

## 2.8 Leerdoelen en toetsen Achtergrondinformatie

### Toetsen

#### Inleiding

Bij de bepaling van de waardering die de leerlingen voor een leraar hebben, speelt de manier waarop deze omgaat met toetsing en beoordeling een belangrijke rol. Gelegenheid tot vragen stellen in de les ervoor, snel nakijken, duidelijke toetsen en toegankelijk zijn voor bezwaren tegen de correctie zijn daarbij belangrijk. Bij dat laatste moeten leerlingen het gevoel hebben dat er serieus naar hen geluisterd is. Alleen dan zullen ze een besluit om het cijfer niet te veranderen kunnen aanvaarden, zij het vaak onder protest. Besteed daarom veel aandacht aan de zorgvuldigheid ten aanzien van beoordeling en de manier waarop je op dat punt met leerlingen omgaat. Er is geen beter middel denkbaar om een goede relatie met een klas te bederven dan het geven van een in de ogen van de leerlingen onrechtvaardige toets, het maken van tendentieuze opmerkingen bij het teruggeven (zoals: “Jij had natuurlijk weer een onvoldoende”) of het niet willen luisteren naar bezwaren tegen de correctie. De leraar doet er dan ook goed aan zoveel mogelijk te laten zien op welke wijze hij tot zijn beoordeling is gekomen, en ruime aandacht te besteden aan de bezwaren van leerlingen daartegen.

Leraren hebben wel eens de neiging om een klas waarmee ze een slechte relatie hebben af te straffen door het geven van moeilijke toetsen of door laag te cijferen in de hoop dat de leerlingen daardoor tot inkeer zullen komen. Het tegenovergestelde is meestal het geval. Het gebruik van cijfers en toetsing als straf is niet alleen theoretisch niet verantwoord, maar werkt in de praktijk alleen maar averechts.

Spreek toetsen ruim van tevoren af en hou je aan de schoolregels zoals ‘niet meer dan twee toetsen per dag’. Leerlingen stellen het zeer op prijs als je met hen overlegt over een geschikte datum, in plaats van die als een dictaat af te kondigen.

#### Normeren

Het maken van een normering begint met het maken van het antwoordmodel van de toets. Zet daarna op dat antwoordmodel eerst een voorlopige en ten slotte een definitieve normering, bijvoorbeeld op de volgende manier.

**Basisvraag** – Kies een vraag uit waarvan het antwoord uit eenvoudige verschillende delen bestaat, bijvoorbeeld: “De grootheden A en B zijn recht evenredig. Geef op drie verschillende manieren aan wat dat betekent.” Je kunt dan voor dat onderdeel 3 of 2 punten geven. In het laatste geval krijgt de leerling bijvoorbeeld 1 punt als twee antwoorden goed zijn en niks als er maar een goed is.

**Punten per vraag** – Gebruik die basisvraag als uitgangspunt om het aantal punten voor de andere vragen vast te stellen. Geef niet meer dan 4 en bij uitzondering 5 punten voor een vraag. Streef niet naar een totaal van 100. Bij een toets van een lesuur leidt dat tot ongeveer 7 punten per vraag. Zo’n gedetailleerde normering kost veel tijd omdat er dan per vraag veel meer mogelijkheden zijn tot het gedeeltelijk toekennen van punten aan een onvolledig of deels onjuist antwoord.

**Onderverdeling per vraag** – Bekijk dan elke vraag apart en stel vast hoe je de beschikbare punten verdeelt over de verschillende aspecten van het bedoelde antwoord. Soms kan dat leiden tot een verandering van het aantal punten voor een vraag.

**Omrekenen naar eindcijfer** – Tel dan de punten op. Nu wordt er meestal als laagste cijfer de 1 gehanteerd, omdat de 0 als eindcijfer officieel niet bestaat. Bij het eindexamen krijgt de kandidaat 10 punten vooraf, en iemand die een leeg blaadje met alleen zijn naam erop inlevert haalt dus een 1. Om dat te bereiken tel

je bij het totaal nog 10% op. Is het totaal bijvoorbeeld 31, dan wordt het maximaal te behalen aantal punten dus  $31 + 3 = 34$ . Elke leerling begint dus al met 3 punten. Stel dat een leerling voor zijn werk 21 punten haalt. Zijn totaal is dan  $21 + 3 = 24$  en het cijfer wordt  $24/34 \times 10 = 7,1$  of 7 – dat hangt ervan af hoe precies je de waardering vaststelt.

