

Toetsen om te leren

Paul van den Bos

Vrije Universiteit Amsterdam
Faculteit der Gedrags- en Bewegingswetenschappen,
Universitair Centrum voor Gedrag en Bewegen

Tekst t.b.v. didactische professionalisering van docenten in het hoger onderwijs

Versie september 2016

INHOUD

- 1.1. Inleiding. Summatieve en formatieve toetsing
- 1.2. Een onbalans tussen de inspanning van docent en student
- 1.3. De relatie tussen toetsing en het leergedrag van studenten; diepteleren en oppervlakteleren
- 1.4. 'Constructive alignment'. Afstemming van doelstellingen, leeractiviteiten en toetsing
- 1.5. Toetsen op hoog niveau volgens Dee Fink
- 1.6. Toetsmomenten en het bevorderen van zelfstudie
- 1.7. De rol van toetsing bij het komen tot zelfstandig leren
- 1.8. Samenvattend: meer halen uit toetsing
- 1.9. Referenties

Toetsen om te leren

“Als ze hun zesje binnen hebben, staan ze nauwelijks meer open voor feedback terwijl er nog zoveel te verbeteren valt” (docent HBO, voorjaar 2011)

1.1. Inleiding

Deze uitspraak verwijst naar twee functies die toetsing in het onderwijs kan hebben: het komen tot een cijfer enerzijds en het geven van inhoudelijke feedback anderzijds. Beide functies van toetsing zijn nodig in het onderwijs, maar hebben een verschillende betekenis voor het werk van docenten en voor het leergedrag van studenten.

Het komen tot een cijfer is de functie die doorgaans het eerst genoemd wordt. Toetsing is nodig om te weten of de studenten het geleerde hebben begrepen, of de stof voldoende is ‘blijven hangen’. Zo ja, dan kunnen de studiepunten worden toegekend en is het betreffende vak afgesloten. Zo nee, dan moet een hertentamen worden afgelegd of moet het betreffende vak nog eens worden overgedaan. Deze functie wordt de *summatieve* functie van toetsing genoemd, vooral gericht dus op selectie: mag je wel of niet door naar een volgend treetje in het onderwijsprogramma.

Dee Fink (2003) noemt deze functie van toetsing ‘*backward looking assessment*’, waarmee hij bedoelt dat deze vorm van toetsing in de eerste plaats is bedoeld om het leerproces met betrekking tot dit vak of deze module af te sluiten, en niet zozeer om inhoudelijke informatie te geven over het vervolg van het leerproces.

De tweede functie van toetsen heeft vooral betrekking op de inhoudelijke kant van toetsen. Het gaat daarbij niet in de eerste plaats om het geven van een cijfer, maar juist om het geven van informatie over wat de student precies goed en fout heeft gedaan, welk gedeelte van de stof reeds beheerst wordt en welke weg nog moet worden gegaan om alle stof onder de knie te krijgen. Dit wordt de *formatieve* of *diagnostische* functie van toetsen genoemd. Deze vorm van toetsing helpt studenten zelf te bepalen waar ze staan, hoe ze het tot nog toe hebben gedaan en wat er nog moet gebeuren. Dee Fink (2003) spreekt hier over ‘*forward looking assessment*’, waarmee hij bedoelt dat deze vorm van toetsing is gericht op wat er nog moet worden geleerd in de toekomst. De feedback die door docenten wordt gegeven op het werk van studenten (opdrachten, taken, tussentijdse toetsen, etc.) moet studenten helpen om gericht te studeren en geeft docenten zelf tussentijds informatie over waar de studenten staan, zodat het leerproces eventueel tussentijds bijgestuurd kan worden.

De eerste functie van toetsen – de summatieve functie – komt het meeste voor in het hoger onderwijs. Het is zelfs zo dat er in het HBO er een sterke tendens bestaat om het aantal toetsen en het aantal toetsmomenten terug te brengen tot één keer – aan het eind van elke onderwijsperiode. Daarnaast bestaat er een tendens tot standaardisering en digitalisering van toetsing.

1.2. Een onbalans tussen de inspanning van docent en student

In de titel van het boek van Dee Fink (2003): 'creating significant learning experiences' zit al wat kritiek verpakt op de sterke gerichtheid op summatieve toetsing. Deze vorm van toetsing zou gemakkelijk leiden tot uitstelgedrag van studenten, leergedrag dat is gericht op het afvinken van tentamens en niet zozeer op betekenisvol leren.

