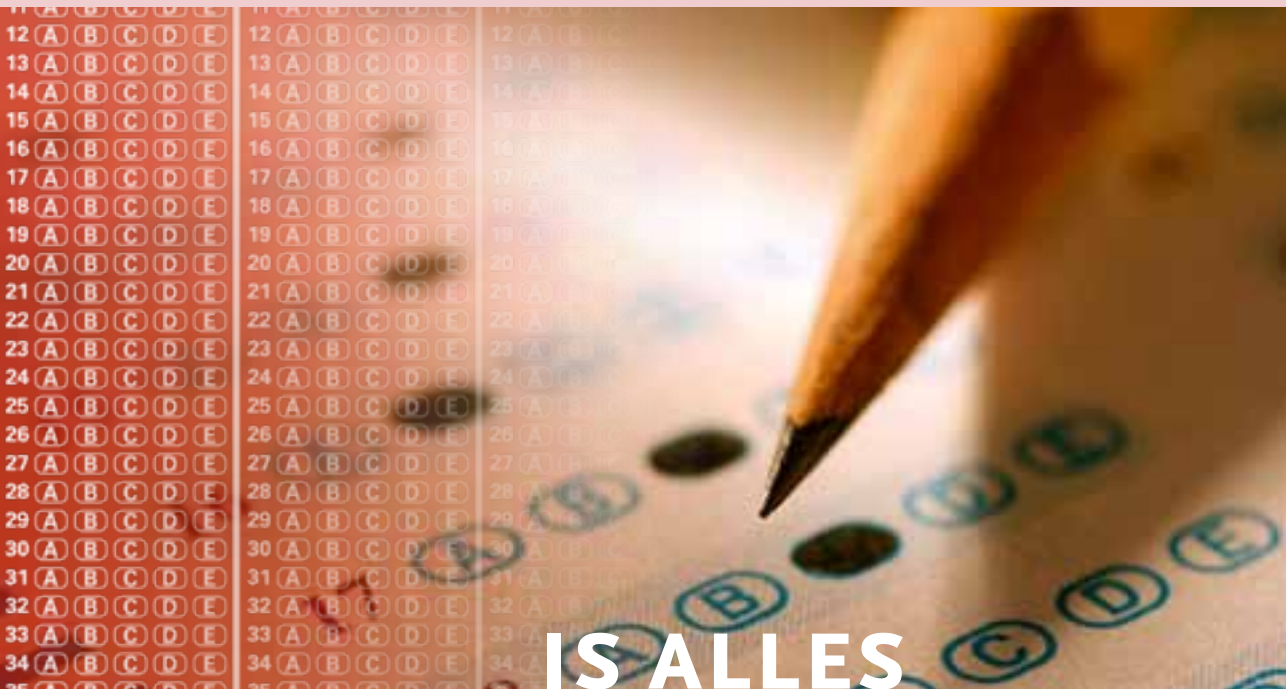


Jacqueline Bulterman & Bram de Muynck



IS ALLES VAN WAARDE MEETBAAR?

Toetsing en vorming
in het onderwijs

Is alles van waarde meetbaar? is geschreven in samenspraak met mbo- en hbo-docenten die studenten opleiden voor het primair onderwijs. Het is het resultaat van onderzoek dat plaatsvond binnen het project 'Voortgangstoetsing in de doorstroom mbo-hbo', een samenwerkingsproject van Driestar Hogeschool (hbo), Christelijke Hogeschool Ede (hbo) en het Hoornbeeck College (mbo).

Is alles van waarde meetbaar? kwam tot stand met steun van de stichting SURF. SURF initieert, regisseert en stimuleert ICT-vernieuwingen in het hoger onderwijs en onderzoek, door onder meer het financieren van projecten. Meer informatie over SURF is te vinden op de website (www.surf.nl).

Deze uitgave is beschikbaar onder de licentie Creative Commons Naamsvermelding 3.0 Nederland.



IS ALLES VAN WAARDE MEETBAAR?

IS ALLES VAN WAARDE MEETBAAR?

Toetsing en vorming in het onderwijs

Jacquelien Bulterman-Bos & Bram de Muynck

Een publicatie verzorgd door de lectoraten Docent en Talent (CHE)
en Christelijk Leraarschap (Driestar Hogeschool)



Buijten & Schipperheijn *Motief* – Amsterdam

Dit project is tot stand gekomen met steun van Stichting SURF, de organisatie die ICT vernieuwingen in het hoger onderwijs en onderzoek initieert, regisseert en stimuleert door onder meer het financieren van projecten. Meer informatie over SURF is te vinden op de website (www.surf.nl).

This project was made possible by the support of SURF, the higher education and research partnership organisation for Information and Communications Technology (ICT). For more information about SURF, please visit www.surf.nl.

In SURF werken universiteiten, hogescholen en onderzoeksinstituten samen aan grensverleggende ICT-innovaties. Hierdoor kan het hoger onderwijs en onderzoek optimaal gebruik maken van de mogelijkheden van ICT om zo de kwaliteit van onderwijs en onderzoek te verbeteren. SURF bestaat uit drie organisaties: SURFfoundation, SURFnet en SURFdiensten.



Deze uitgave is beschikbaar onder de licentie Creative Commons Naamsvermelding 3.0 Nederland.

© 2014 Buijten & Schipperheijn *Motief*
Reeks DIXIT, deel 8

Voor zover de afbeeldingen in deze uitgave auteursrechtelijk beschermd zijn, is contact gezocht met de rechthebbenden. Mocht iemand menen rechten te kunnen doen gelden op een van de in deze uitgave opgenomen afbeeldingen, dan gelieve hij/zij contact op te nemen met de uitgevers: Buijten & Schipperheijn te Amsterdam (rechten@buijten.nl).

Ontwerp omslag: Aperta, Jan Johan ter Poorten
ISBN 978-90-5881-787-7

Inhoud

Woord vooraf	7
Hoofdstuk 1	
Naar een evenwichtig toetsbeleid	9
Hoofdstuk 2	
Twee paradigma's	20
Hoofdstuk 3	
Toetsen vanuit het instrumentele paradigma	35
Hoofdstuk 4	
Van instrumentele toetsing naar assessment	51
Hoofdstuk 5	
Slingeren tussen instrumenteel en vormend paradigma	68
Hoofdstuk 6	
Normatieve professionalisering	83
Hoofdstuk 7	
Wereldbeelden, maakbaarheid en toetsing	102
Hoofdstuk 8	
Het onmeetbare in het toetsplan	116
Hoofdstuk 9	
Doel en middelen	133
Bijlage	137
Literatuur	138
Personalia	143

Woord vooraf

Dit boekje is een handreiking voor visievorming op toetsing in het licht van de bredere taak van het onderwijs. Met gespreksvragen en een scala aan casussen, ontleend aan allerlei onderwijspraktijken, maken we visievorming op toetsing tot een praktische aangelegenheid. In eerste instantie richten we ons op docenten in hbo en mbo, maar ook op docenten in het vo en het po. In tweede instantie is het boekje gericht op beleidsmedewerkers, leidinggevers, opleidingscoördinatoren, directeuren en ambtenaren, politici, eigenlijk op iedereen die in het onderwijs verantwoordelijkheid draagt.

De tekst is te gebruiken bij scholings- en bezinningsmomenten over toetsing. De hoofdstukken kunnen het beste in samenhang met elkaar gelezen worden. Maar het is ook mogelijk de afzonderlijke hoofdstukken voor gesprekken in teams te gebruiken.

