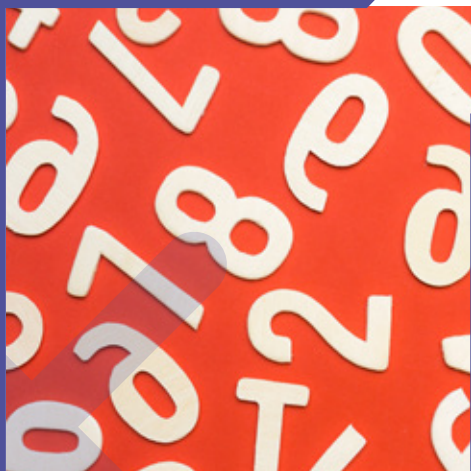
“Meteen aan het begin toen ik via mijn werk bij de Universiteit Maastricht bij ‘Onderwijsontwikkeling en Research’ in het onderwijs terecht kwam. Toetsing is voor mij veruit het leukste onderwerp. Leren en toetsen zijn twee kanten van dezelfde medaille. Daarnaast ben ik altijd geïnteresseerd geweest in probleemoplossing. En binnen toetsing bestaan uitdagingen genoeg.”

Wie zijn jouw leermeesters geweest?

“Cees, Cees en Cees (van der Vleuten, red.). Ook Wynand Wijnen en Henk Schmidt (cognitief psycholoog, destijds vakgroepvoorzitter van ‘Onderwijsontwikkeling en Research’). De

overeenkomst tussen deze drie mensen zit in hun karakter. Deze professionals zaten er niet voor zichzelf, zij gaven juist anderen een podium, verbonden anderen en lieten anderen het beste uit zichzelf halen. Ik probeer dit ook door te geven.

Wat mij nog bijstaat van Wynand is dat hij een presentatie moest houden op een congres over ontwikkelingen binnen het medisch onderwijs. Wellicht door miscommunicatie of misschien een verkeerde programmering deed hij dat zonder voorbereiding. De wijze waarop Wynand dat deed was onovertroffen. Hij kreeg de hele zaal op zijn hand, onnavolgbaar was dat. Complexe onderwerpen kon hij heel helder en begrijpelijk neerzetten, en dat zo voor de vuist weg.”



**Ik ben niet tegen getallen,
wel tegen het misbruiken
van cijfers om kwaliteit
te beschrijven.**

Wie beschouw je als vooraanstaand in de toetswereld?

“Cees van der Vleuten vanzelfsprekend en ook David Boud. David Boud was in de jaren negentig al bezig met toetsen voor leren, en heeft een heel oorspronkelijke kijk op toetsing. Teveel vooraanstaande personen om op te noemen eigenlijk, vooral experts uit Australië, de VS en Canada, die ook in het medisch onderwijs zitten. Deze personen zijn in Nederland minder bekend.”

Wat drijft je?

“Pure lol en passie in mijn gezin, binnen het onderwijs en op het gebied van toetsinnovatie, zonder daarbij innovatie als doel te hebben. Wat ik ook doe, ik doe het met passie, snel en explosief. Mijn prioriteit ligt bij mijn gezin, als vader wil ik niet vervangbaar zijn. Mijn kinderen kunnen geen vervangende vader kiezen. Voor mijn werkgever is het mogelijk binnen zes maanden een vervanger voor mij te vinden en aan te stellen.”

Wat voor visie heb je op toetsing en hoe ziet de toekomst van toetsing er in jouw ogen uit?

“Bij toetsing in de toekomst moeten we ons realiseren dat studenten dan veel meer dingen tot hun beschikking hebben, dan dat ze nu gebruiken. Ze zijn meer verbonden met de maatschappij dan wij waren. Ze zijn lid van verschillende community's. Ze kunnen veel breder en makkelijker communiceren dan wij konden via vele verschillende media. Ze maken ook deel uit van verschillende kennisnetwerken. Ze kunnen op veel meer verschillende wijzen iets creëren (video, audio, powerpoints, prezi's, enz.) en iets ontwikkelen, zijn beter in staat om een boodschap over te brengen en dit zit allemaal in één apparaat: hun telefoon of tablet. En dat gebruiken ze.

Studenten worden partner in hun eigen toetsing, moeten snel informatie kunnen vinden en op waarde schatten. Dat vereist veel kennis in hun 'biologische' geheugen maar die zal anders georganiseerd moeten zijn dan toen ik

studeerde. Studenten kunnen nog meer dan nu verbonden zijn met de wereld en we zullen studenten minder als individu moeten toetsen maar meer als deel van een netwerk. Toetsing zal minder puur op kennis gericht zijn en meer gericht zijn op hoe de student zijn kennis, zijn netwerk en alle andere voordelen mobiliseert om zijn – toekomstig – werk te doen. Als we onze studenten daarin niet faciliteren/trainen, missen ze de aansluiting, of zoals ze hier zeggen 'If you are not at the table, you will end up on the menu'. Als we als onderwijsorganisatie zelf niet nieuwe ontwikkelingen oppakken, dan pakken andere organisaties het op. Commerciële organisaties zullen dan in de toekomst onderwijs over gaan nemen. En denk aan micro credentials die buiten het onderwijs om, waarde gaan krijgen. Nederland loopt hier mijns inziens in voorop en gaat hierin heel hard.

