# **Formatieve toetsing**

‘*Toetsing om te leren, geen toetsing van het leren*.’

### **Inleiding**

Het verschil tussen summatieve en formatieve toetsing is het verschil tussen de gerichtheid op het product en het proces.

Summatieve toetsen sluiten het leerproces af en leiden tot een cijfer.

Formatieve toetsing geeft de student kennis over de vorderingen in een leerproces zodat deze student en zijn/ haar docent zicht hebben op de aanpak die tot een later succes zullen leiden. Formatieve toetsen leiden tot het bijstellen van gedrag en zijn daar ook op gericht. Het geeft de lerende onder andere de kans zijn werkwijze bij te stellen om zo betere resultaten te krijgen. Terwijl summatieve toetsen leiden tot een fait a compli: zo is het., niets aan te doen. Vaak is formatieve toetsing tot een soort tussentijdse summatieve toetsing verworden. De student krijgt een cijfer of aantekening maar heeft niet meer de mogelijkheid dingen anders te doen, zich te verbeteren.

### 

### **Feedback**

Naar feedback is veel onderzoek gedaan. Vooral John Hattie die een grote meta-analyse van 6000 verschillende onderzoeken deed, heeft hier veel aandacht aan gegeven. In zijn onderzoek laat hij zien dat Feedback een grote invloed heeft op het leren van mensen.

Een onderzoek van Hattie en Timperley laat zien dat feedback vooral effectief is als het gegeven wordt op de taak, het proces en de taak-regulatie. Feedback over de persoon (Jij bent goed) heeft veel minder effect. Zij stelden vast dat:

*“Leren kan worden verbeterd door de mate waarin studenten de uitdagende leerdoelen delen, zichzelf toetsen, evaluatie strategieën gebruiken, en technieken om fouten op te sporen ontwikkelen. Het vergroten van hun competentiegevoel om uitdagender taken aan te pakken, leidt tot meesterschap en begrip van leerstof.”*

### 

### **Formatieve toetsing (formatieve assessment)**

In 1998 laten Paul Black en Dylan William zien dat het gebruik van formatieve assessment leerresultaten met zo’n 25 percentielpunten kan laten groeien. Dit heeft de aandacht getrokken van velen. Veel uitgevers hebben zich hier op gestort en leveren formatieve toetsen. Hier moet je kritisch op zijn. Testuitgevers zagen hun kans schoon en herbenoemden hun formele summatieve toetsen als formatieve toetsen. Maar formatieve toetsing werkt het sterkst als het dicht op het leerproces staat en studenten daarover informatie geeft.

Formatieve toetsing is niet stevig gedefinieerd. Eigenlijk is vooral de functie ervan gedefinieerd: het geeft informatie die weer teruggezet in het systeem mensen helpt betere resultaten te hebben in het vervolg van het proces. Het is eerder toetsing om te leren dan toetsing van het leren.

### 

### **Heldere doelen en voortgang in leren**

Hoewel er vaak heldere einddoelen worden geformuleerd is het onbekend via welke stappen deze doelen moeten worden bereikt. De referentieniveaus zijn hier een aardig voorbeeld van. Het eindniveau is helder maar wat zijn de logische tussenstappen. We spreken dan over tussendoelen. Rond het stellen van doelen is veel onderzoek gedaan. Hieruit kunnen een paar conclusies worden getrokken:

* Het stellen van doelen heeft een duidelijk effect op leerresultaten;
* Specifieke doelen hebben meer opbrengst dan algemene doelen;
* Doelen moeten op het juiste niveau zijn geformuleerd voor een student om het beste effect te hebben. (zone van de naaste ontwikkeling)

### **De anatomie van formatieve toetsing**

In de jaren zestig wordt formatieve toetsing door Benjamin Bloom het eerst genoemd. Dit leidde echter nooit tot uitwerking op de klassenvloer.

Black en Williams definieerden in 1998 voor het eerst formatieve assessment *als alle activiteiten die leraren en/of studenten ondernemen om informatie te geven als feedback om leeractiviteiten te veranderen waarin zij actief zijn*. Hierin is dus vooral de functie van het formatieve assessment helder. In 2006 specificeert Popham formatieve toetsing als

* een proces en niet als een specifieke toets;
* wordt niet alleen gebruikt door leraren maar door leraren en studenten;
* het vindt plaats tijdens het onderwijs;
* het geeft toets-gebaseerde feedback aan studenten en leraren;
* het heeft de functie om leraren en studenten te helpen om aanpassingen te doen die studentprestaties rond een bepaald leerdoel verbeteren.

In 2008 stelt hij zijn definitie bij als “*een* ***gepland*** *proces waarin leraren en studenten toets-gebaseerd bewijs gebruiken om aan te passen wat ze op dit moment doen.”*

Het verschil tussen formatieve en summatieve toetsing is niet de vorm van de toets maar het moment waarop: als een kok de soep proeft is het formatief en als de klant dat doet is het summatief.

Eigenlijk gaat het dus niet om de toetsing maar om de gegevens die eruit voortkomen: we kunnen dus spreken van formatieve gegevens en summatieve gegevens.

De summatieve score is geen gemiddelde van de cijfers van de formatieve toetsen maar is het eindresultaat, het niveau waarop de student het leerproces afsluit.

Summatieve scores zijn niet zomaar af te leiden uit de formatieve scores. Het is niet zo dat de laatste formatieve score automatisch de summatieve score is. De score van een test bestaat altijd uit twee elementen: de werkelijke score + de fout score. Er zijn allerlei elementen die de student beïnvloeden tijdens een toets en die een foutmarge maken in de score. Dit geldt ook voor de formatieve toetsing in de klas. Daarom moeten meer elementen worden meegewogen als men tot een summatief oordeel komt.