**Controle** – Deze stap is zeer belangrijk. Krijgt de leerling die in jouw ogen een 6 verdient dat eindcijfer ook? Loop de toets na en stel bij elke vraag vast of een leerling die een 6 moet kunnen halen die vraag moet kunnen maken. Reken dan uit of hij die 6 ook haalt. Als het resultaat aanzienlijk boven of onder de 6 zit, moet de normering worden bijgesteld.

**Standaardfouten** – Bedenk hoe je rekenfouten en fouten met afronden, significantie, eenheden en dergelijke rekent. Daarvoor is de examennorm voor het omgaan met dit soort fouten bruikbaar, maar vergeet niet dat een punt aftrek bij een toets met een maximum van 40 punten zwaarder telt dan bij een examen met een maximum van 100.

Overigens is het toetsen natuurlijk ook een meting. Je meet in hoeverre de leerling voldoet aan de eisen. Uit fysisch oogpunt is de meting zeer gebrekkig. Een ander meetinstrument (toets) zal bij een aantal leerlingen leiden tot een ander meetresultaat. De nauwkeurigheid van de meting is dan ook laag en het geven van een cijfer tot een decimaal achter de komma is eigenlijk onverantwoord. Leerlingen en ouders protesteren nogal eens tegen afrondingen tot een 10-puntsschaal. Sommige scholen hanteren als tussenvorm de  $5\frac{1}{2}$ , de  $6^-$  en de  $6^+$ . Ook hanteren scholen wel een 4-puntsschaal (goed-voldoende-bijna voldoende-onvoldoende) of zelfs een 2-puntsschaal (voldoende-onvoldoende). Dergelijke vormen worden in de bovenbouw zelden toegepast, maar in andere landen is dat heel gebruikelijk: in Engeland hanteert men A t/m F, en in Duitsland 1 t/m 5.

### Corrigeren

Dan begint de correctie. Schrijf beslissingen over de normering van vaak voorkomende fouten in een andere kleur op het antwoordmodel. Dat bevordert de consistentie van de correctie. Het schrijven van toelichtingen bij fouten van leerlingen is voor de leerling prettig en heel nuttig, maar kost zeer veel tijd. Veel gemaakte fouten kunnen beter klassikaal besproken worden.

De ervaring leert dat de normering op enkele punten bijgesteld moet worden na het corrigeren van zo'n vijf toetsen. Doe dat niet al te rigoureuus: het kost veel tijd en het effect ervan op het eindresultaat is meestal marginaal. Tijdens de correctie zul je nogal eens door het onbehaaglijke gevoel beslopen worden dat de toets te gemakkelijk of te moeilijk was. Begin dan alleen in extreme gevallen opnieuw met nakijken. Maak gewoon de correctie af en bekijk de toets nog eens kritisch. Was hij echt te simpel of te moeilijk? Vraag het bij twijfel aan een collega.

Verandering van de norm kan achteraf gebeuren door verandering van het aantal punten dat de leerling ook zonder prestatie krijgt (die 10%). Een norm die achteraf aangescherpt wordt is moeilijk aan leerlingen te verkopen. Ze zullen dan snel het gevoel hebben onrechtvaardig behandeld te worden. Ze zijn daar nogal gevoelig voor, en vergeten snel de keren dat de norm versoepeld werd. Als je een slechte relatie met een klas hebt, is het meestal niet verstandig een norm achteraf aan te scherpen – zelfs niet als dat eigenlijk zou moeten. Dat zou de relatie met de klas blijvend kunnen beschadigen.