Het begrip 'zesjescultuur' komt in dit verband veelvuldig terug in de media de laatste jaren. En ook over het 'uitstelgedrag' van studenten zijn al vele columns en artikelen geschreven.

Dit verklaart ook de wat teleurgestelde reactie van de betreffende docent in het citaat aan het begin van deze tekst. Docenten steken vaak veel tijd en moeite in het nakijken van tentamens en werkstukken; voorzien ze soms uitgebreid van feedback; toch nemen studenten dit vaak voor kennisgeving aan en lijken er weinig mee te doen. Met andere woorden toetsing betekent in dit geval voor docenten een grote investering in tijd en moeite, terwijl het leerrendement gering is. Voor studenten betekent toetsing in dit geval vaak: een flinke investering van tijd en moeite voorafgaand aan de toets en dan ook nog vaak kórt voorafgaand aan de toets, en een opluchting na afloop van de toets, terwijl gelijk al vooruitgekeken wordt naar het volgende vak.

Deze situatie laat een onbalans zien tussen een grote investering door docenten aan de ene kant en een beeld van studenten die slechts minimaal bereid zijn zich in te zetten aan de andere kant.

Het is de vraag hoe deze onbalans tussen de investering van docenten en de inzet van studenten kan worden beïnvloed. Inhoud en organisatie van de toetsing spelen daarbij een grote rol

1.3. De relatie tussen toetsing en het leergedrag van studenten; diepteleren en oppervlakteleren

Over de invloed die toetsing heeft op het leergedrag van studenten is reeds veel gepubliceerd, bijvoorbeeld door Graham Gibbs. In zijn meest recente publicatie (Gibbs, 2010) trekt hij nog eens de belangrijkste conclusie uit eerder onderzoek: "What they found was that, unexpectedly, what influenced students most was not the teaching but the assessment (p.2) ".

"If you are under a lot of pressure then you will just concentrate on passing the course. I know that from bitter experience. One subject I wasn't very good at I tried to understand the subject and I failed the exam. When I re-took the exam I just concentrated on passing the exam. I got 96% and the guy couldn't understand why I failed the first time. I told him this time I just concentrated on passing the exam rather than understanding the subject. I still don't understand the subject so it defeated the object, in a way."

(Gibbs, 1992, p. 101)

Studenten stemmen hun leergedrag af op wat in de toets wordt gevraagd. Als de toets het mogelijk maakt om met oppervlakkig studeren toch nog een voldoende te halen zullen studenten gemakkelijk dit type studeren toe gaan passen. Al was het maar dat voor studenten niet alle vakken even belangrijk zijn en zij dus vaak zullen afwegen hoeveel tijd zij willen investeren in het halen van een tentamen, zelfs als ze weten dat oppervlakkig studeren op de lange

termijn niet erg zal renderen. Het citaat van een student oceanografie dat Gibbs hierboven aanhaalt is veelzeggend.

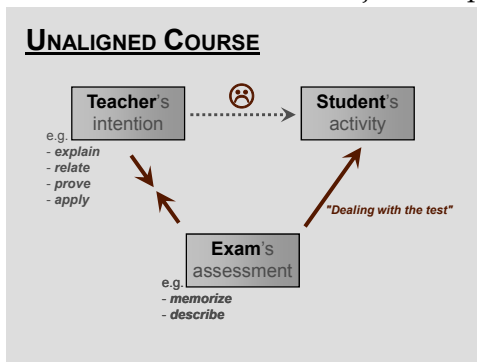
In het citaat van de student wordt verwezen naar twee soorten leren: leren gericht op het halen van het tentamen en leren gericht op 'understanding'. In de literatuur worden deze twee soorten meestal dieptelers en oppervlaktelers genoemd. Zie bijvoorbeeld Ramsden (1992).

Dieptelers is gericht op het verwerven van inzicht en op het ontwikkelen van leerstrategieën die studenten helpen meer zelfstandig te studeren ('duurzaam' studeren!). Bij dieptelers is men bewust met de leerstof bezig en probeert men deze te koppelen aan wat men al weet. Dit leidt tot een beter begrip, langduriger onthouden van het geleerde, zodanig ook dat het kan worden toegepast in wisselende dagelijkse situaties.