De inhoud is ontwikkeld binnen het project 'Voortgangstoetsing in de doorstroom mbo-hbo'. In dit project hebben een aantal collega's van het Hoornbeeck College voor (reformatorisch) mbo, met locaties in Amersfoort, Apeldoorn, Goed, Kampen en Rotterdam, Driestar Hogeschool in Gouda en Christelijke Hogeschool Ede, digitale toetsen ontwikkeld. Tegelijkertijd met de toetsen is een doorlopende leerlijn rekenen, taal en Engels gemaakt die begint bij de opleiding tot onderwijsassistent in het mbo en eindigt als leraar basisonderwijs in het hbo. Dit pakket van leerlijnen en digitale toetsen biedt mbo-studenten oefenmogelijkheden. Zo kunnen ze hun kansen vergroten om als onderwijsassistent door te stromen naar de pabo en de opleiding tot leraar basisonderwijs succesvol af te ronden.

Aan ons als lectoren in het hbo is gevraagd een bezinnende tekst te schrijven voor docenten. Het werkplan bevatte de volgende doelstellingen. (a) De publicatie is erop gericht om docenten te helpen zich te verhouden tot toetsing en toetsgestuurd leren: waarom welke toetsing in te zetten op welk moment en bij welke lesinhoud; (b) de publicatie maakt duidelijk wat het doel en de rol van toetsgestuurd leren is binnen een beroepsopleiding; (c) de publicatie maakt duidelijk hoe toetsgestuurd leren past in een gewenste studiecultuur op mbo en hbo.

Het project vond plaats terwijl toetsing in het hbo sowieso al hoog op de agenda stond. Bezinning op toetsen kan alleen zinvol zijn als je hier de doelstelling van mbo en hbo bij betreft: beroepsvorming. Wij zien beroepsvorming als

een voor ieder uniek proces waarin jongeren hun talenten verkennen en ontwikkelen. Vorming is zo belangrijk dat we niet met docenten kunnen spreken over 'visie op toetsen' zonder het te hebben over de uniciteit van ontwikkeling en de idealen die docenten daarbij hebben. Zo is dit een boekje geworden dat niet alleen gaat over (digitale) toetsing, maar ook over vorming.

We spreken onze dank uit aan de deelnemers die aan het project digitale toetsing en door middel van interviews aan dit boek hebben meegewerkt: Nico den Besten, Carla van Hengel, Anja van der Hoek, Johan de Jong, Janneke de Jong – Slagman, Johan Melse (Driestar Hogeschool); Jan Willem Chevalking, Susan van der Ree - Buiten, Henk Rietdijk, Jurriaan Steen, Mirjam Trapman, Adé de Vries, (CHE); Aart Aarnoutse, Adrie van Binsbergen, Hester Bos – Willemsen, Frans Jan Hertog, Linda Kusters, Harm Visser, Hanneke Zaayer – Groenewoud (Hoornbeeck College). Onze zeer gewaarde collega Johan de Jong (Driestar Hogeschool) is ons helaas plotseling overleden in de laatste fase van het project. De gesprekken waren leerzaam en inspirerend en brachten ons meer dan eens op nieuwe gedachten. Een aantal van de genoemde collega's heeft meegelezen, evenals Peter Blokhuis en Marieke de Leeuw (CHE), Piet Murre (Driestar Hogeschool) en Johan de Visser en Ad Blonk (Hoornbeeck College). Hun suggesties hebben het boek nog dichterbij de praktijk gebracht, waarvoor we hen erkentelijk zijn. Verder danken wij de projectleiding in de personen van Laura Boele-de Bruin en Kees Ritmeester voor de coördinatie, begeleiding, suggesties en geduld.

Hoofdstuk 1

Naar een evenwichtig toetsbeleid

Tijdens de colleges van Dagmar wordt altijd veel gelachen. Dagmar is een muzische persoonlijkheid die studenten met haar kwinkslagen op het verkeerde been zet om vervolgens via een paar andere kwinkslagen belangrijke noties onder de aandacht te brengen. Haar vak ontwikkelingspsychologie wordt niet in hapklare brokken afgeleverd. Ze wil de hbo-student juist tot nadenken prikkelen en dat is te merken. Als de colleges van Dagmar zijn afgelopen, stroomt de gang vol met leven alsof studenten nog nagenieten van wat ze hebben meegemaakt.

Plezier is belangrijk maar Dagmar daagt haar studenten flink uit. Met een half woord neemt ze geen genoegen. Ze eist dat studenten begrippen uit de literatuur op een juiste manier gebruiken. Tijdens de colleges wordt dit geoefend. Alle studenten maken in hun stage een kort portret van een kind dat halverwege de cursus wordt ingeleverd. Deze portretten vormen voor Dagmar een rijke bron van inspiratie die ze in haar onderwijs volop benut.

De opleiding vraagt van docenten om toetsplannen te schrijven. De achterliggende reden hiervoor zijn zorgen over het niveau van het hbo. Ook Dagmar moet exact aangegeven welke doelstellingen haar cursussen hebben en hoe objectief gemeten wordt of deze doelstellingen zijn gehaald.

Dagmar is de eerste die staat voor kwaliteit in het hbo. Toch heeft ze een ongemakkelijk gevoel. Past deze instrumentele benadering bij haar visie op onderwijs? Zij wil met haar onderwijs de talenten van studenten ontwikkelen. Dat is iets anders dan hen programmeren volgens een vooraf geplande route. Onderwijs betekent voor haar vorming. Kritisch leren denken staat in al haar colleges centraal. Dagmar weet zeker dat ze studenten daarbij verder helpt, maar hoe kan ze dit hard maken? Inderdaad is kennis van begrippen voor haar

ook belangrijk maar ze voelt weinig voor onderwijs dat hapklare brokken toetst en nauwelijks aansluit bij de praktijk.

Dagmar stelt de directie voor een studiemiddag aan het onderwerp toetsbeleid te besteden. Het motto voor deze studiemiddag is de retorische vraag: 'Is alles van waarde meetbaar?'

De middag stelt haar voor verrassingen. Dagmar krijgt bijval maar een deel van de docenten is het pertinent niet met haar eens. Ze stellen dat leren handelen in de praktijk inderdaad belangrijk is maar dat hbo'ers daarbij vooral ook in staat moeten zijn om abstract te denken. Voor havisten, mbo'ers en zelfs voor vwo'ers betekent dat een enorme sprong in hun intellectuele ontwikkeling. Ze moeten leren om grotere hoeveelheden stof te overzien. Als die niet getoetst wordt, lezen studenten de literatuur nauwelijks. Abstract leren denken komt dan ook onvoldoende uit de verf.

Na afloop spreekt Dagmar de middag na met de directie. Het was een zinnige discussie maar er waren twee kampen. Verder dan een of-of houding waren ze niet gekomen. Nóg niet...

Aandacht voor toetsen

Toetsen is een belangrijk onderwerp van discussie die in alle geledingen van het onderwijsbestel speelt. Veel instellingen zijn bezig met het verbeteren van het toetsbeleid. Op diverse plaatsen worden gestandaardiseerde toetsen geïntroduceerd. Terwijl dit gebeurt, verschijnt er een rapport dat wijst op de smalle kijk op onderwijskwaliteit (Onderwijsraad, 2013) en ontstaan er collectieven die ervoor pleiten om toetsen af te schaffen. Hoewel de situaties in het hbo, mbo, het voortgezet onderwijs en het basisonderwijs drastisch van elkaar verschillen, spelen overal vergelijkbare thema's. Daarbij ontstaan kampen van voor- en tegenstanders. Ook het onderwijsbeleid slingert door de jaren heen van de ene kant naar de andere. Docenten die al wat langer meedraaien, verzuchten dat er maar weinig nieuws is onder de zon.

In dit boek proberen we niet met de laatste mode mee te doen. Wel willen we de vragen van vandaag serieus nemen. Toetsen zijn onderwerp van gesprek vanwege de kwaliteit van het onderwijs. De lat moet hoger, zo wordt gesteld. Hier worden ook bedenkingen tegen geuit. Zijn toetsen adequate

middelen om de kwaliteit te verbeteren? Gaat het in het onderwijs niet om méér dan kennis maar ook om vorming van jonge mensen in het licht van hogere waarden?