Artsen moeten zich blijven ontwikkelen omdat de geneeskunde zich ontwikkelt, de moderne patiënt heeft zich vaak beter voorbereid en verwacht dat zijn ervarings-expertise als patiënt serieus genomen wordt. De dialoog tussen patiënt en arts verandert sterk en de arts werkt al lang niet meer op de oude wijze als autoriteit. In toenemende mate wordt van de arts verwacht dat hij actuele problemen in de actuele setting op moet lossen. Ook daarbij zie je in sommige landen al dat de arts veel meer een leider (of volger) wordt in een samenwerkingsverband, dan de individuele expert.

Waar ben je het meeste trots op?

"Ik doe niet aan trots. Ik kom uit een normaal middenklasse gezin, ik kon een studiebeurs krijgen, en had geluk dat ik werd ingeloot. Ik ben blij en dankbaar dat ik geneeskunde heb kunnen studeren en ik heb geboft met mijn baas en promotor en vele anderen dichtbij mij. Ik probeer dingen zo goed mogelijk te doen. Ik ben dus heel blij met hoe dingen gelopen zijn en over het algemeen vind ik mijn werk leuk en heb ik een harmonisch gezin. Daar heb ik

zelf niets aan gedaan. Mijn levensles is dat ik dankbaar ben dat ik nu hier ben en zit."

Wat zou je anders hebben gedaan?

"Niets. Ik heb tijdens mijn opleiding geneeskunde niet zo hard gelopen als ik moest lopen. Ik heb daarna in eerste instantie gekozen voor psychiatrie. Toen heb ik keuzes gemaakt met de informatie van toen. Door voortschrijdend inzicht kwam ik tot andere keuzes. Ik ben ook heel duidelijk de eerste 10 tot 15 jaar voor mezelf en mijn eigen ontwikkeling gegaan en heb toen ook mijn netwerk ontwikkeld. Maar dan heb je wel genoeg voor je eigen ego gedaan. Dan komt het neer op het delen van je kennis, inzichten en ervaringen en kun je voor anderen wat gaan betekenen."

Welke boodschap heb je voor de lezers van Examens?

"Wees trots op het Nederlandse onderwijs en de Nederlandse onderwijscultuur. En dit geldt ook voor het medisch onderwijs. Het is niet toevallig dat er zoveel gebeurt in Nederland. De Nederlandse volksaad is van proberen en ervoor gaan. Met name bij innovatieve ontwikkelingen. Met onderwijsontwikkeling en -onderzoek neemt Nederland in de wereld een vooraanstaande plaats in. Helaas zie ik veel kritische reacties op collegae, vliegen afvangen, doe dat niet, waardeer het talent en de kennis van anderen."

Referentie

- Van der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. J., & van Tartwijk, J. (2012). A model for programmatic assessment fit for purpose. *Medical Teacher*, 34, 205-214. <https://doi.org/10.3109/0142159X.2012.652239>

Programmatisch toetsen: de touchstones en de goudstukken

Touchstone

Een touchstone is een zwarte steen (lydiet) die vroeger door de goudsmid werd gebruikt als toetssteen door een streep van een gevalideerd goudstuk op de steen te trekken. Door vergelijking van de spiegeling van een nieuwe streep van een aangeboden goudstuk met de gevalideerde norm (de gevalideerde streep) op de touchstone kon het goudgehalte van het aangeboden goudstuk worden geschat. Tegenwoordig wordt het begrip ‘touchstone’ gebruikt als maatstaf of criterium waarop zaken worden beoordeeld.

Programmatisch toetsen

Aan Lambert Schuwirth werd de vraag voorgelegd of het begrip ‘toetsen’ binnen ‘programmatisch toetsen’ niet een verkeerde connotatie heeft. Toetsen wordt in Nederland veelal geassocieerd met een tentamen/examen met een beslissing zoals behaald of niet behaald. Terwijl programmatisch toetsen vooral gericht is op het aantonen van een ontwikkeling die een student doormaakt aan de hand van bewijzen die de student vermeldt, datapunten genaamd. Deze datapunten worden niet beoordeeld met behaald/niet behaald, ze geven als geheel en in relatie met elkaar wel een indruk over de ontwikkeling van de student. Als binnen een leerlijn een datapunt nog niet ‘aan de maat’ is, kan de student op een later moment of zelfs met een andere datapunt of combinatie van datapunten aantonen dat hij zich wel ontwikkelt richting de bekwaamheid die verwacht wordt. Het is het totaalbeeld aan datapunten dat inzicht moet geven in de bekwaamheid van de student op basis waarvan uiteindelijk (na bijvoorbeeld een langere onderwijsperiode van een semester of een heel onderwijsjaar of opleiding) een high-stake-beslissing (behaald/niet behaald) kan worden genomen. Hiervoor zijn doordachte en voldoende datapunten noodzakelijk (aggregatie) samengebracht in een programma.