Oppervlaktelers daarentegen is gericht op het wat meer gedachteloos opnemen van kennis en informatie en deze opslaan als losstaande feiten. Oppervlaktelers bevordert niet het onthouden voor de langere termijn, bovendien leidt het tot kennis die niet flexibel is en niet makkelijk in een nieuwe context kan worden gebruikt.

1.4. 'Constructive alignment'. Afstemming van doelstellingen, leeractiviteiten en toetsing

De relatie tussen de toetsing en de leeractiviteiten van studenten staat ook centraal in de 'constructive alignment theorie' van Biggs (1996). In deze theorie wordt betoogd dat wanneer we in het onderwijs tot dieptelers willen komen, er ook doelstellingen op

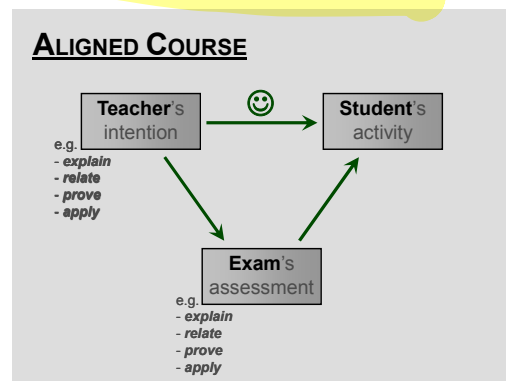


hoog niveau moeten worden nagestreefd, dat de leeractiviteiten daarop afgestemd moeten zijn en dat de toetsing dan ook op dit hoge niveau gericht moet zijn. Hij beschrijft het veelkomende voorbeeld van een 'unaligned', een niet afgestemde, cursus (zie afbeelding). De docent streeft wel doelstellingen na op hoog niveau – bijvoorbeeld het kunnen leggen van verbanden en de kennis kunnen toepassen – en benadrukt dit in de colleges. In de toetsing worden echter alleen vragen gesteld die een

beroep doen op het reproduceren van de kennis. Volgens Biggs zal dit ertoe leiden dat studenten zich in hun leergedrag al snel zullen gaan richten op het leren voor de toets en minder de noodzaak inzien van een actieve participatie in de colleges.

In een afgestemde cursus daarentegen (zie afbeelding) worden niet alleen doelen op hoog niveau nagestreefd, maar is ook de toetsing daarop afgestemd: er wordt meer gevraagd naar inzicht, het leggen van verbanden en de toepassing van de leerstof. Gedacht kan daarbij worden aan bijvoorbeeld praktijkcasussen. Studenten zullen zich in zo'n situatie eerder genoodzaakt voelen leeractiviteiten te ontplooiën die helpen om te komen tot een dieper inzicht in de stof, zeker wanneer dit wordt ondersteund door colleges en studieopdrachten die qua

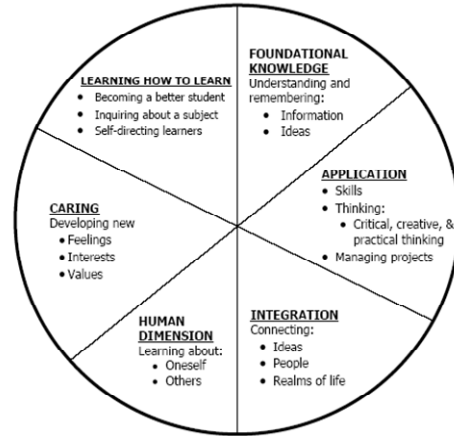
moelijkheidsgraad voorbereiden op het tentamen.



Blijkens het bovenstaande heeft toetsing dus een belangrijke sturende functie voor wat betreft het *niveau* waarop gestudeerd wordt door studenten.


1.5. Toetsen op 'hoog niveau'; betekenisvol leren volgens Dee Fink

De 'constructive alignment' theorie van Biggs(1996) is een theoretisch model voor het opbouwen van onderwijs. Dee Fink (2003) echter heeft dit model nader uitgewerkt tot een bruikbaar stappenplan voor het ontwerpen van cursussen in het hoger onderwijs. Ook dit model draait om de samenhang tussen leeractiviteiten, doelstellingen en toetsing en daarbij wordt ook veel nadruk gelegd op het bevorderen van onderwijs op een hoog niveau. Dee Fink spreekt over 'significant learning' - betekenisvol leren - en wijst daarbij ook op de 'duurzaamheid' van het leren.