De bovenstaande casus betreft een bekwame en betrokken professional. Haar tegenstanders zijn dat echter ook. Ze verschillen fundamenteel van mening. Toch moet het te vormen toetsbeleid door het hele team gedragen worden. Dit boek schrijft niet voor hoe het ideale toetsbeleid eruit ziet maar nodigt professionals uit om hier zelf invulling aan te geven. Al dan niet in gesprek met elkaar. Met praktijkvoorbeelden, achtergrondinformatie, resultaten van onderzoek, discussievragen en een zelftoets dagen we de lezer uit om de eigen visie te verhelderen en te herijken.

Visie op toetsing hangt samen met visie op onderwijs. In dit hoofdstuk gaan we in op de functies van onderwijs. Vervolgens geven we eerst een omschrijving van basisbegrippen over toetsing en evaluatie, en kijken we welke plaats toetsing heeft in de verschillende sectoren van het onderwijs. Dan volgt de zelftest waarmee de lezers hun onderwijsvisie kunnen verhelderen. Ten slotte volgt een korte schets van wat de lezers van de rest van dit boek kunnen verwachten.

Functies van onderwijs

Om grip te krijgen op de plaats die toetsing en evaluatie in het onderwijs verdienen, is het goed om stil te staan bij de functies die onderwijs vervult. We volgen daarvoor een indeling van Biesta (2012) en de omschrijvingen die hij geeft: de kwalificatiefunctie, de socialisatiefunctie en persoonsvorming. Bij de kwalificatiefunctie gaat het om het aantonen van geschiktheid: kinderen en jongeren worden ‘gekwalificeerd’ om een vervolgopleiding te gaan volgen of om een bepaald beroep te gaan beoefenen. De socialisatiefunctie betreft het ingroeien in bestaande culturele tradities en praktijken. Leerlingen en studenten worden op school ingevoegd in een bestaande orde waarin bepaalde normen, waarden en gebruiken vanzelf spreken. Bij persoonsvorming gaat het niet om de omgeving maar om wat leerlingen of studenten zelf meebrengen en hoe zij daarin groeien. Volgens Biesta staan deze drie functies altijd in een zekere verhouding tot elkaar. Ze kunnen elkaar in evenwicht houden en tegelijk kan de ene functie de andere ook misvormen. Wanneer de kwalificatiefunctie en de socialisatiefunctie heel veel gewicht krijgen, kan dit bijvoorbeeld ten koste gaan van de persoonsvorming. Wanneer de nadruk op persoonsvorming wordt gelegd, kan dit

weer ten koste gaan van de kwalificatiefunctie. Verderop zullen we op deze functies terugkomen.

Begrippen

Dit boek gaat niet alleen over toetsing met meer keuzetoetsen maar over allerlei manieren van evaluatie waarbij informatie verzameld wordt over de vorderingen van studenten: observaties, mondelinge of schriftelijke overhoringen, toetsen, praktijkopdrachten, klassieke examens, opstellen, werkstukken, (mondelinge) boekbesprekingen, portfolio's etc.

Summatieve evaluatie vindt plaats aan het eind van (een deel van) het leerproces en telt mee voor de beslissing die genomen wordt over het vervolgtraject: mogen studenten door, krijgen ze een diploma?

Formatieve evaluatie gebeurt terwijl het leerproces nog aan de gang is en is bedoeld om het in de gewenste richting bij te sturen of te ondersteunen (Dochy & Janssens, 2003).

Het woord *evaluatie* duidt dus op dit brede scala van middelen om de vorderingen van studenten vast te stellen, terwijl toetsen één soort middelen zijn. Als we spreken van *toetsing*, dan duiden we op een evaluatievorm met vragen waarbij de antwoorden van tevoren vast staan. Er kan sprake zijn van (digitale) toetsing door middel van meerkeuzetoetsen of van toetsen met vragen waarbij een uitgebreid antwoordmodel is ontwikkeld. Als de antwoorden van tevoren vast liggen, kan min of meer objectief gemeten worden hoeveel antwoorden er goed zijn. Bij toetsen horen dus geobjectiverde normen. Wij gebruiken in dit boek het woord toetsing als er gemeten wordt volgens zo eenduidig mogelijke standaarden. Toetsen kunnen zowel formatief als summatief gebruikt worden. Een diagnostische toets is een voorbeeld van een toets die een formatieve functie vervult.

Competentieassessments zijn in het beroepsonderwijs belangrijke evaluatievormen. Het Engelse woord *assessment* is synoniem met 'evaluation' en betekent waardebeoordeling. Als we het woord *evaluatie* of *assessment* gebruiken, bedoelen we een beoordeling van de waarde van het werk van studenten die niet noodzakelijk door middel van toetsen plaatsvindt. Het woord 'oordeel' in 'beoordeling' geeft aan dat er meer nodig is dan anonieme instrumenten en geobjectiverde normen.



In het (middelbaar) beroepsonderwijs wordt tegenwoordig gesproken van toetsingskaders, toetsbeleid en toetsplannen. Het is onwaarschijnlijk dat deze toetsingskaders en dit toetsbeleid alleen maar toetsen bevatten met vragen waarbij de antwoorden min of meer vast staan. Hoe komt het dat hier het woord toetsing wordt gebruikt? Spreekt daaruit de wens om toch vooral streng genoeg te zijn? Hoe dan ook, wij gaan ervan uit dat het toetsbeleid en het toetsplan de kwaliteit van onderwijs moet dienen. In de loop van dit boek bepalen de lezeres zelf welke invulling zij willen geven aan deze gebieden.

Toetsen en evaluatie in de verschillende sectoren van het onderwijs

In vrijwel alle geledingen van het onderwijs staat toetsing op de agenda. Hieronder laten we zien wat er speelt in het po, het vo, het mbo en het hbo. Daarnaast spelen er ook zaken door alle geledingen van het onderwijsbestel heen. Om de prestaties van leerlingen op het gebied van taal en rekenen te verbeteren, heeft de overheid in 2010 referentieniveaus opgesteld. Hierin staat aangegeven welke taal- en rekenprestaties van leerlingen op een bepaald niveau verwacht mogen worden. Momenteel speelt de discussie over hoe deze taal- en rekenprestaties getoetst zullen worden.

Primair onderwijs (po)

Nederlandse kinderen maken gemiddeld 72 toetsen voordat ze de overgang maken naar het voortgezet onderwijs. Deels gaat het hier om toetsen waarmee de ontwikkeling van leerlingen bij rekenen en taal gevolgd wordt. Ze zijn landelijk genormeerd en staan los van de methodes waarmee op de scholen wordt gewerkt. De resultaten worden opgeslagen in het (geautomatiseerde) leerlingvolgsysteem. De individuele ontwikkeling van iedere leerling kan op die manier gemakkelijk in kaart worden gebracht. Ook kunnen gegevens van groepen uit de toetsresultaten worden afgeleid, zodat de leraar ook op het niveau van de groep zijn handelen kan bijstellen (zie hierover uitvoerig in De Muynck, Both & Visser-Vogel, 2013). De Cito-eindtoets was aanvankelijk bedoeld als een ondersteuning voor een goed advies over het voortgezet onderwijs. De leraar had een belangrijke stem in dit advies en de score kwam op de tweede plaats. Geleidelijk heeft de Cito-score steeds meer gezag gekregen. Methodegebonden toetsen die aansluiten bij het specifieke onderwijs, zijn bij de 72 toetsen nog niet meegerekend. Deze toetsen horen ook bij andere vakken dan rekenen en taal.

