

Hoeveel centimeters gaan er in 4,6 meter?

Columniste Aleid Truijens is geschrokken van het rekenniveau van eerstejaars PABO-studenten (*De Volkskrant* 23/6/2007): velen zijn nog niet eens in staat om de bovenstaande opgave uit de CITO-rekentoets te beantwoorden. Truijens is verkeerd geïnformeerd. Ten eerste vergeet ze te vermelden dat de rekentoets voor PABO-studenten een tijdslimiet heeft (2 minuten per vraag). Ten tweede, en dat is erger, heeft ze niet begrepen dat het hier een **adaptieve plaatsingstoets** betreft. Je kunt die toets vergelijken met een sorteermachien voor kippe-eieren. De zwaar onvoldoende kriel-eitjes glijden in de eerste bak, de iets te kleine eieren in de tweede bak, en zo door tot de vierde bak, waar de eieren inglijden die ruim boven de maat zijn. Truijens citeert in haar column opgaven die vermoedelijk met name bedoeld zijn om misera-bele kriel-eitjes te identificeren.

Adaptieve toetsen zijn zeer efficiënt om studenten in niveaugroepen te verdelen, waarna elke groep een aangepaste onderwijsbehandeling krijgt aangeboden. Het schokkende is echter dat de PABO-rekentoets niet alleen gebruikt wordt om aankomende studenten in niveaugroepen onder te brengen, al naargelang de grootte van hun rekendeficiënties, maar ook als selectietoets aan het eind van het eerste studiejaar. De eieren uit de eerste en tweede bak worden ongeschikt verklaard voor verdere deelname aan de opleiding. De vraag is nu: is een adaptieve plaatsingstoets-met-tijdslimiet geschikt om keihard te differentiëren tussen voldoende (vierde en derde bak) en onvoldoende (tweede en eerste bak)?

Hoezo niet? Om dat te argumenteren moeten we de metafoor van het eiersorteermachien verlaten: studenten zijn geen eitjes maar mensen van vlees en bloed die anderhalf uur lang een stressvolle computertoets moeten afleggen. We kunnen beter een andere metafoor kiezen: polsstokhoogspringen. Kandidaten die niet hoger dan 3 meter kunnen springen, die vallen af. In een adaptieve toets wordt de lat om te beginnen op 2,5 meter gelegd. Kom je daar niet binnen twee minuten overheen, dan mag je niet nogmaals de 2,5-meterlat proberen, maar moet je terug naar 2 meter. Ondertussen krijgt de kandidaat niet te horen hoe hoog de lat in eerste instantie lag, en hij krijgt evenmin te horen dat de lat vervolgens verlaagd is. Uiteindelijk wordt de kandidaat afgewezen, terwijl hij niet de informatie en de gelegenheid heeft gekregen die hij nodig had om aan te tonen dat hij in staat is over de 3-meterlat heen te komen.¹

Het CITO beroemt zich op het feit dat adaptieve toetsen veel goedkoper zijn dan traditionele tentamens. Een eiersorteermachine is inderdaad veel goedkoper dan intensieve metingen met een schuifmaat. Maar als er, zoals in een tentamen, grote belangen op het spel staan, dan verdient het de voorkeur een gevoelige schuifmaat te gebruiken die met voldoende nauwkeurigheid en transparantie het gebied tussen zakken en slagen aftast.

"Meneer Jansen, waarom ben ik gezakt voor mijn rekentoets? Ik had de meeste opgaven goed!"

"Ja meid, dat zit zo: je hebt wel veel opgaven goed gescoord, maar dat waren nou juist de *makkelijke* vragen, die niet meetellen voor een voldoende. De computer heeft die vragen geselecteerd omdat je de *moeilijker* vragen niet stante pede wist te beantwoorden."

"Maar meneer Jansen, waarom zijn die vragen dan in het tentamen opgenomen?"

"Ja, dat weet ik ook niet! Het CITO zegt dat zo'n toets veel goedkoper is."

Ik ben geen toetsdeskundige, maar mijn intuïtie zegt dat er iets mis is als men de CITO-rekentoets voor pass/fail-selectie gebruikt. In de Nederlandse leraarsopleidingen zijn ongetwijfeld mensen die deskundig zijn op het gebied van toetsconstructie. Hebben zij zich vergewist van de merites van adaptieve toetsen in het algemeen en van adaptieve toetsen-met-tijdslimiet in het bijzonder? Als ze in de bestaande psychometrische literatuur geen bevestiging kunnen vinden voor de stelling dat zo'n toets geschikt is als summatief, selectief tentamen om de schapen van de bokken te scheiden, laten zij dan hun stem verheffen!²

Wes Holleman

weblog onderwijs 25-6-2007

www.onderwijsethiek.nl

¹ Bij polsstokhoogspringen daarentegen wordt de lat telkens één tree hoger gelegd. Als de sprong bij dezelfde tree driemaal mislukt, mag je niet meer naar een hogere tree meedingen.

² Zo'n literatuurstudie zal trouwens niet gemakkelijk zijn. Amerikaans onderzoek moet met grote argwaan worden bekeken, aangezien in het Amerikaanse onderwijsbestel nauwelijks onvoldoendes worden uitgedeeld. Er wordt gescoord op een schaal van A t/m F, waarbij alleen de slechtste 5 of 10% een onvoldoende (F) krijgt. Amerikaanse tentamens en examens zijn meestal niet *criterion-* maar *normreferenced*; ze hebben primair tot doel studenten al naargelang hun prestaties in een rangorde te zetten..., netzoals de eieren in een eiersorteermachien.