Het gaat in het hoger onderwijs niet alleen om het leren van theorie voor het onderwijs van nu, maar de inzet moet breder zijn: het gaat ook om leerervaringen die bruikbaar zijn in het persoonlijk en het beroepsmatig leven en om het leren van vaardigheden die levenslang gebruikt kunnen worden. Bij het bepalen van de doelstellingen van een cursus moet de horizon dus wat verder gelegd worden dan alleen het einde van de cursus.

De taxonomie van betekenisvol leren van Dee Fink (2003) en de taxonomie voor cognitieve vaardigheden van Bloom

Cognitief	Foundational knowledge	Basisbegrippen beheersen	
	Application	Kennis kunnen toepassen	
	Integration	Het geleerde kunnen koppelen aan andere vakgebieden en aan ervaringen uit het dagelijkse leven	
Affectief	Human dimension	Kunnen aangeven wat de betekenis van het geleerde is voor het persoonlijk leven en voor de omgang met anderen	
	Caring	Het voelen van betrokkenheid bij en passie voor het onderwerp	
Regulatief	Learning how to learn	Inzicht krijgen in de eigen leerstrategieën en het eigen leergedrag	

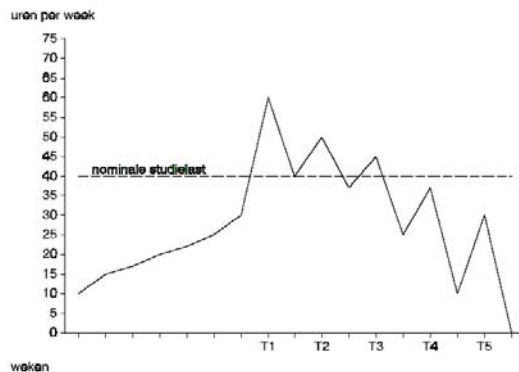
Dee Fink bepleit ervoor eerst na te denken over wat je graag zou willen dat de leerresultaten zijn na één of twee jaar en van daaruit terug te redeneren naar wat dat betekent voor de doelstellingen, toetsvormen en leeractiviteiten van de te ontwerpen cursus. Hij noemt dit 'backward design'.

Dee Fink onderscheidt zes elementen van significant leren waar het in iedere cursus om zou kunnen gaan (zie tabel). In dit overzicht is goed te zien dat betekenisvol leren een breed terrein beslaat. Het betreft zowel cognitieve vaardigheden (foundational knowledge, application en integration) als affectieve vaardigheden (human dimension, caring) en regulatieve vaardigheden (learning how to learn).

Voor wat betreft de cognitieve vaardigheden omvat de taxonomie van betekenisvol leren ook alle bekende elementen uit de taxonomie voor cognitieve vaardigheden van Bloom. Op p. 79 van zijn boek laat Dee Fink zien dat al onder application cognitieve vaardigheden als kritisch denken, analyseren, beoordelen en creëren worden geschaard, waarmee de hoogste trapjes uit de taxonomie van Bloom zijn afgedekt.

1.6. Toetsmomenten en het bevorderen van zelfstudie

Toetsing heeft ook grote invloed op de momenten waarop gestudeerd wordt door studenten. De grafiek in figuur 1 laat een beeld zien dat van toepassing is op veel opleidingen in het hoger onderwijs. De tentamens voor meerdere cursussen zijn



Figuur 1 Schematische weergave van de tijdsbesteding van studenten gedurende een semester met 5 cursussen parallel

geconcentreerd aan het einde van de onderwijsperiodes, er is geringe studieactiviteit van studenten in het begin van de periode en toenemende studielast naarmate het tentamen dichterbij komt. Nog los van het gebrek aan efficiency voor opleidingen is er een onderwijskundige reden waarom deze situatie niet wenselijk is. Gijsselaer en Schmidt (1993) maar ook time-on-task onderzoek van latere datum (Cotton, 1999) heeft laten zien dat niet in de eerste

plaats de hoeveelheid *contacttijd*, maar met name de hoeveelheid tijd die studenten aan de *zelfstudie* besteden en ook de *kwaliteit* van de zelfstudie uitmaakt voor hun studieresultaten.