De toetsdruk in het basisonderwijs is toegenomen. Op veel scholen wordt flink geoefend voor de toets. De laatste jaren zijn oefenbureaus ontstaan waar kinderen door hun (betalende) ouders naartoe gestuurd worden om zo een beter advies voor het vervolgonderwijs te krijgen. De meningen over toetsen zijn verdeeld. Sommige scholen vinden dat toetsen leerkrachten handreikingen geven om hun onderwijs beter af te stemmen op wat leerlingen nodig hebben. Anderen wijzen erop dat de toetsen slechts een klein deel van de totale ontwikkeling volgen en daarmee onrecht doen aan leerlingen. In het basisonderwijs wordt opbrengstgericht gewerkt: van scholen wordt verwacht dat zij de ontwikkeling van leerlingen volgen, hun onderwijs afstemmen op wat leerlingen nodig hebben en zo de opbrengsten (de toetsresultaten) verhogen. In de begroting van 2013 heeft de overheid de doelstelling geformuleerd om de gemiddelde score van de Cito-eindtoets op de basisschool omhoog te laten gaan van 535,5 (in 2012) naar 537 in 2015. Tijdens het debat in de Tweede Kamer in september 2013 was hier verzet tegen en heeft de regering deze doelstelling geschrapt.

Voortgezet onderwijs (vo)

In het voortgezet onderwijs worden bij de meeste vakken schriftelijke overhoringen, proefwerken en repetities gegeven. Deze worden door iedere school en soms door individuele docenten, zelf gemaakt. Ook kent het vo praktische

opdrachten, profielwerkstukken, boekverslagen, verslagen van practica etc. De eindexamens worden voor de helft door de overheid gemaakt en genormeerd. In het programma van toetsing in afsluiting (PTA) staat precies hoe het school-examencijfer tot stand komt. De inspectie houdt nauw in de gaten of de resultaten van het schoolexamen in de pas lopen met de resultaten van het centraal schriftelijk examen. Omdat de evaluatie in het vo verschillend is voor iedere school, is er gepleit voor een landelijke diagnostische toetsing aan het einde van de onderbouw voor Nederlands, Engels en wiskunde/rekenen. Momenteel zijn deze toetsen niet verplicht maar sommige scholen nemen ze wel af. De overheid heeft samen met de vo-raad (een landelijke organisatie van vo-scholen) de ambitie geformuleerd om ieder jaar meer scholen opbrengstgericht te laten werken. In 2018 zou 90% van de scholen opbrengstgericht moeten werken en dus ook gebruik moeten maken van de daartoe bestemde toetsen.

Het middelbaar beroepsonderwijs (mbo)

In het mbo wordt competentiegericht gewerkt. De manier waarop beroeps-competenties geëvalueerd worden, is sterk in ontwikkeling. Aanvankelijk vond evaluatie vooral plaats door middel van portfolio's waarmee de student gaandeweg bewijzen verzamelde voor verworven competenties. Vaak betrof het opdrachten met steeds een afzonderlijke beoordeling. Deze waren vaak nogal theoretisch van aard. Dit is aan het verschuiven in de richting van het beoordelen van concrete handelingen in een concrete en complexe situatie. Er wordt dan gesproken over 'proeven van bekwaamheid'. Iedere mbo-opleiding heeft een kwalificatiedossier. Het werkveld heeft een groot aandeel in de beoordeling en in de vaststelling van het niveau. Toch blijft het mbo zelf de verantwoordelijkheid houden. Daarom is op dit moment voor de scholen een belangrijke rol weggelegd om de kwaliteit te waarborgen. De opdrachten moeten nauwkeurig geformuleerd worden en de beoordelaars moeten geschoold worden. In deze scholing wordt aandacht besteed aan hoe je werkt met de criteria die hiervoor zijn opgesteld en waar de cesuur ligt. Er wordt gestreefd naar een strikte scheiding tussen begeleiden en beoordeling. Al enkele jaren bestaat vrij algemeen het beeld dat de kwaliteit van mbo-diploma's onder de maat is. De inspectie ziet nauw toe op de kwaliteit van de beoordeling.

Het hoger beroepsonderwijs (hbo)

In het hbo bestaat een grote variëteit aan toetsen. Ook daar zijn bij iedere opleiding beroepscompetenties geformuleerd. De meeste opleidingen kennen een

mix van kennistoetsen en praktijkgerichte opdrachten die beoordeeld worden. De opleiding wordt afgesloten met een praktijkgericht onderzoek. Net als het mbo krijgt het hbo het verwijt dat diploma's te gemakkelijk zouden worden toegekend. Ook daar wordt nagedacht over manieren om de kwalificatiestructuren te verbeteren. Hogescholen maken toetsplannen; tentaminering gaat gepaard met toetsmatrijzen. De betrouwbaarheid en de validiteit van toetsen is onderwerp van analyse. Speciale scholing van docenten wordt georganiseerd. De overheid heeft ook een meer overkoepelende manier om de kwaliteit te belonen. De minister maakt met instellingen zogenaamde prestatieafspraken over hoe zij de kwaliteit kunnen verbeteren. Worden die afspraken niet nagekomen, dan wordt de subsidie gekort. Het idee is dat de instellingen de eigen verantwoordelijkheid serieus nemen en doelgericht streven naar betere diploma's, betere doorstroom van studenten, betere scholing van docenten enzovoort.

De lerarenopleiding

Docenten spelen een belangrijke rol als het gaat om de kwaliteit van onderwijs. Op de kwaliteit van de lerarenopleidingen wordt extra nauw toegezien. De pabo's kennen een verplichte landelijke Cito-toets voor rekenen en taal. Voor veel vakken is een kennisbasis geformuleerd. Al deze kennisbasissen bij elkaar vormen teveel stof voor de opleiding. Er wordt nagedacht over manieren om het niveau van de havo- en mbo-studenten te verhogen, of aan toelatingstoetsen op het gebied van geschiedenis, aardrijkskunde, natuur en techniek voor studenten die hierin geen examen hebben gedaan. Op die manier zal de kennis en de vaardigheden van de aankomende docenten toenemen. Lerarenopleidingen verwachten dat daarmee ook de kwaliteit van het primair en voortgezet onderwijs zal verbeteren.

Zelftest onderwijsvisies

Visies op toetsing hangen nauw samen met visies op onderwijs. Het is onze bedoeling om reflectie op beide visies te bevorderen. Voordat we een indruk geven van de richting waar dit boek naartoe gaat, krijgen de lezers de kans om hun eigen visie op onderwijs te verhelderen met de onderstaande test (Kaldeway, De Heer & Bulterman-Bos, 2014). Achter in dit boek staat hoe de uitslag van de test berekend kan worden.

Vragenlijst onderwijsvisies

In de volgende lijst worden steeds twee stellingen voorgelegd. De stellingen vormen geen tegenstellingen, ze sluiten elkaar niet uit. Aan u wordt gevraagd uw positie in dit spectrum aan te geven: als u moet kiezen, ligt uw prioriteit of affiniteit dan meer aan de ene of andere kant? Daarbij gaat het om uw 'ideale' visie op onderwijs, even ongeacht de externe eisen die eventueel aan het onderwijs gesteld worden. Denk ook niet aan een bepaald schooltype of een bepaalde opleidingsfase maar aan onderwijs in brede zin, van primair tot hoger onderwijs.

Waar positioneert u uzelf?



