

Meervoudige **meerkeuzevragen** tegen **het licht** gehouden

□ Irene Biemond en Jolanda Soeting; met een reactie van Silvester Draaijer

Met meer dan gemiddelde belangstelling lezen Irene Biemond en Jolanda Soeting in *EXAMENS* nummer 4 van 2007 het artikel van Silvester Draaijer e.a. over hun experiment met beeldschermtoetsing. In deze reactie noemen zij een aantal bezwaren tegen de in het artikel van Draaijer e.a. beschreven meervoudige meerkeuzevragen. De redactie vroeg Draaijer te reageren op Biemond en Soeting. Deze interessante discussie beoogt een bijdrage te leveren aan de verhoging van de kwaliteit van beeldschermtoetsing.

Inleiding

In hun artikel ‘Een experiment met beeldschermtoetsing’ beschrijven Draaijer e.a. hoe zij in de digitale tentamens tandheelkunde toetsen of studenten in staat zijn de juiste informatie te vinden en te selecteren. Hiervoor wordt gebruikgemaakt van tandheelkundige casussen met meervoudige meerkeuzevragen. Studenten kiezen per vraag de juiste alternatieven uit een aantal uitklapmenu's. Draaijer e.a. bepleiten dat deze vraagvorm beter aansluit bij het toetsen van een concept (zoals de ziekte van Graves) dan enkelvoudige meerkeuzevragen.

Vraagvormen

Voor diverse opdrachtgevers ontwikkelen wij regelmatig beeldschermtoetsen. Daarbij maken we, indien zinvol, gebruik van andere vraagvormen dan de

traditionele multiplechoicevraag. De laatste jaren zien we steeds meer variatie in vraagvormen, ongetwijfeld door de opkomst van toetsservicesystemen. Zoals vaak gaat deze variatie gepaard met enige of zelfs veel huiver van degenen die liever vasthouden aan het traditionele en vertrouwde. We vinden het dan ook aangenaam te lezen dat bij het Academisch Centrum Tandheelkunde serieus werk wordt gemaakt van het gebruik van nieuwe vraagvormen.

Toch bleven we na lezing van het artikel met gemengde gevoelens achter. Tijdens onze cursus toetsconstructie en in adviestrajecten benadrukken we altijd dat de vraagvorm moet passen bij de inhoud van de vraag én voldoende duidelijk moet zijn. Dit lijkt in de gegeven voorbeelden niet optimaal te zijn gerealiseerd. We zien de volgende twee belangrijke bezwaren bij de voorbeelden van gatenvragen (de figuren 4 en 6 uit het oorspronkelijke artikel zijn hier opnieuw opgenomen).

1. Door de vraagvorm wordt de toetsvraag meer dan wenselijk een puzzel voor de kandidaat.
2. De verschillende onderdelen van de toetsvraag zijn van elkaar afhankelijk; het antwoord op een deel van de vraag is richtinggevend voor het latere deel.

Bezwaren

Het eerste bezwaar wordt ons inziens vooral veroorzaakt door het werken met lijstvensters/uitklapmenu's in het toetsservicesysteem dat men gebruikt. Deze methode zorgt ervoor dat een kandidaat nooit in een oogopslag alle antwoordmogelijkheden kan zien. Afhankelijk van het aantal uitklapmenu's in de vraag en het aantal opties per menu kan het aantal antwoordcombinaties enorm groot zijn. Hierdoor wordt een groot beroep gedaan op het (fotografisch) geheugen van de kandidaat.

Het tweede bezwaar wordt het best zichtbaar bij de gatenvraag uit figuur 6. In het eerste menu kan men kiezen uit verschillende toedieningsvormen van paracetamol, zoals tabletten, stroop of infuusvloeistof. Het vierde menu biedt de keuze van de dosering gecom-



Door de opkomst van toetsservice-systemen zien we steeds meer variatie in vraagvormen

De ziekte van Graves is , waarbij produceert. Deze zijn gericht tegen . Door binding hieraan wordt gestimuleerd

Figuur 4 De open vraag over de ziekte van Graves omgevormd naar een gatenvraag met uitklapmenu's

Het recept voor paracetamol t.b.v patiënt dhr. A.H. luidt als volgt:

R/ paracetamol

dtd no. S.