In een recent onderzoek van Schmidt e.a. (2010) in het medisch onderwijs komt dit ook duidelijk naar voren: de hoeveelheid tijd die studenten kunnen besteden aan zelfstudie is het meest bepalend voor hun studieresultaten en studieduur. Een overmaat aan contacttijd - met name hoorcolleges - kan zelfs een contra-productief effect hebben.

In zijn studie met de titel 'de ware aard van uitstellen' stelt Vos (1989) dat uitstelgedrag uiteindelijk niet veel meer is dan een algemeen menselijk principe waaraan niet alleen studenten, maar ook docenten, ouders en managers onderhevig zijn: "kort voor een persoon verwacht dat er iets gaat gebeuren ontstaat de neiging om zich erop

voor te bereiden'. Het bestaan van een deadline verhoogt het gevoel van urgentie en zorgt ervoor dat men aan het werk gaat. Ook Vos wijst erop dat met name de zelfstudietijd effectief is voor het studeren en doet wat suggesties voor het op elkaar afstemmen van zelfstudie en contacttijd. Hij stelt dat tot op zekere hoogte meer contacturen meer zelfstudieactiviteit oproepen mits de contacturen gerelateerd zijn aan de toetsing en de zelfstudie ook daadwerkelijk ondersteunen.

Hij pleit er voor contacttijd meer volgens een voorbereiding-feedbackmodel in te vullen: zorg ervoor dat er regelmatig momenten zijn (hoor- of werkcolleges of andere contacttijd waarin opdrachten worden besproken, tentamens, proeftentamens, etc.) waarop studenten feedback krijgen op de activiteiten die ze hebben ontplooid en plan in het rooster ook daadwerkelijk tijd waarin studenten zich kunnen voorbereiden op deze feedbackmomenten. Aangezien studenten met name gaan studeren wanneer er een toets op komst is, verdient het aanbeveling te zoeken naar mogelijkheden om de toetsing te spreiden.

1.7. De rol van toetsing bij het komen tot zelfstandig leren

Spreiding van toetsing en een toename van het aantal momenten dat studenten feedback ontvangen op hun studieprestaties levert ook nog een derde belangrijk voordeel op. Het helpt de studenten het eigen leerproces beter te sturen. We vatten daarbij toetsing dus breed op. Het gaat dan niet alleen om de summatieve toetsing, want daarvan is de feedbackfunctie niet erg groot. Het gaat vooral om het gebruik van de formatieve toetsing. Alle momenten waarin studenten inhoudelijke informatie ontvangen over de geleverde prestaties - in de interactie met studenten, via kleine opdrachten in de colleges, via thuisopdrachten, presentaties, etc - al dan niet meetellend voor een cijfer, gelden als formatieve toetsmomenten. Deze vorm van toetsing helpt de student te bepalen waar hij staat in het leerproces en welke weg hij nog heeft te gaan. Bovendien kan toetsing aan studenten informatie verschaffen over de manier van studeren en de effectiviteit ervan.

Verder geeft toetsing informatie aan docenten over de vorderingen en de aanpak van studenten. Deze feedback en monitoring functie van toetsing, wordt gezien als het middel bij uitstek om de studenten te leren het eigen leerproces in de hand te nemen. Vermunt (1991) noemt dit het vergroten van de regulatieve vaardigheden van studenten, waarvan bekend is dat ze een grote voorspellende waarde hebben voor studiesucces. Uit de omvangrijke literatuur over feedback in het onderwijs (zie bijvoorbeeld Hattie & Timperley, 2007) komt een aantal centrale kenmerken van effectieve feedback naar voren.

Allereerst zijn er voor goede feedback heldere doelstellingen nodig en heldere criteria waaraan de leerprestaties moeten voldoen. Dee Fink (2003) bepleit het gebruik van rubrics: schalen waarin gedetailleerd is aangegeven wat het beheersingsniveau is voor een bepaalde studietaak. Rubrics maken peer-assessment en self-assessment mogelijk. Studenten kunnen - eventueel met hulp van medestudenten - het eigen werk beoordelen en zelf bepalen wat de volgende stap is in het leerproces.