1. In het onderwijs is het van belang dat lerenden aan algemene eisen vanuit de samenleving leren voldoen.	0 0 0 0 0	In het onderwijs is het van belang de eigen talenten van de lerenden ruimte te geven.
2. Onderwijs moet erop gericht zijn dat lerenden plezier in leren ontwikkelen.	0 0 0 0 0	Onderwijs moet lerenden stimuleren een zo hoog mogelijk niveau te bereiken.
3. In het onderwijs is het van belang scherp te hebben wat lerenden zich precies aan kennis, inzicht en vaardigheden eigen moeten maken.	0 0 0 0 0	In het onderwijs is het van belang in te spelen op de eigen motivatie van lerenden om op hun eigen moment iets te leren.
4. Het is belangrijk bij het aanbieden van de leerstof rekening te houden met de specifieke, verschillende manieren waarop lerenden het beste leren.	0 0 0 0 0	In het onderwijs kun je het beste werken met beproefde vakmethodieken die lerenden naar het leerdoel brengen.
5. Het onderwijs moet in elk geval zorg dragen voor beheersing van de basisvaardigheden in geletterdheid (taal) en gecijferdheid (rekenen/wiskunde).	0 0 0 0 0	Het onderwijs moet evenredig aandacht hebben voor de communicatieve, praktische, creatieve en cognitieve ontwikkeling.
6. Het is van belang dat lerenden van zichzelf in de verschillende leergebieden sterke en zwakke punten leren ontdekken.	0 0 0 0 0	Het is van belang dat lerenden met regelmaat door toetsing worden geïnformeerd over het niveau van hun prestaties in de verschillende leergebieden.
7. Het is van belang dat lerenden bij het leren voor (onderdelen) van vakken ruimte krijgen voor keuzes die bij hen passen en voor eigen accenten.	0 0 0 0 0	Het is van belang dat lerenden voor alle verschillende (onderdelen van) vakken een bepaald basisniveau bereiken.

Het vervolg van dit boek

We begonnen dit hoofdstuk met Dagmar en de spanning die zij ervaart met toetsen. In de casus klinken twee paradigma's door: het instrumentele paradigma en het vormingsparadigma. In hoofdstuk 2 worden beide paradigma's toegelicht en wordt het spanningsveld tussen die twee duidelijk. Beide paradigma's lijken iets waardevols te vertegenwoordigen en tegelijk gaat het om tegenpolen.

Hoofdstuk 3 geeft vanuit het instrumentele paradigma handreikingen om de kwaliteit van toetsing en evaluatie te verbeteren.

In hoofdstuk 4 komt een andere kant aan de orde: is het inderdaad waar dat de kwaliteit van onderwijs verbetert, of wordt deze ook versmald? Bijna zouden we vergeten dat het onderwijs in de vorige eeuw ook een periode kende waarin de instrumentele benadering centraal stond. Vanuit de kritiek hierop ontstond echter aan het eind van de vorige eeuw een reeks onderwijsvernieuwingen, die ontleend waren aan het vormingsparadigma. We geven concrete handreikingen om leerprocessen te evalueren vanuit het vormingsparadigma.

In hoofdstuk 5 komen ook de bezwaren tegen het vormingsparadigma aan de orde die duidelijk werden toen de commissie Dijsselbloem de onderwijsvernieuwingen onderzocht. Als gevolg van deze bezwaren hebben veel onderwijsinstellingen op dit moment weer te maken met een instrumentele golf. De contouren van een tegenreactie zijn echter al volop zichtbaar.

We prikkelen de lezer met wetenschappelijk onderzoek en praktijkvoorbeelden die laten zien dat het onderwijs vanuit beide paradigma's verbeterd kan worden. Wat zegt dit over de aard van het onderwijs? Zo komen we in hoofdstuk 6 op het begrip normatieve professionalisering. Daarmee wordt het cruciale belang van docenten duidelijk. Ze staan voor de uitdaging om voortdurend met paradoxen om te gaan. Kwaliteit van onderwijs is daarom niet in systemen te vangen. Toch wordt dit steeds weer geprobeerd, met een voortdurende slingerbeweging als gevolg.

In hoofdstuk 7 stellen we de vraag waar deze slingerbeweging vandaan komt. We concluderen dat deze samenhangt met een beeld van de werkelijkheid dat afkomstig is uit de verlichting. Als gevolg van een bepaalde vorm van objectiviteitsdenken is praktische professionaliteit, ook die van docenten, naar de achtergrond verdrongen ten gunste van rationele, bewijsbare kennis. In de wetenschap staat het objectiviteitsdenken echter ook onder kritiek.

Vanuit deze kritiek presenteren we een alternatief werkelijkheidsmodel. Wat gemeten kan worden, krijgt een volwaardige plaats in dit werkelijkheidsmodel. Tegelijkertijd wordt ook een volwaardige plaats ingeruimd voor wat niet gemeten kan worden maar wat toch voor iedere student belangrijk is.

Het nieuwe model benadrukt dat beoordeling van het werk van studenten om persoonlijke betrokkenheid en vakmanschap van docenten vraagt. In hoofdstuk 8 geven we vanuit dit model een praktische handreiking.

Hoofdstuk 9 staat in het teken van de vraag naar doel en middelen. Zijn toetsen middelen om het leerproces te bevorderen of is het leerproces een middel om de toetsresultaten te bevorderen? De lezer mag deze vraag zelf beantwoorden.

Hoofdstuk 2

Twee paradigma's

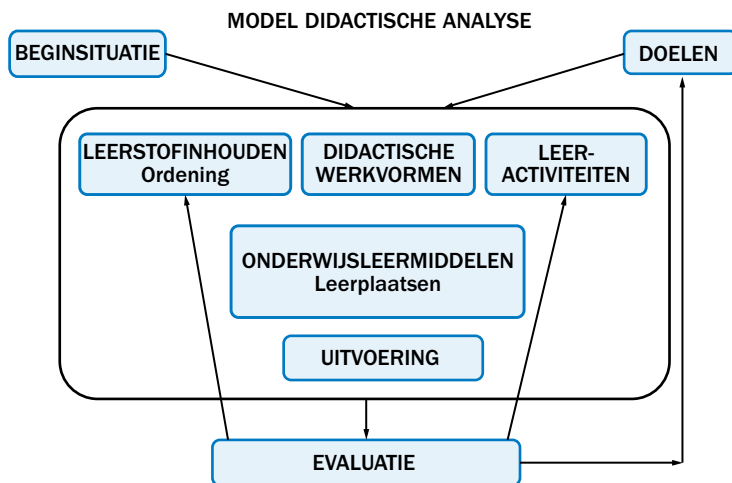
De begrippen 'toetsen' en 'evalueren' zijn eenvoudig te omschrijven als een activiteit waarbij nagegaan wordt of onderwijsdoelen zijn bereikt. Welke middelen je daarvoor kiest wordt weer bepaald door hoe je over onderwijs denkt. In dit hoofdstuk leggen we twee manieren van denken uit, die sturend zijn voor onderwijs en voor bijbehorende manieren van toetsen en evalueren. De eerste noemen we de instrumentele manier van denken, de tweede de vormingsbenadering. We spreken over paradigma's omdat ze elk voor zich staan voor een samenhangend pakket aan ideeën. In de loop van dit boek zal blijken dat het onderwijsbeleid dan weer eens de ene, dan weer eens de andere kant op gaat, een slingerbeweging die voortduurt. Ons streven is om de lezer instrumenten aan te reiken om zélf verder te komen in de verantwoording van de vraag wat goed onderwijs is en hoe je leerprestaties vaststelt. Onze eigen mening blijft zoveel mogelijk op de achtergrond. Na een verkenning van het instrumentele paradigma en een verkenning van het vormingsparadigma komen we tot een vergelijking waar de verschillen tussen de twee paradigma's op een rijtje komen te staan.

De instrumentele benadering

De instrumentele manier van denken komt op verschillende manieren in het onderwijs naar voren. Kort aangeduid: we doelen op *de benadering die processen van leren en ontwikkeling probeert te vatten in termen van objectiveerbare en controleerbare doelstellingen*. We beginnen met een korte uitleg over didactische analyse, een basismodel waarin een aantal didactische handelingen zoals doelstellingen, instructie en toetsing met elkaar in verband worden gebracht (Van Gelder, Oudkerk Pool, Peters & Sixma, 1973, De Corte, Geerlings, Lagerweij, Peters & VandenBerghe, 1981, Tielemans, 2001, Mylle, 2011). Vervolgens leggen we uit hoe de instrumentele benadering gebruikt wordt om het onderwijs te verbeteren en waarom objectieve metingen daarbij van belang worden geacht.

