

De kwaliteit van schoolexamens (II)

Ik ben maar een eenvoudige lesboer. Zo begon hij zijn bespiegeling. Ik laat ze de stof doorploegen en dan krijgen ze er een proefwerk over. Het doel van die toets is vast te stellen of ze de stof voldoende beheersen. Wie de stof voldoende beheerst, krijgt een zes. Wie iets beter presteert, krijgt een zeven en wie de toets heel goed gemaakt heeft, krijgt een acht of hoger. Mijn tienpuntsschaal is eigenlijk een combinatie van twee schalen: (a) voldoende of onvoldoende? en (b) in welke mate (on)voldoende? Mijn toetsen zijn er niet op gemaakt om exact vast te stellen of iemand een zeven dan wel een acht verdient. Als ik een acht geef, bedoel ik dat hij beter gepresteerd heeft dan een klasgenoot die een zeven heeft gekregen, maar ik heb niet de pretentie te kunnen verantwoorden hoévél beter een acht is in vergelijking met een zeven. In absolute zin discrimineren mijn toetsen slechts binnen het cijfergebied rond de cesuur, dus tussen 4,5 en 6,5. Dat spreekt vanzelf: ik wil niet dat ze ten onrechte slagen en ook niet dat ze ten onrechte zakken. Aan het eind van het trimester bereken ik voor iedere leerling een rapportcijfer, een soort gewogen gemiddelde van de toetscijfers. Wiskundig gezien is dat uiterst aanvechtbaar: van scores op een ordinale schaal valt geen rekenkundig gemiddelde te berekenen.¹ Maar ik vind het prima zo: (a) ik meet (absoluut) of hun toetsprestatie voldoende is; (b) de leerlingen komen (relatief) te weten hoe goed ze zijn in vergelijking met hun klasgenoten; en (c) op het rapport komt een cijfer waarin, naast mijn absolute beoordelingen (a), ook hun relatieve positie ten opzichte van hun klasgenoten (b) enigszins is meegewogen.

Nu ik er zo over nadenk, hebben de hogere cijfers, in de range tussen 7 en 10, voor mij een heel ander doel dan de cijfers tussen 4,5 en 6,5. De hogere cijfers dienen er vooral toe de zesjescultuur bij de leerlingen te doorbreken, hen tot bovengemiddelde prestaties aan te zetten en hen daarvoor een pluim te geven. Zoiets als een tien met een griffel? Nee, een zeven is een zes met een griffel, een acht representeert een zes met twee griffels, een negen staat voor een zes met drie griffels, en een tien is voor de meester.

Maar van de bovenbouwcoördinator hebben we onlangs opdracht gekregen minder scheutig te zijn met hoge cijfers. De school wordt namelijk door de *Onderwijsinspectie* gestraft als het rekenkundig gemiddeld cijfer van het schoolexamen (gemiddeld SE-cijfer, gerekend over alle kandidaten en alle vakken) teveel boven het gemiddeld cijfer van het centraal schriftelijk (CE) ligt.² Er is zelfs een *wetswijziging* in de maak waardoor de school haar examenlicentie kan kwijtraken als de discrepantie tussen het gemiddeld SE- en CE-cijfer te groot is.³

Dat betekent in feite dat we SE-tentamens moeten gaan construeren die niet alleen discrimineren tussen voldoende en onvoldoende toetsprestaties (in de range tussen 4,5 en 6,5), maar ook tussen beheersingsniveaus die *in absolute zin* een zes, zeven, acht, negen of tien waard zijn. Ik denk dat zoiets met veel kunst- en vliegwerk te doen is, maar dan moet je ellenlange toetsen en beoordelingsmatrijzen maken: deel I is bedoeld om de range tussen 4,5 en 6,5 af te dekken, deel II voor de range tussen 6,5 en 7,5 en deel III voor de beheersingsniveaus die hogere cijfers waard zijn. Ik denk dat het kan, maar volgens mij moet je dat niet willen, want op die manier ontaardt een onderwijsinstelling in een toetsfabriek.

Volgens mij moet de Onderwijsinspectie een andere maat bedenken voor de kwaliteit van het schoolexamen. Bijvoorbeeld: stel het aantal kandidaten die voor het gehele eindexamen geslaagd zijn op 100% en bereken vervolgens het percentage dat gezakt zou zijn als de SE-cijfers buiten beschouwing zouden zijn gelaten. Op die manier komen scholen aan de schandpaal als ze zwakke kandidaten matsen, maar gaan ze vrijuit als hun ordinale schaalcijfers in de range van 7 tot 10 iets te hoog zijn uitgevallen.

Wes Holleman

weblog onderwijs 28-6-2010

<http://www.onderwijsethiek.nl>

¹ Op Noord-Amerikaanse scholen hanteert men om die reden een letter- in plaats van een cijferschaal: van A+ (uitmuntend) tot D- (onder de maat). Maar zij bezondigen zich wel aan de gewoonte op basis van deze ordinale letterschaal een GPA (Grade Point Average) in cijfers te berekenen. Relatieve beoordeling en classificatie van prestaties (grading on the curve) is in hun competitieve schoolbestel niet ongebruikelijk.

² *De kwaliteit van schoolexamens (I)*: <http://www.onderwijsethiek.nl/?p=1689> (15/5/2010).

³ <http://www.rijksoverheid.nl/ministeries/ocw/documenten-en-publicaties/persberichten/2010/06/25/examenlicentie-kwijf-bij-grote-verschillen-examencijfers.html>.