

## DE CESUURBEPALING

Uit: Bijlage 4, [Hoofdstuk 11 Toetsbeleid](#), Handboek Onderwijskwaliteit

Na het inlezen van de antwoorden van de studenten en het inlezen van de antwoordsleutel van het tentamen wordt voor iedere student de toetsscore bepaald. Om deze toetsscores vervolgens naar cijfers te kunnen omzetten dient een docent allereerst te bepalen welke toetsscore overeenkomt met het cijfer 6.0. Het bepalen van deze grens tussen voldoende en onvoldoende wordt de cesuurbepaling genoemd. Daarna kunnen ook de overige toetsscores worden omgezet tot cijfers (zie Bijlage 3, H 11 Toetsbeleid, Handboek Onderwijskwaliteit).

Voor het vaststellen van de (voorlopige) norm voor een voldoende zijn diverse methoden mogelijk. Het belangrijkste onderscheid hierbij is of het om een absolute dan wel een relatieve methode gaat.

Bij een absolute methode ligt de norm in de stof en staat dus vooraf vast. Bij een relatieve methode daarentegen is de norm afhankelijk van de groepsprestatie: hierbij wordt de norm dus pas achteraf vastgesteld.

Wie bij gebruik van een relatieve methode toevallig in een hoog scorende groep zit en relatief laag scoort, heeft dus pech; met diezelfde score zou in een minder goed scorende groep een hoger cijfer behaald zijn (wat heel goed het verschil wel/niet een voldoende zou kunnen betekenen).

Omdat relatieve methoden dus groepsafhankelijk zijn, wordt het gebruik ervan dringend afgeraden.

Bij de cesuurbepaling volgens een absoluut criterium is er meer dan één manier om tot de cesuur te komen. Het is aan de docent te bepalen welke manier hij/zij daarbij wil gebruiken.

De mogelijkheden zijn onder andere:

- de '60-procent-methode';
- de methode Ebel;
- de methode Nedelsky;
- de kernitem-methode van De Groot.

De meest gebruikelijke methode: de 60-procent-methode, is vrij rigide. Als uitgangspunt hierbij geldt dat het tentamen zelf en het onderwijs dat aan het tentamen vooraf is gegaan, van goede kwaliteit zijn. Deze aanname is in de praktijk natuurlijk niet altijd even onproblematisch. Dit betekent dat eventuele kwaliteitsgebreken van tentamen en/of onderwijs geheel voor rekening van de student komen.

De methodes Ebel en Nedelsky, en de kernitemmethode, zijn wat genuanceerder. Bij de methode Ebel wordt bijvoorbeeld ook de moeilijkheid en de relevantie van de vragen meegenomen in de bepaling van de grens voldoende/onvoldoende. Ook de methode Nedelsky en de kernitemmethode berusten op inschattingen van de docent, zij het dat de aannamen bij de kernitem-methode wat minder stringent zijn.

Uit onderzoek is echter gebleken dat docenten doorgaans niet goed in staat zijn de moeilijkheid van hun eigen vragen in te schatten, hetgeen betekent dat er nogal wat vraagtekens te plaatsen zijn bij de bruikbaarheid van deze meer genuanceerde methoden in de praktijk. Dit geldt met name voor de methoden Ebel en Nedelsky (zie voor een meer gedetailleerde bespreking van deze methoden Dousma & Horsten, 1989).

Om die reden wordt docenten toch aangeraden, ofwel de 60-procent methode, ofwel de kernitem-methode te gebruiken. Beide methoden worden hier nader toegelicht.

### ***De 60-procent-methode***

De benaming van deze methode is enigszins misleidend, omdat hiermee wordt gesuggereerd dat altijd zou moeten worden uitgegaan van zestig % 'goed' voor een voldoende. Dit is in de praktijk weliswaar vrij gebruikelijk, maar zeker geen wet van Meden en Perzen.

Het is in principe aan de docent te bepalen welke score behaald moet worden (met andere woorden, hoeveel van de stof beheerst moet worden), om een voldoende te halen. Vaak ook geldt hiervoor een vakgroeps- of facul- teitsnorm. In zo'n geval zal een docent gegronde redenen moeten hebben om van deze norm af te wijken.

Wanneer het vereiste percentage 'goed' is vastgesteld, moet voor het bepalen van de cesuur nog worden gecorrigeerd voor de raadkans.

### **Voorbeeld van een cesuurberekening volgens de 60-procent-methode**

Voorbeeld van de cesuurberekening (zak/slaaggrens) voor een meerkeuzetentamen.

Stel: een meerkeuzetentamen bestaat uit 60 vier-keuzevragen. Voor het berekenen van de cesuur wordt eerst de raadkans van het totaal afgetrokken: dat is voor 4-keuzevragen 25% van het totaal, dus een score van 15. Voor deze eerste 15 punten worden dus geen cijferpunten toegekend.

In dit geval bepaalt de docent dat men 60% goed moet hebben om een voldoende te krijgen. Die 60% wordt berekend over de score na aftrek van de kansscore: dat is in dit voorbeeld 60% over 45 = 27.