Didactische analyse

Het (beroeps)onderwijs heeft brede doelstellingen die een beroep doen op een totaalpakket aan kennis en vaardigheden. Het model van de didactische analyse deelt die doelen op in kleine onderdelen, die afzonderlijk getoetst kunnen worden. Een rekenvaardigheid als 'vermenigvuldigen' kan bijvoorbeeld opgedeeld worden in een aantal deelvaardigheden. Neem de vermenigvuldiging 34×22 . Studenten moeten zich kunnen voorstellen wat vermenigvuldigen werkelijk betekent. Om dat te kunnen, moeten ze begrijpen dat ze de hoeveelheid '22' twee keer kunnen nemen en beide hoeveelheden bij elkaar op kunnen tellen, maar ook bijvoorbeeld drie keer of 34 keer. Vermenigvuldigen is dus hetzelfde als herhaald optellen. Wie begrijpt wat vermenigvuldigen is, begrijpt ook wat optellen is. Optellen is dus een deelvaardigheid van vermenigvuldigen. Zo zijn allerlei leertaken in deeltaken op te splitsen. Wie het economisch systeem in de VS wil vergelijken met dat in Nederland, moet eerst iets weten van economische systemen. Het idee van de didactische analyse is dat je iedere leertaak kunt opsplitsen in aan te leren en toetsbare vaardigheden.



Afbeelding 2.1. Model didactische analyse

Naar: Gelder, L. van, Oudkerk Pool, T., Peters, J., & Sixma, 1973.

Voor een recente bewerking zie Mylle, 2011, p. 40

In het model van didactische analyse is het van belang dat de doelstellingen in verband gebracht worden met de beginsituatie (zie afbeelding). Het verschil tussen de beginsituatie en de doelstellingen leidt vervolgens tot een les of tot een serie lessen waarin didactische principes worden toegepast. Omdat men

goed over doelstellingen heeft nagedacht en zicht heeft op de kloof tussen de beginsituatie en de doelstellingen, kan de instructie zorgvuldig uitgekiend worden. Vervolgens neemt het model didactische analyse aan dat door middel van een toets objectief vastgesteld kan worden of de (deel)doelstellingen zijn gehaald. Dit gebeurt met behulp van (meet)instrumenten, daarom spreken we van een instrumentele benadering (Thijs & Van den Akker, 2009, p. 17-18). Het woord ‘objectief’ geeft aan dat er gestreefd wordt naar het scheiden van feiten van meningen. De toets beperkt zich, zover mogelijk, tot wat als ‘waar’ wordt beschouwd of tot wat als ‘juist’ wordt beschouwd. Over spelling bestaan bijvoorbeeld afspraken. Een student wordt geacht de regels te kennen en deze correct toe te passen. Dat er discussie over die afspraken mogelijk is – om de zoveel tijd worden spellingregels aangepast – wordt in het instrumentele paradigma niet benadrukt.

Didactische analyse kan worden toegepast van kleuteronderwijs tot en met universitair onderwijs en van voorbereidend beroepsonderwijs tot en met hoger beroepsonderwijs. Het bekende model van ‘directe instructie’ is hier bijvoorbeeld van afgeleid. In een collegereeks over cognitieve gedragspsychologie kun je kennisdoelen zo opsplitsen dat je de betekenis van begrippen als stimulus en respons afzonderlijk kunt toetsen.

Gespreksvraag: Herkent u het model didactische analyse (zie afbeelding 2.1.) in uw eigen vakgebied of in de eigen lespraktijk? Kunt u een opdracht in uw eigen vakgebied analyseren? Welke voorkennis hebben studenten nodig om deze opdracht te kunnen doen?

Het uiteenrafelen van doelen, de didactische analyse, is een klus op zich (zie uitvoerig hierover Mylle, 2011, 55-76, Tielemans, 2011, 21 e.v.). De uiteen-gerafelde doelen vormen samen een leerlijn. Een belangrijk uitgangspunt van de instrumentele benadering is dat studenten zich lineair ontwikkelen. Informatie over de mate waarin doelstellingen gehaald zijn, dient tevens om de beginsituatie te bepalen ten behoeve van de volgende stap op de leerlijn. Zo komt een cyclus tot stand: de toetsing van een stap die gemaakt is, levert weer informatie op voor de volgende stap. Iedereen doorloopt dezelfde leerlijn, zo is de aanname. Daarmee beschouwt de instrumentele benadering de ontwikkeling van studenten als voorspelbaar.

Tegelijkertijd is ook duidelijk dat niet alle studenten hetzelfde aankunnen. De toetsen bieden objectieve informatie om te differentiëren binnen de groep, of

te selecteren naar verschillende schooltypes. Er is binnen dit paradigma dus wel aandacht voor niveaoverschillen, maar minder voor de eigenheid van de studenten.

Het instrumentele paradigma draait om eendimensionale doelstellingen. Niet de gehele persoon van de student staat daarbij centraal evenmin als de manier waarop de student de opdrachten persoonlijk invult. Het gaat alleen om de prestatie van studenten op een duidelijk omschreven taak waarbij van tevoren in grote mate vaststaat hoe die prestatie eruit moet zien.

Opbrengstgericht werken

Opbrengstgericht werken is tegenwoordig een gevleugelde term in onderwijsland. De overheid streeft ernaar de opbrengsten van het onderwijs (dat wil zeggen de scores van leerlingen op toetsen) te verhogen en neemt aan dat dit mogelijk is door de opbrengsten nauwkeurig te volgen. De evaluatieve cyclus van opbrengstgericht werken (OGW) heeft veel weg van het model didactische analyse. Beide passen in de instrumentele benadering.



Afbeelding 2.2.

Ook hier zien we dat de resultaten van metingen gebruikt worden voor het nemen van beslissingen, waarna nieuwe doelen vastgesteld worden. OGW wordt toegepast op diverse niveaus in het onderwijs: op het niveau van de individuele leerling, de groep, de school, de schoolvereniging, het hele land. De aanname is dat er streefdoelen gesteld worden, die net iets hoger liggen

dan de huidige scores. Op grond van de huidige scores kan een analyse gemaakt worden van de subdoelen die de leerlingen nog niet beheersen. Op grond daarvan kan de instructie verbeterd worden doordat hij beter is toegespitst op wat studenten nodig hebben. Zo wordt verondersteld dat prestaties in het onderwijs toenemen door te meten. Het instrumentele paradigma is optimistisch over de mate waarin processen binnen de school maakbaar zijn (De Mynck, Both & Visser-Vogel, 2013).

Metten, economie en ranking

Doordat er veel meetinstrumenten beschikbaar zijn, ook internationaal, is het mogelijk vergelijkingen te maken tussen scholen onderling en tussen het onderwijs in diverse landen. Zo krijgen toetsen nog een andere functie erbij. Er ontstaan zogenaamde ranglijsten van goed naar minder goed. Ranglijsten worden gemaakt van landen, (hoge)scholen en universiteiten en van studenten, waar eigenlijk niet van?

De instrumentele manier van denken wordt toegepast zoals hij ook in bedrijven en fabrieken wordt gebruikt. Net als bij bedrijven kan men immers met begin- en eindmetingen vaststellen wat het rendement is van een investering. Zo kan bijvoorbeeld geprobeerd worden om met behulp van Cito-scores het rendement van het onderwijs te bepalen. Men toetst het instroomniveau en kijkt wat de uiteindelijke opbrengst is en neemt aan dat die gegevens iets zeggen over de kwaliteit van het onderwijs. Hierover is veel discussie.

De vormingsbenadering

Het vormingsparadigma legt andere accenten dan het instrumentele paradigma. Onder vorming verstaan we de *eigen identiteitsontwikkeling van studenten die plaatsvindt onder invloed van prikkeling van buitenaf die zij op hun eigen manier in zich opnemen*. Dit wordt ook wel aangeduid met het Duitse woord *Bildung*. Daarin herkennen we het woord 'beeld'. Het woord heeft in de christelijke traditie de connotatie dat je niet zomaar iemand bent, maar bent geschapen naar het beeld van God. Mensen hebben dus een hogere roeping. Het vormingsideaal leeft echter niet exclusief binnen een christelijke context. In hoofdstuk 5 zullen we zien dat diverse onderwijsvernieuwingen van de laatste decennia op concepten uit de vormingsbenadering waren geïnspireerd.