### **Formatieve gegevens als bron voor gedragsverandering**

Voor de student en leraar zijn formatieve gegevens een aanleiding om hun gedrag en de effectiviteit ervan te overwegen.

Voor de student geldt dat hij herkent waarin hij zich moet verbeteren. Dus is vooral de kwalitatieve kant van de feedback interessant. De kwantiteit, ik heb er tien fout, is veel minder interessant dan de vaststelling dat ik op een verkeerde manier getallen optel.

Voor de leraar geldt dat deze vaststelt welke leerstof opnieuw moet worden aangeboden en misschien ook op welke manier.

Formatieve gegevens leiden dus vooral tot een kwalitatieve verbetering.

### **Goed volgen van resultaten**

Voor zowel studenten als leraren is het belangrijk om resultaten goed te begrijpen. Het gaat hierbij vooral om een kwalitatieve inschatting en normering. Wat betekent dit aantal fouten voor de volgende stap die de student en de leraar zouden moeten zetten in dit leerproces. Om dit goed te doen, moet vooral de toets zo worden gemaakt dat deze dit soort informatie biedt. Als een toets de student alleen bevraagd op 1 niveau weten docent en student alleen iets over het leren van gisteren en niet over dat van morgen. Het samenstellen van toetsen moet dus gebeuren op grond van het hele leerproces in een module of lessencyclus. Hieronder gaan we hier verder op in.

### **De noodzaak voor een nieuwe schaal**

In scholen worden veel verschillende schalen gebruikt. Als we naar de werkwijze kijken van leraren dan laat dit zien dat schalen vooral verschillend geïnterpreteerd worden. Met dezelfde toetsuitslag heb je bij leraar A een 8 en bij leraar B een 7. Er is dus behoefte aan een schaal die minder mogelijkheid biedt aan interpretatie, die eenduidiger is.

Bij de 100 punten schaal zie je vooral verschillen in de zwaarte die aan de verschillende soorten vragen worden gekoppeld.

Bij de rubrics (rubica terra = rode aarde, werd gebruikt om aan te geven als iets belangrijk was) zie je vaak verschillende invullingen van de niveaus. Dit kan bij de ene leraar veel moeilijker zijn dan bij de andere.

### 

### **Simpele en complexe inhouden vaststellen**

Om op een goede manier rubrics in te vullen moet je een aantal stappen zetten. De eerste stap is het vaststellen van het leerdoel, het doel van het onderwijs dat je gaat geven. Twee leerdoelen van een biologieleraar in een lessencyclus rond erfelijkheid zouden kunnen zijn:

1. Studenten kunnen de verschillen zien in de werkelijke wereld tussen niet-erfelijke en erfelijke eigenschappen;
2. Studenten kunnen voorspellingen doen met behulp van het kruisingsschema (vierkant van Punnett)

Nu maakt de leraar voor ieder doel een hoger en een lager doel.

Voor leerdoel 1 zou het er dan als volgt uit kunnen zien:

*Complexer*

Studenten zijn in staat om te beargumenteren hoe erfelijke en niet erfelijke eigenschappen elkaar beïnvloeden

*Richtdoel*

1. Studenten kunnen de verschillen zien in de werkelijke wereld tussen niet-erfelijke en erfelijke eigenschappen;

*Eenvoudiger*

Studenten kunnen de goede uitspraken herkennen over geïsoleerde voorbeelden van erfelijke en niet erfelijke eigenschappen.

Als deze trappen gemaakt zijn kan er een schaal worden gemaakt.

|  |  |
| --- | --- |
| Score 4.0 | Complexere inhoud |
| Score 3.0 | Richt doel |
| Score 2.0 | Eenvoudiger inhoud |
| Score 1.0 | Met hulp deels succesvol op doel 2.0 en doel 3.0 |
| Score 0.0 | Met hulp niet succesvol |

Natuurlijk worden deze doelen ook in de taal van de student omgezet als dat nodig is.

Deze schaal kan een onderverdeling krijgen die het mogelijk maakt beter tussen studenten te differentiëren. Hieronder daar het model van.

|  |  |
| --- | --- |
| Score 4.0 | Complexere inhoud |
| Score 3.5 | Met beheersing va 3.0 gedeeltelijke beheersing 4.0 |
| Score 3.0 | Richt doel |
| Score 2.5 | Geen fouten of weglatingen in 2.0 en gedeeltelijke beheersing van 3.0 |
| Score 2.0 | Eenvoudiger inhoud |
| Score 1.5 | Gedeeltelijk succes in 2.0 maar nog grote fouten en weglatingen in 3.0 |
| Score 1.0 | Met hulp gedeeltelijke beheersing van 2.0 en 3.0 |
| Score 0.5 | met hulp gedeeltelijk succes in 2.0, niet in 3.0 |
| Score 0.0 | Met hulp geen succes |

### **Toetsing bij de rubrics**

Hierbij past een vorm van toetsing die de verschillende niveaus meet:

|  |  |
| --- | --- |
| Sectie A | 2.0 – 10 mutliple-choice vragen die feitelijk van aard zijn maar wel van belang voor het onderwerp |
| Sectie B | 3.0 – vier open vragen die studenten vragen principes uit te leggen, voorbeelden te geven van generalisaties uit de les |
| Sectie C | 4.0 – twee open vragen die van studenten vraagt om gevolgtrekkingen te doen of toepassingen te bedenken die het klassenniveau overstijgen. |

Deze schaal in drie delen past binnen de logica van de drie niveaus van omgaan met de inhoud, zoals in Kunst en Wetenschap van het Lesgeven in de vragen 2, 3 en 4 is weergegeven. Lesgeven moet hier ook opgericht zijn dus.