Tijdens het corrigeren kom je altijd zwakke punten van de toets op het spoor. Verbeter die na de correctie meteen, dan heb je volgend jaar een betere toets tot je beschikking. Uitstellen van het verbeteren tot volgend jaar werkt niet: de meeste verbeteringen zijn dan allang weer vergeten. Het noteren van later aan te brengen verbeteringen kost evenveel tijd als het verbeteren zelf. Houd ook aantekening van de scores van de leerlingen, zodat je later weet of het een (te) moeilijke of juist een gemakkelijke toets was.

### Analyseren

Na het corrigeren kun je desgewenst de toets verder analyseren door voor elke vraag de p-waarde en de D-waarde te bepalen met behulp van een *scorematrix* (zie figuur 1).

**Moelijkheidsgraad** – De *moelijkheidsgraad* (p-waarde) van elke vraag is de gemiddelde score van alle leerlingen op de vraag, gedeeld door de maximale score op diezelfde vraag. Hoe lager de p-waarde van een vraag is, des te moeilijker was die vraag.

**Discriminatie-index** – De *discriminatie-index* (D-waarde) geeft aan in hoeverre de score op een vraag overeenstemt met de score op de gehele toets. De D-waarde bereken je als volgt:

- Verdeel de leerlingen in drie ongeveer even grote groepen: de topscorers, de middenmoters en de laagscorers.
- Bepaal voor elke vraag de gemiddelde score van de topscorers en de gemiddelde score van de laagscorers.
- Trek deze twee gemiddelde scores van elkaar af, en deel het verschil door de maximale score op de vraag.

	vraag	1	2	3	4	5	6	7	totaal	
leerling	max	3	5	8	3	4	3	4	30	
A		3	4	7	2	3	3	3	25	top
B		3	4	7	2	2	3	3	24	
C		3	5	8	2	1	2	3	24	
D		3	0	7	2	3	3	3	21	midden
E		3	0	8	1	3	2	4	21	
F		3	5	8	2	1	0	1	20	
G		3	1	3	2	2	3	4	18	laag
H		2	0	6	2	2	3	2	17	
I		3	0	8	1	1	3	1	17	
J		3	0	7	3	0	3	0	16	
p-waarde		0,97	0,38	0,86	0,63	0,45	0,83	0,60		
D-waarde		0,08	0,82	0,17	0,00	0,19	-0,11	0,31		

Figuur 1 – Voorbeeld scorematrix.

Zo krijg je een D-waarde die ligt tussen +1 en -1. Een D-waarde van +1 betekent dat de topscorers de vraag allemaal helemaal goed hebben, en de laagscorers allemaal helemaal fout. Zo'n vraag maakt een goed onderscheid tussen de goede en de zwakke leerlingen. Als daarentegen de D-waarde -1 is, hebben de topscorers het allemaal fout, en de laagscorers het allemaal goed. Dan is er iets mis met de vraag – bijvoorbeeld een addertje onder het gras dat jij niet zag bij het opstellen van de vraag, maar de goede leerlingen wél. In de praktijk ligt de D-waarde ergens tussen de uitersten. Een makkelijke vraag (hoge p-waarde) zal een D-waarde hebben in de buurt van 0 (ga na waarom).

### Uitdelen

Bedenk vooraf of de leerlingen de toets mogen houden. Als je de toets inneemt, kan hij verbeterd en in een volgend jaar nog eens gebruikt worden.

Deel de toetsen zelf uit (als je tenminste de namen van de leerlingen al kent) en zeg daarbij iets tegen elke leerling, maar zeker tegen de leerling die een onvoldoende terugkrijgt. Laat merken dat je begrijpt wat dat voor die leerling betekent. Zo'n opmerking kan perspectief bieden, bijvoorbeeld: "Je werkt er wel aan, maar toch heb je de meeste kennisvragen fout. Kijk de volgende keer goed naar wat je moet weten, dan gaat het misschien beter." Of: "Het is nog wel niet voldoende, maar het gaat naar mijn idee steeds beter. Gewoon volhouden, dan komt het wel goed." Bij sommige leerlingen zie je zelf ook geen perspectief. Bij die leerlingen kun je de gezamenlijkheid van het probleem benadrukken, bijvoorbeeld met een opmerking als: "Tja, het is helaas weer mis. Ik heb het heel goed nagekeken maar het is nou eenmaal niet meer dan dit. Ik vind het heel jammer. Laten we ervoor zorgen dat het in ieder geval niet slechter wordt." Er zijn natuurlijk ook leerlingen die niet werken. Die categorie krijgt dan bijvoorbeeld te