We zullen nu met behulp van voornamelijk het klassieke denken over vorming enkele aspecten van het vormingsparadigma apart onderscheiden. We maken

daarbij gebruik van recente verwoordingen en interpretaties van het vormingsbegrip zoals die te vinden zijn bij Vanderstraeten, 2001, Van Crombrugge, 2011, Gude, 2012, Pikkareinen, 2012 en De Muynck & Vos, 2014.

De persoon vormt zich in de culturele omgeving

Het is onjuist het woord ‘vormen’ op te vatten als ‘gevormd worden’ in de zin dat studenten zich volgens een vooraf bepaald format zouden moeten ontwikkelen. Met vorming wordt juist geduid op een proces waarbij de zich vormende mens zelf actief is. In de vormingsbenadering staat de vrijheid in het leerproces voorop. Studenten worden niet in de eerste plaats geprogrammeerd volgens een geplande route, maar beslissen wat betreft vorm en inhoud mee over wat ze gaan doen. Vrijheid betekent ook dat de lerende persoon onverwachte ontdekkingen mag doen die geheel buiten het curriculum vallen. Wat gechargeerd uitgedrukt biedt de docent veel aan en laat tegelijk de ontwikkeling haar gang gaan. In het vormingsparadigma is er niet de neiging om te controleren of doelen precies zijn gehaald. Docenten die denken vanuit vorming, hebben een sterk vertrouwen in de uitkomsten van de vorming, ook zonder dat zij grip willen hebben op deze uitkomsten.

Dat vorming een vrij proces is, betekent niet dat de docent een vrijblijvende en afwachtende houding aanneemt, integendeel zelfs. Biesta (2012, p. 75 e.v.) spreekt in dit verband van de pedagogiek van de onderbreking. Studenten stellen vanuit zichzelf vragen, omdat zij een intrinsieke motivatie hebben de werkelijkheid te begrijpen maar daardoor alleen ontstaat geen diepgang. Ze hebben de ontmoeting nodig met iets of iemand die hen in hun ‘natuurlijke’ leerproces onderbreekt en iets nieuws inbrengt wat zij nog niet eerder zijn tegengekomen. Het ‘onderbroken worden’ correspondeert niet met wat in een instrumentele benadering aan leerlijnen en doelstellingen kan worden uitgedacht, want daar gaat het om voorspelbaarheid. Hier gaat het erom studenten op een ander been te zetten en zelf te leren nadenken (zie kader).

Roefs (2010) ontdekte dat docenten in het hoger onderwijs vooral als inspirerend ervaren worden als zij hun studenten met voorbeelden of stellingen in aanraking brengen, die hen in de war brengen. Een student zegt over zo’n docent: “Hij zet mij aan tot nadenken over specifieke onderwerpen. Ik word echt geprikkeld om over bepaalde thema’s dieper na te denken en daar de goede argumenten voor te vinden” (p. 62). De les heeft een uitwerking die niet goed

in een leerdoel te vangen is. Het leren gaat ook door buiten de hogeschool. Dezelfde student zegt: “Daar ga ik dan ook na de lessen met mensen over in discussie, met mijn ouders of vrienden. Zo ontdek ik nieuwe interesses” (p. 62). Niet door een uitgekiend programma te volgen verwerven studenten nieuwe inzichten, maar juist doordat ze op een ander been worden gezet. Dit op het andere been zetten heeft in het vormingsparadigma een brede strekking: onderwijs moet uitdagen om de grote levensvragen te helpen stellen, zoals: wat is een menswaardig bestaan, hoe gaan we om met fouten, met schuld (individueel en collectief). Vormingsgericht onderwijs thematiseert dus ook de morele dimensie van het menselijk handelen. Vorming gaat niet alleen over kennis en vaardigheden maar ook over wie je bent, over zelfkennis en over hoe je in de wereld staat.

Gespreksvraag: Zijn er momenten in uw lespraktijk waarin u studenten probeert te ‘ontregelen’?

Docenten die betrokken zijn bij het vormingsproces realiseren zich dat hun bijdrage voornamelijk bestaat uit het voorzien van de juiste prikkels. Behalve ontregelende vragen kunnen deze prikkels ook voortkomen uit een diepgaande beschouwing van de docent. Hoe de persoon van de student zich vormt, onttrekt zich voor een groot deel aan de waarneming. Volgens de Duitse bildungs-traditie gebeurt vorming in het innerlijk. “Es bildet ein Talent in der Stille, sich ein Charakter in dem Strom der Welt” (geciteerd door Gude, 2012 p. 35). Dat betekent dat docenten hun studenten niet voortdurend lastig moeten vallen. Er is innerlijke rust en rijpingstijd nodig om vorming te laten gebeuren. Volgens de vormingsbenadering kan de voortdurende monitoring van vaardigheden en het werken van het ene toetsmoment naar het andere een gevaar zijn omdat er dan geen ‘rust is om te rijpen’.

Vorming komt tot stand door een samenspel van impliciete en relatief onvoorspelbare factoren, zoals de houding van de opleider, de sfeer in de opleiding, vriendschappen, invloeden binnen het gezin waaruit de studenten afkomstig zijn en onverwachte gebeurtenissen. Omdat dit samenspel aan factoren zo belangrijk is, is het essentieel welk klimaat een opleiding heeft en is het steeds de vraag hoe een optimaal klimaat gerealiseerd kan worden. Dat betekent niet dat het vormingsproces maakbaar en dus beheersbaar is. Kern van de vormingsbenadering is dat de persoon geroepen is zich te verhouden tot alle prikkels die op hem of haar af komen (Von Hentig, 1996). Daarom

moet vorming niet verward worden met zelfsturing. Vormend onderwijs beoogt iets los te maken bij studenten waardoor zij zichzelf gaan vormen. Dat persoonlijke kwaliteiten en talenten van studenten actief worden ingeschakeld is daarom essentieel in deze visie.

Gespreksvraag: Waar is in uw lespraktijk sprake van 'open' ruimte? Wat belemmert en bevordert het ontstaan van open ruimte?

Hoewel de student zichzelf vormt is de omgeving er bij betrokken. Vorming gebeurt in relatie tot de cultuur waarin men leeft, zo zeggen de klassieke vormingsfilosofen. Daarom is vorming niet gelijk aan 'zelfontplooiing'. Dat woord suggereert dat er een soort pure kern zou zijn, die op een autonome manier tot ontplooiing kan komen en daarin ook een eigen route volgt. In de vormingsbenadering is het juist zo dat een leerling zich eerst allerlei inzichten en verbanden eigen maakt die er in de cultuur al zijn, zich daarin opgenomen weet, maar vervolgens een eigen oorspronkelijke bijdrage daaraan levert (Gude, 2012, p. 43-44). De gevormde mens zal zelf namelijk ook invloed gaan hebben op de wereld om hem heen.

In het verleden gingen Westerse landen zich steeds meer als natiestaten ontwikkelen (Taylor, 2007; Gude, 2012). In Nederland wilde de overheid in de negentiende eeuw onder meer door middel van het onderwijs de bevolking burgerlijke en christelijke deugden bijbrengen. Het vormingsdenken staat kritisch tegenover overheden die er behoefte aan hebben de bevolking te leren zich aan te passen aan de gewenste orde. In sommige landen zoals Singapore (zie het kader op pag. 32) wordt het onderwijs nog steeds voor dit doel gebruikt. We spreken dan van staatspedagogiek, waar het vormingsdenken wars van is. Omdat vorming het belang van kritisch denken benadrukt, kan het een maatschappijkritische uitwerking krijgen.