De grens voldoende/onvoldoende komt nu te liggen op 27 + de raadkans (15): men moet dus 42 van de 60 vragen goed hebben voor een voldoende (een zes). In formule:

$$\text{Cesuur} = 0.25 \times 60 + 0.6 \times (60 - 0.25 \times 60) = 42.$$

### ***De kernitemmethode van De Groot***

In het verwerkingsprogramma van het Onderwijscentrum bestaat de mogelijkheid de cesuur via het aanwijzen van kernitems vast te stellen.

Een kernitem is een 'kritische' tentamenvraag die aan een aantal eisen moet voldoen. De vraag betreft de kern, de essentie van de leerstof en is onderscheidend tussen hen die deze kernstof beheersen en hen voor wie dit niet geldt.

Een kernitem 'mikt' dus als het ware voor een belangrijk onderdeel van de leerstof op de grens tussen voldoende en onvoldoende (De Groot, 1964).

De basisgedachte van de kernitem-methode is verder als volgt. Ieder kernitem dat voldoet aan de eisen, is op zichzelf genomen een geschikte toets om te zien of een leerling een voldoende verdient. De afname van één ideaal kernitem zou rechtstreeks het percentage studenten met een voldoende op kunnen leveren, namelijk zijnde het percentage goede antwoorden op dat kernitem. Eén kritische vraag is echter onvoldoende betrouwbaar om als criterium voor voldoende en onvoldoende te dienen, zodat men een aantal vragen als kernitem aanwijst (geadviseerd wordt 20-25% van het totale aantal vragen) en de procedure herhaalt. De kernitems worden niet in eerste plaats gebruikt om aan studenten een plus of min uit te delen - hoewel die plussen en minnen natuurlijk wel in hun totaalscore meespelen - maar om **tot een schatting te komen van het percentage studenten dat 'verdient' te slagen**. In de praktijk komt het er op neer dat het gemiddelde percentage goede antwoorden van de kernitems (de gemiddelde p-item van de kernitems) gelijk gesteld wordt aan het percentage studenten dat een voldoende krijgt voor het tentamen als geheel.

De antwoorden van de studenten op de kernitems worden dus gebruikt als schatting van het percentage studenten dat een voldoende verdient. Omdat men gewoonlijk meer wilt weten dan alleen wie voldoende en wie onvoldoende heeft en ook binnen het voldoende en onvoldoende gebied wil differentiëren, neemt men naast de kernitems een flink aantal andere items in het tentamen op. De antwoorden van de studenten op het tentamen als geheel (dus kernitems en de rest van de vragen) bepaalt uiteindelijk de eindscore van de individuele studenten.

#### *Voorbeeld:*

Een toets bestaat uit 30 vier-keuzevragen. De docent heeft vooraf zeven items als kernitems aangewezen. Na de verwerking van het tentamen blijken de p-items van deze vragen respectievelijk: 0.50, 0.63, 0.64, 0.74, 0.72, 0.44 en 0.58. De gemiddelde proportie goede antwoorden van deze kernitems bedraagt 0.61. Dit betekent dat 61% van de studenten een voldoende verdient. Met behulp van het cumulatieve percentage overzicht van de toetsscores kan nu de daarbij behorende cesuurscore afgelezen worden. In het verwerkingsprogramma wordt dit automatisch verzorgd. De docent dient alleen op het verwerkingsformulier de vraagnummers van de kernitems aan te geven.

*Opmerking:*

In het verwerkingsprogramma wordt gerekend met de 'gewone' p-item. Dit betekent dat er van uit wordt gegaan dat de studenten niet raden. Docenten die de kernitem-methode willen gebruiken met een correctie voor raden, kunnen dit doen door met de hand de voor raden gecorrigeerde gemiddelde p-item van kernitems te bepalen en met behulp van het cumulatieve percentage overzicht van de toetscores de uiteindelijke cesuurscore te bepalen. Deze score kan dan vervolgens weer in het verwerkingsprogramma worden ingevoerd.

**Samenvatting van de werkwijze bij het bepalen van de cesuur**

Kies voor het bepalen van de grens voldoende/onvoldoende bij voorkeur één van de hier genoemde absolute cesuurmethoden: de 60-procent-methode, de kernitem-methode, de methode Ebel of de methode Nedelsky. De laatste twee methoden zijn in de praktijk wat minder gebruikelijk en worden hier niet nader omschreven (zie voor meer informatie o.a. Dousma & Horsten, 1989). Denk er aan dat bij de 60-procent-methode gecorrigeerd moet worden voor de raadkans.

Bij het gebruik van de kernitem-methode moeten de kernitems worden aangegeven.

Wanneer u heeft gekozen voor een van de andere absolute methoden, kunt u ermee volstaan de cesuurscore aan te geven op het verwerkingsformulier.

Alle hier aangegeven methoden zijn slechts hulpmiddelen om op meer rationele wijze de grens tussen voldoende en onvoldoende te trekken. Welke methode u in een concreet geval wilt gebruiken, hangt af van het gewicht dat u wilt toekennen aan de diverse factoren die bij de methoden een rol spelen. Over het algemeen kan gezegd worden dat het voor de docent en de student duidelijk en inzichtelijk moet zijn hoe de cesuur bepaald wordt. Gegoochel met ingewikkelde formules is af te raden.