1^e uitklapmenu

- tabletten
- stroop
- zetpillen
- kauwtabletten
- infuusvloeistof

2^e uitklapmenu

- 750 mg
- 600 mg
- 500 mg
- 50 mg
- 5 mg
- 3 mg
- 6 gram
- 5 gram
- 3 gram

3^e uitklapmenu

- 100
- 24
- 6
- 5
- 2
- 1

4^e uitklapmenu

- maximaal 8 tabletten per dag
- elke 2 uur een tablet
- alle tabletten 1 uur vóór de ingreep innemen
- zo nodig innemen
- 4-6 maal daags 5 ml stroop nemen
- elke 6 uur infuus laten infusen

Figuur 6 Complete gatenvraag met uitklapmenu's en alternatieven

bineerd met de toedieningsvorm (tabletten, stroop, infuus). Het spreekt voor zich dat sommige combinaties bij voorbaat uitgesloten zijn: bijvoorbeeld de toedieningsvorm infuus in combinatie met een aantal tabletten. Maar ook de aantallen kunnen niet altijd met elkaar worden gecombineerd. In het voorbeeld geldt bovendien dat er slechts één combinatie met stroop of een infuus beschikbaar lijkt te zijn. Een dergelijke afhankelijkheid, die zich overigens ook in het andere voorbeeld voordoet, zou ons inziens deels de oorzaak kunnen zijn van de later beschreven hoge betrouwbaarheid van de toets. Je krijgt namelijk een punt voor elke goede keuze en een bonuspunt voor een geheel goed antwoord. Het ligt voor de hand dat

veel kandidaten ofwel (bijna) alle keuzes fout hebben, ofwel (bijna) alle keuzes goed. Door de scoringsmethode wordt het discriminerend vermogen van de vraag ons inziens kunstmatig verhoogd.

Enkelvoudige meerkeuzevragen

Onze voorkeur gaat in dit geval daarom toch uit naar een enkelvoudige meerkeuzevraag. De kunst voor de toetsconstructeur hierbij is te vragen naar het aspect (toedieningsvorm, dosering enzovoort) dat het belangrijkste is voor de betreffende context. Zonder dat wij op de hoogte zijn van het juiste antwoord, valt dan te denken aan de volgende vraag.

*De kunst is te vragen naar het aspect
dat het belangrijkste is
voor de betreffende context*



Inleidende casus enzovoort

De heer A.H. krijgt 6 tabletten paracetamol van 500 mg voorgeschreven.

Wanneer kan deze patiënt de tabletten het best innemen?

- A elke 2 uur een tablet
- B alle tabletten 1 uur voor de ingreep
- C naar behoefte

Iemand heeft de ziekte van Graves.

Welk hormoon maakt hij of zij door deze ziekte meer dan normaal aan?

- A groeihormoon
- B histamine
- C insuline
- D schildklierhormoon

Draaijer e.a. geven in hun artikel aan dat ze graag in één vraag willen toetsen wat de oorzaak van de ziekte van Graves is, omdat een opeenvolging van enkelvoudige meerkeuzevragen tot afhankelijkheid tussen de vragen leidt. Zoals we hiervoor aangaven treedt dat probleem ons inziens net zo sterk of zelfs sterker op bij het ontwikkelen van meervoudige meerkeuzevragen. Ons advies aan de auteurs is dan ook om in dit geval toch te kiezen voor enkelvoudige meerkeuzevragen (waarbij sommige vragen niet in één toets kunnen worden opgenomen). Concreet levert dat bijvoorbeeld de volgende enkelvoudige meerkeuzevragen op:

Wat voor aandoening is de ziekte van Graves?

- A een allergie
- B een auto-immuunziekte
- C een infectieziekte
- D een stofwisselingsziekte

Tegen welke cellen produceert iemand met de ziekte van Graves antistoffen?