Er wordt aan studenten veel autonomie geboden om te werken aan bewijzen. Concrete, eenduidige en beoordeelbare leerdoelen geven de studenten (én de docenten) handvaten om hun ontwikkeling inhoud te geven en zichzelf te beoordelen of hun bewijzen van ontwikkeling ‘aan de maat’ zijn. Voor docenten wordt de rol van coach, begeleider en/of mentor steeds belangrijker naast die van de docent die doceert. Het begeleiden van de studenten betekent een belangrijke feedback en feed-forward-gevende taak op basis van de leerdoelen die sturing daarvoor geven. Studenten en docenten moeten loskomen van het idee van toetsen met cijfers, die tot nu toe vaak de ontwikkeling van de student een maat gaven. Toetsen kunnen nog steeds worden gebruikt, zelfs met een cijfer maar dan vooral met de functie voor het leren, en zonder dat het cijfer de beslissing van behaald/niet behaald produceert, maar meer om geïntegreerd te worden met andere informatie (zoals in de klinische praktijk ook cijfermatige labwaarden worden geïntegreerd met kwalitatieve informatie uit de anamnese en het lichamelijk onderzoek). De toetsresultaten alsmede de feedback die de docenten en peers de student geven, worden afgezet tegen de leerdoelen als norm, de eerste goudstreep op de touchstone. Aangegeven wordt hoever de student nog af is van, of al verder is dan de goudstreep die de norm is; dus van de leerdoelen.

Touchstone, touch en aanraken

Lambert gaf aan dat in het woord 'touchstone' het woord 'touch' zit dat aanraken betekent. En dat is waar het bij programmatisch toetsen ook om gaat; het virtueel aanraken van de student, de student begeleiden, ondersteunen, ernaast staan, feedback en feedforward geven, er dus zijn voor de student; die 'touch' geven. Hij zei hierover letterlijk: 'In onderwijs kun je niet succesvol zijn als je de student niet aanraakt, en hierbij is sprake van tweezijdigheid.'

Wij als interviewers vonden 'Touchstone' een schitterend begrip en een sterke metafoor. Zou het een idee zijn om datapunten om te dopen in goudstukken? Dat klinkt toch veel sympathieker. De glans van een datapunt wordt vergeleken met de spiegeling van de leerdoelen op de touchstone.

Harry en Jan

GeSignaleerd

Fraude of duwtje in de rug? ChatGPT houdt gemoederen in hoger onderwijs flink bezig

9 juli 2023, 08.00 uur · Aangepast 9 juli 2023, 16.22 uur · Door Trix Pieterse & Lotte van den Berg

Examencommissies en docenten op het HBO en de universiteit hebben hun handen vol aan de gevolgen van ChatGPT. Het is voor docenten steeds lastiger om te bepalen of studenten hun teksten zelf hebben geschreven en de grens tussen hulp bij het schrijven en het plegen van plagiaat lijkt te vervagen. Desondanks heeft ChatGPT ook veel potentie, zeggen docenten.

"Sommigen collega's vinden dat we ChatGPT moeten omarmen, anderen zeggen dat we het einde van de wereld hebben bereikt", zegt Talita Ferrantelli, docent bij Politics, Psychology, Law and Economic (PPLE) aan de Universiteit van Amsterdam.

Afgelopen jaar voerden Ferrantelli en haar collega's meerdere discussies over gebruik van het computerprogramma, maar tot consensus kwam het nog niet. "Er is nu wel een protocol, maar iedere docent beslist zelf hoe ze ermee omgaan."

ChatGPT, een computerprogramma dat op aanvraag coherente teksten kan schrijven, is populair onder studenten. Het programma heeft antwoorden op alle vragen, kan feedback geven en schrijft beter dan de gemiddelde student. ChatGPT is als een slimme klasgenoot, maar dan een die altijd klaarstaat om je te helpen.

Voor docenten en examencommissies betekent ChatGPT dan ook een radicale verandering. Niet langer kunnen ze ervan uitgaan dat studenten hun eigen teksten hebben geschreven. Want wat als de tekst briljanter is dan de student zelf? En hoe controleer je nog op plagiaat?

Lees verder op <https://www.nhnieuws.nl/nieuws/321430/fraude-of-duwtje-in-de-rug-chat-gpt-houdt-gemoederen-in-hoger-onderwijs-flink-bezig>