

FOUTE ANTWOORDEN: ONWETENDHEID OF DEVIANT GEDRAG?

W. Bender, onderwijskundige, hoofd Bureau Onderwijs Ontwikkeling
Geneeskunde (BOOG), Rijksuniversiteit Groningen.

De geprecodeerde vraagvorm, bestaande uit een uitspraak waarop de geëxamineerde kan kiezen uit de antwoorden "juist", "onjuist" en "weet niet", wordt op veel plaatsen in het medisch onderwijs gebruikt. Ik laat de rechtvaardiging en gerechtvaardigheid van dit veelvuldig gebruik hier achterwege. In plaats daarvan ga ik nader in op een empirische eigenaardigheid van toetsen, die bestaan uit vragen van bovenstaand type (in het vervolg aangeduid als j/o/?-items).

Het materiaal waaraan de genoemde eigenaardigheid zich laat demonstreren, bestond uit zes verschillende toetsen, die in totaal aan 52 verschillende groepen werden voorgelegd. De toetsen betroffen ofwel algemeen medische kennis op het niveau van het artsexamen ("voortgangstoetsen") ofwel discipline-georiënteerde kennis. Het aantal vragen lag tussen 80 en 120. De 52 getoetste groepen waren medische studenten (derde- tot en met zesdejaars), huisartsen-in-opleiding, huisartsen en aan perifere ziekenhuizen verbonden specialisten. De groepsgrootte lag tussen 25 en 250, met uitzondering van twee groepen met elk 10 personen.

Omdat de groepen qua gemiddelde leeftijd en bijgevolg qua kennis en competentie nogal uiteenliepen, is het niet verwonderlijk dat de gemiddelde groepsresultaten -uitgedrukt als percentage goede antwoorden- eveneens aanzienlijke variatie vertoonden. De laagste groepsscore was 25% en de hoogste score bedroeg 75%. Deze grote spreiding kan overigens, behalve met leeftijds- en competentiever- schillen, ook te maken hebben met verschillende moeilijkheidsgraden van de zes toetsen. Het verbazing- wekkende echter is, dat de pro-

portie foute antwoorden -binnen zekere grenzen- niet varieert met leeftijd, kennis, competentie of moeilijkheidsgraad. Anders gezegd: het gemiddelde percentage foute antwoorden is een constante. De waarde van die constante ligt tus- sen de 19% en 29%, met een gemid- delde waarde van 23%.

Het bovenstaande roept tenminste twee vragen op: waarom worden er überhaupt zoveel foute antwoorden gegeven en waarom gaat het om een constant percentage?

In het algemeen betekent een fout antwoord ofwel foutieve kennis ofwel een onfortuinlijke gissing voortkomend uit onwetendheid of- wel een logisch gevolg van een inferieure vraagredactie. De twee- de verklaring is hier echter niet of slechts in beperkte mate van toe- passing. Immers, de derde optie bij de j/o/?-vraag probeert gissen te voorkomen en doet dat ook, getuige het feit dat ruimschoots gebruik wordt gemaakt van het weet niet-antwoord. Er is dus geen reden om aan te nemen, dat het constante fout-percentage moet worden verklaard via verkeerd uitgevallen gissingen. Ook de der- de verklaring, de inferieure vraagredactie, is niet van toepas- sing. Het betrof in alle gevallen zorgvuldig geredigeerde toetsen met vragen waarvan de formulering ook na item-analyse geen reden tot bezorgdheid gaf.

De constante fout zou derhalve te maken moeten hebben met foutieve kennis; daarbij blijft het echter raadselachtig dat het volume fou- tieve kennis niet varieert met de gemiddelde leeftijd van de partici- perende groepen. Een mogelijkheid echter om dit raadsel op te lossen, zou kunnen liggen in een uncon- ventionele kijk op de term "fou- tieve kennis".