- A mestcellen
- B T4-receptoren
- C TSH-receptoren

Casus

We zijn enthousiast over het gebruik van casussen. Dit biedt de toetsconstructeur de mogelijkheid de toepassing van kennis in een authentieke context te toetsen. Daarom tot slot nog een suggestie om de casussen ook op een andere manier in te zetten. Bij de casus met patiëntinformatie, zoals in figuur 1 van het oorspronkelijke artikel, kan de vraag worden gesteld welke aandoening de patiënt mogelijk heeft. Dit is gemakkelijk te realiseren in een enkelvoudige meerkeuzevraag:

Welke aandoening heeft deze patiënt?

- A de ziekte van Graves
- B de ziekte van Kawasaki
- C de ziekte van Wegener

Een dergelijke vraag doet ons inziens meer recht aan de wens van de auteurs om aankomende tandartsen te toetsen op hun vaardigheden in het vinden en selecteren van juiste informatie dan het werken met meervoudige meerkeuzevragen. >>



Verdere technische verfijning van toetssoftware is zeker nodig

Reactie van Silvester Draaijer

Met belangstelling hebben we de reactie op ons artikel over het gebruik van de gatenvragen bij het tentamen tandheelkunde van Irene Biemond en Jolanda Soeting gelezen. Als belangrijkste conclusies stellen zij dat het gebruik van eenvoudige meerkeuzevragen vanuit toetstechnisch oogpunt te prefereren is boven het gebruik van 'gekunstelde' gatenvragen.

De techniek geeft vraagontwikkelaars nieuwe mogelijkheden. Als die niet benut worden kunnen toetsen net zo goed met pen en papier worden afgenomen. Vanuit het oogpunt van verhogen van de acceptatie van digitaal toetsen met gesloten vraagvormen bij een groter publiek (docenten, studenten) is het toepassen van andere vraagvormen dan alleen enkelvoudige meerkeuzevragen volgens ons noodzakelijk. Goed gecontroleerde experimenten zoals beschreven in ons artikel zijn nodig om daarover enige empirische en bruikbare informatie te krijgen. Het beschreven experiment geeft aan dat het gebruik van andere vraagvormen geen uitzonderlijk rare toetsresultaten oplevert. Dat is al een hele winst denken wij.

We zijn het echter eens met de schrijvers dat de technische imperfectie van een systeem als Questionmark Perception (maar ook van andere

systemen) negatieve invloed kan hebben op de prestaties van de student. Het verspringen van regels en het gebrek aan overzicht over alle mogelijke antwoorden is een mogelijke bedreiging voor de validiteit en de betrouwbaarheid. Verdere technische verfijning van toetssoftware is zeker nodig.

De observatie van de schrijvers dat er in het aangedragen voorbeeld een aantal onmogelijke combinaties zitten of juist één set bij een ander combinatie is juist. We denken echter dat het niet strikt noodzakelijk is dat alle combinaties aanwezig zijn om van een goede vraag te kunnen spreken. Ook bij vierkeuzevragen is vaak maar een beperkt aantal combinaties ondergebracht in de alternatieven.

Het feit dat het discriminerend vermogen van de getoonde vraag kunstmatig verhoogd wordt is ons inziens niet erg als het relatieve gewicht van de vraag in de toets dat toestaat. Er is namelijk toetsinhoudelijk pas een probleem als blijkt dat de item-testcorrelatie (zeker van zo'n meervoudige meerkeuzevraag) heel laag of negatief zou zijn. Dan moet een dergelijke vraag nader bekeken worden (is de vraag duidelijk gesteld?, is de stof wel behandeld?, zijn de juiste alternatieven gemarkeerd?) en mogelijk aangepast of verwijderd worden uit de toets, net zoals bij gewone meerkeuzevragen.

□ Drs. I.E.A. Biemond en drs. J. Soeting zijn beiden werkzaam als opleidingskundig adviseur bij Teelen Kennismanagement. Ir. S. Draaijer is als adviseur ICT en Onderwijs verbonden aan het Onderwijscentrum Vrije Universiteit te Amsterdam. s.draaijer@ond.vu.nl; irene.fiermond@teelen.nl