horen: “Het is een 4, en we weten allebei hoe dat komt. Misschien wordt het tijd dat je daar eens iets aan gaat doen.” Dan ligt het probleem bij de leerling. Dit is hét moment om irritante leerlingen met een onvoldoende uiterst vriendelijk te benaderen: uiteindelijk vang je meer vliegen met stroop dan met azijn. Dat kan bijvoorbeeld met: “Ik weet niet waarom je een onvoldoende hebt, maar het zou wel eens iets te maken kunnen hebben met het feit dat we steeds ruzie hebben. Ik wil daar best eens over praten. Als jij dat ook wilt moet je maar eens langskomen.” Dat doet hij overigens meestal niet. Dat je een leerling die het goed gedaan heeft moet prijzen hoeft geen betoog. Een simpel woord als “een 8, prima” is al voldoende.

Maak geen sarcastische of cynische opmerkingen bij het teruggeven van een toets, tenzij je precies weet waarom je dat bij die ene leerling doet. Soms werkt het, maar meestal wordt het je zeer kwalijk genomen, met name door andere leerlingen die het horen. Voor leerlingen valt dat gedrag van een leraar in de categorie ‘machtsmisbruik’ en ‘natrappen’. Grapjes met een goedge maakte toets mogen wel: met een somber gezicht naar een leerling toelopen, dan even wachten en vervolgens een 9 neerleggen wordt vaak zeer gewaardeerd.

### Bespreken

Je kunt de toets bespreken voor of na het uitdelen. Voor het uitdelen is de betrokkenheid van de leerlingen bij de bespreking vaak te gering, na het uitdelen te groot. In het laatste geval lukt het bijna nooit om de bespreking in alle rust en stilte te laten verlopen. Streef daar dan ook niet naar, maar zeg bijvoorbeeld wel: “Normaal wil ik dat het stil is als ik iets uitleg. Bij het bespreken van een toets ben ik daar wat minder streng in.” Het is verstandig alleen die vragen te bespreken die veel leerlingen fout hadden. Schrijf de normering op het bord.

Na de klassikale bespreking van de toets en de norm komen leerlingen die het niet eens zijn met hun beoordeling vaak allemaal tegelijk naar je toe. Ga daar dan niet op in en verwijst ze naar een later moment, als de klas aan het werk is. Stuur iedereen naar zijn plaats en laat ze één voor één bij je komen of loop zelf rond. Zorg ervoor dat iedereen die wat tegen je wil zeggen dat ook kan doen. De leerling hoeft het dan niet eens te zijn met je uiteindelijke beslissing, maar moet die wel kunnen aanvaarden. Als hij dat niet kan, verstoort dat je relatie met die leerling. Biedt de leerling dan bijvoorbeeld aan om het werk door een collega te laten nakijken, en stel hem voor dat jullie beiden zich bij diens beslissing zullen neerleggen.

Sommige leraren vragen de leerlingen terug te komen in de pauze. Dat verdient echter geen aanbeveling: leraar en leerlingen hebben een pauze nodig, en het klassenmanagement moet zo zijn dat die pauzes ook als zodanig gebruikt kunnen worden. Als er in de les geen tijd beschikbaar is of als er veel leerlingen met commentaar zijn kunnen ze ook hun commentaar op het blaadje schrijven, zodat de leraar er dan thuis eerst nog eens rustig naar kan kijken.

### Bronnen

Verhagen, P.A.J. (1993). *Leerdoelen en toetsen*. Utrecht: CDβ.

Poorthuis, H. (2015). *Theorietoets*. Utrecht: HU.