Verder maakt het gebruik van rubrics mogelijk dat er frequenter getoetst kan worden en dat de feedback onmiddellijk beschikbaar is, wat ook bijdraagt aan de effectiviteit van de feedback. Tenslotte zijn de meeste auteurs het erover eens dat positieve feedback, die gericht is op de leerprestatie en niet op de persoon zelf, het meest oplevert.

Voor het geven van feedback hanteert Dee Fink (2003) het FIDeLity-principe: Frequent, Immediate, Discriminating (based on criteria and standards) en Lovingly (supportively).

1.8. Samenvattend: meer halen uit toetsing

Het begrip toetsing wordt hier nogmaals breed opgevat. Toetsing betreft niet alleen de summatieve toetsing aan het eind van een onderwijsperiode, maar ook alle vormen van tussentijdse feedback en formatieve toetsvormen.

Toetsing vervult dan een drietal krachtige functies:

- Aanzetten tot studeren op hoog niveau. De inhoud van de toetsing heeft direct invloed op de studieaanpak van studenten. Toetsen op hoog niveau leidt tot leeractiviteit op hoog niveau.
- Aanzetten tot zelfstudie. Toetsmomenten hebben invloed op de timing van het leren van studenten. Door contactmomenten, en toetsmomenten goed op elkaar af te stemmen kan invloed uitgeoefend worden op hoeveel tijd studenten aan de zelfstudie besteden en hoe effectief deze zelfstudie is.
- Aanzetten tot zelfstandig studeren. Heldere doelen en gedetailleerd uitgewerkte toetscriteria maken peer-assessment en self-assessment mogelijk, maken het mogelijk vaker te toetsen en stellen studenten daarmee in staat het leerproces meer in de eigen hand te nemen.

Het loont dus de moeite om goed na te denken over hoe toetsing in het eigen onderwijs kan worden ingezet teneinde zoveel mogelijk te kunnen profiteren van de kracht die toetsing kan hebben. In een *toetsplan* wordt dit denkproces weergegeven. Daarin wordt weergegeven welke doelstellingen op welk niveau men nastreeft. Hoe de toetsing wordt gespreid over de cursus en wat de kwaliteit van de toetsing is. We bespreken deze punten hieronder.

1.9. Referenties

Berkel, H. v. & Bax, A. (2006). *Toetsen in het Hoger Onderwijs*. Houten: Bohn Stafleu Van Loghum.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32, 347-364.

Brakel, G. van & Heijmen-Versteegen, I (2003); *Continue Zicht, Toetsing als spiegel voor zelfgestuurd leren en competentiegericht opleiden*; SKIF-publicatie. Te downloaden SKIF-site

Cotton, K. (1991). *Educational time factors*. Portland, OR: Portland, OR: NW Regional Educational Laboratory.

Dee Fink, L. (2003). *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. San Francisco: Jossey-Bass.

Gibbs, G. (2010). *Using assessment to support student learning*. Leeds Metropolitan University.

Gijselaer, W. H. & Schmidt, H. G. (1993). Investeren in instructietijd: spaarzaamheid loont de moeite. *Onderzoek van Onderwijs*, 22, 57-60.

Ramsden, P. (1992). *Learning to Teach in Higher Education*. London & New York: Routledge.

Schmidt, H. G., Cohen-Schtonus, J., van der Molen, H. T., Splinter, T. A. W., Bulte, J., Holdrinet, R. et al. (2010). Learning more by being taught less: a "time-for-selfstudy" theory explaining curricular effects on graduation rate and study duration. *Higher Education*, 60, 287-300.

Ten Dam, G. (1993). Zin in zelfstudie . OC Bulletin 33. Twente, Technische Universiteit.

Vermetten, Y (e.a); (2004); *Inzet van Assessment: Waarom, wat, hoe, wanneer en door wie?. Beslismodel voor een beargumenteerde keuze van assessmentvormen in onderwijs en opleiding*. OTEC. Open Universiteit. Heerlen.

Vos, P. & . (1998). Over de ware aard van uitstellen. *Tijdschrift voor Hoger Onderwijs*, 16, 259-274.