Het is een bekend feit, dat medici nogal van elkaar verschillen in de manier waarop ze medische problemen oplossen. Geef tien dokters dezelfde tien patiënten en het resultaat is honderd verschillende diagnoses en honderd verschillende behandelingen. De overdrijving is in het voorgaande natuurlijk groterlijks aanwezig, maar het is realistisch om te stellen, dat in veel opzichten consensus met betrekking tot medisch management ontbreekt. Welnu, wanneer probleemoplossend gedrag van medici wordt gekenmerkt door een gebrek aan consensus, dan zou dat kenmerk zich ook weleens kunnen voordoen op het strikt cognitieve terrein. Dat zou kunnen betekenen, dat de constante fout bij j/o/?-toetsen moet worden verklaard als gebrek aan consensus. Dit zou betekenen, dat een fout antwoord eerder deviant gedrag is dan foutieve kennis of onwetendheid. Anders gezegd: door het geven van een fout antwoord verklaart de examinandus dat hij het niet eens is met de antwoordsleutel van de examiner. Het aardige van deze verklaring is, dat zij experimenteel kan worden onderzocht. Daartoe werd de volgende opzet bedacht.

Aan drie groepen, elk bestaande uit circa 65 huisartsen, werd een toets voorgelegd, bestaande uit 120 vragen op het gebied van de oncologie. De vragen waren door een panel van (niet deelnemende) huisartsen en universitaire docenten geselecteerd op relevantie. Halverwege de toets werden groep 1 en groep 2 in algemene zin gewezen op de mogelijkheid, dat op sommige vragen meningsverschillen kunnen rijzen tussen huisartsen en oncologisch specialisten met betrekking tot het goede antwoord. Verder werd groep 1 geïnstrueerd de toets af te maken ALSOF zij geen huisarts maar oncoloog waren. De huisartsen van groep 2 maakten de toets op de gebruikelijke manier, maar kregen halverwege de opdracht om die vragen die naar hun oordeel bij oncologen een an-

der antwoord zouden krijgen dan bij huisartsen, van een merkteken te voorzien. Met andere woorden: groep 2 werd geïnstrueerd contro-versiële items te identificeren.

Groep 3 was de controlegroep en maakte de toets zonder een speciale introductie of instructie. De drie groepen bleken gelijkwaardig te zijn, hetgeen kon worden vastgesteld door vergelijking van de resultaten op de eerste testhelft.

Wanneer de constante fout het gevolg is van deviant gedrag, dan zal groep 1 op deze oncologietoets beter presteren dan de andere groepen. Voorts zou het zo moeten zijn, dat er geen verschil is tussen groep 2 en groep 3; echter, het veranderen van de antwoordsleutel bij de contro-versiële items in het tegenovergestelde, zou de score van groep 2 moeten doen stijgen.

Wel, de resultaten zijn geen tabel of grafiek waard. Géén van de voorspelde effecten deed zich voor, niet bij individuen, niet bij groepen. Ook het wijzigen van de antwoordsleutel, hetzij voor individuen, hetzij voor groepen, gaf geen enkel voedsel aan de veronderstelling dat toetsvragen van het j/o/?-type deviant gedrag kunnen uitlokken. Derhalve houdt een verklaring van het constante foutenpercentage in termen van een welbewuste afwijking van de bedoelingen van de examiner, geen stand. Dat heeft twee plezierige consequenties. In de eerste plaats hoeven we het j/o/?-vraagtype niet af te schaffen (althans niet om redenen van deviant gedrag). In de tweede plaats zal het onderzoek inzake het raadsel van de foutenconstante moeten worden voortgezet.

Summary

Tests of the t/f/? type are attended with a constant percentage of incorrect answers. It was hypothesized that this could be explained as explicit deviant behavior, namely examinees deliberately deviating from the intentions of the

examiner. Instructing examinees to answer as if they were the examiner therefore should improve their score. Instructing examinees to identify items on which they disagree with the examiner, and consequently manipulating the scoringkey of those items, should also improve their score.

However, none of the predicted effects occurred, not with individuals, not with groups. Therefore the explanation of the constant percentage of incorrect answers in terms of deviant behavior has to be cancelled.

Met dank aan Dr. H.J.M. van Rossum voor zijn onverbiddelijke uitwerking van het begrip "gebrek aan consensus".

Van de redactie

De redactie was aangenaam verrast door de Engelse samenvatting bij het artikel van Bender.

Wij beschouwen dit als een compliment. Toekomstige auteurs van dit bulletin die hun artikel ook van een summary willen voorzien, zullen door de redactie positief worden bejegend.

Het raadsel van de "fouten-constante", dat door Bender wordt beschreven, is zeer intrigerend. De redactie is benieuwd naar de ervaringen met j/o/?-items bij andere faculteiten en nodigt dan ook anderen uit hun ervaringen hiermee op schrift te stellen en ter publicatie aan te bieden.