Het vormingsproces heeft dus twee kanten. Aan de ene kant is er de student die zich met zijn of haar gaven op een unieke manier ontwikkelt tot een volwassen professional die zelf verantwoordelijkheid kan dragen. Aan de andere kant is er de omgeving (de opleiding, de onderwijspraktijk, de maatschappij) die dit vormingsproces mogelijk maakt (naar Roeleveld, 2006, p. 35).

"Binnen het vormingsparadigma moeten leraren er rekening mee houden dat vorming 'gebeurt'."

Vorming betreft de hele mens

In het vormingsparadigma is oog voor de hele mens. Het accent van het instrumentele denken op deelvaardigheden wordt als erg smal ervaren. In alle soorten van (beroeps)onderwijs zou men oog moeten hebben voor diversiteit in talenten van studenten. Deze opvatting hangt samen met een, wat we noemen, holistische visie op de mens. De mens is niet alleen een denkend wezen maar heeft ook gevoelens en een wil en is aangelegd op handelen. Hoofd, hart en handen zijn deelaspecten die allemaal belangrijk zijn in de vorming van studenten. Er wordt dus vorming nagestreefd in brede zin, met inbegrip van verstandelijke, emotionele, sociale, geestelijke en morele aspecten. Binnen het christelijke denken over de mens wordt het hart als essentiële levenskern beschouwd. Het hart geeft richting aan het handelen van de leerling of de aanstaande beroepsbeoefenaar.

Bij de ontwikkeling van talenten gaat het niet om afgebakende deelvaardigheden maar om 'levenskennis'. Wanneer kennis wordt gebruikt in levensechte situaties gaat het plotseling om veel meer vragen dan alleen wat 'juist' is of 'onjuist'. Als studenten kennis toepassen, komt het aan op de vraag wie zij zijn en hoe zij in het leven staan. Een levensechte leeromgeving bevordert ook de diepgang. Diepgang betekent dat de zich vormende mens de ruimte krijgt of zelfs uitgedaagd wordt om zich existentiële vragen te stellen. Het motto 'leren voor het leven' staat dus in het verlengde van het vormingsparadigma. Toch betekent dit niet dat het onderwijs louter op praktisch bruikbare (beroeps) kennis gericht dient te zijn. Juist een bredere blik dan de directe beroepsvaardigheden is van belang. Een gevormd iemand wordt geacht geïnteresseerd te zijn in geschiedenis, cultuur, economie en de kwaliteit van de toekomstige samenleving.

Vorming is 'nadenken over' (docent rekenen pabo)

"Vorming zie ik als 'nadenken over', waarbij je naar jezelf leert kijken, dingen ziet in je omgeving. Ik wil mezelf als docent en de wereld eigenlijk openen en daarin de student vormen, openen. De vormende waarde van wiskunde vind ik ook het genieten van de wiskunde, het mooie dat je tot de oplossing van een probleem kan komen. De vormende waarde kan ook de schoonheid van het getallensysteem zijn: niet perfect maar wel heel mooi. Als ik samen met studenten iets van die schoonheid ervaar, dan geniet ik!"

Vorming is voor allen; geen selectie

De stelling dat een gevormd mens geïnteresseerd is in geschiedenis, cultuur, economie en de kwaliteit van de toekomstige samenleving klinkt in sommige oren misschien elitair, maar het vormingsparadigma richt zich juist op alle mensen, zonder onderscheid. Voor Comenius, een belangrijke denker in het vormingsparadigma, was een belangrijk uitgangspunt dat vorming een proces is dat iedereen meemaakt, onafhankelijk van de aard van zijn of haar talent (Pikkareinen, 2012; Van Crombrugge, 2011). Onderwijs binnen een vormingsparadigma heeft daarom ook een niet-selectief karakter. In hoofdstuk 4 zullen we zien dat er in het vormingsparadigma ook geëvalueerd wordt, maar er wordt niet gemeten om te selecteren en leerlingen op te delen in verschillende groepen. Vorming is bedoeld voor allen. Alle leerlingen zijn dragers van de cultuur van de toekomst. Van gevormde mensen wordt verwacht dat ze een kritisch reflectief vermogen ontwikkelen, waardoor ze argumenten kunnen aandragen om bepaald handelen al dan niet te ondersteunen. Het leren vellen van een zelfstandig eigen oordeel telt zwaar. Gevormde mensen staan met interesse in de wereld en beoordelen op grond van argumenten wat er gebeurt. Kritisch leren denken hoort bepaald niet alleen thuis op de universiteit (zie kader hieronder).

3 Teken minder.



afbeelding 2.3.

Zijn rekenlessen geschikt om kritisch te leren denken of staan de antwoorden zo vast dat kritisch denken niet aan de orde is?

Leerlingen uit groep drie zijn bezig met de begrippen 'meer' en 'minder' op de getallenlijn. In de opdracht staan voorwerpen getekend. In het rondje staat hoeveel

het er zijn. De kinderen moeten van ditzelfde voorbeeld minder exemplaren tekenen en het cijfer in het rondje zetten. Een student observeert hoe Christiaan de opgaven maakt. De derde opgave vraagt om minder huisjes te tekenen dan één. Dat brengt hem op een idee: nul is altijd minder dan wat er ook getekend staat. Snel zet hij een nul bij de laatste opgave en klaar is hij! Hij hoeft geen krukjes meer te tekenen. De student vindt Christiaan een slimme jongen en geeft hem een compliment. De mentor die het werk nakijkt, ziet het echter anders. Ze stuurt Christiaan terug om de laatste opgave te verbeteren. De student ziet hem teruglopen met verongelijkte schouders en herinnert zich dat ze Christiaan juist een compliment gegeven heeft. Ze vertelt de mentor dat ze het antwoord van Christiaan niet fout vindt. Nul is toch minder dan vijf? Maar de mentor is dat niet met haar eens. Ze zegt: “Er staat ‘teken minder’, dus Christiaan moet in ieder geval iets tekenen.”

Heeft er iemand gelijk?

De student en de mentor besluiten dat hun ‘verschil van mening’ een enorme kans is om kinderen te leren dat je op verschillende manieren naar de werkelijkheid kunt kijken. Ze leggen hun ‘verschil van mening’ aan de kinderen voor en introduceren een coöperatieve werkvorm, waarbij de discussie streng gestructureerd wordt. De kinderen krijgen allemaal een blaadje. Met een tekening mogen ze aangeven wie gelijk heeft en waarom. Daarna worden de standpunten in hun groepje uitgewisseld, vergeleken en samengevat, weer door middel van tekeningen. Daarop volgt een klassengesprek over de uitkomsten. Met grote tevredenheid kijken de mentor en de student terug op deze les! Het ging over rekenen, maar ook over kritisch denken.

Gespreksvraag: Herkent u in dit voorbeeld iets van het instrumentele paradigma en/of van het vormingsparadigma?

Een vergelijking van de paradigma's

Nu we deze twee manieren van denken naast elkaar hebben geplaatst, komen we tot een vergelijking. We bespreken de vraag of sociale en artistieke doelen meer bij het ene of het andere paradigma thuishoren. En hoe zit het met de neiging om onderwijs vooral in termen van economisch nut te waarderen? Past die bij een van beide benaderingen? Ten slotte zetten we verschillen tussen de twee benaderingen op een rij.

De paradigma's en sociale en artistieke doelen

Hierboven hebben we gezien dat de vormingsbenadering zich met name interesseert voor de vraag 'wie is de leerling?', terwijl de instrumentele benadering zich interesseert voor de vraag 'wat is de leerstof?' Betekent dit dat sociale doelstellingen binnen het vormingsparadigma thuishoren, terwijl cognitieve doelstellingen binnen het instrumentele paradigma thuishoren? Die relatie ligt complexer.

Het instrumentele paradigma richt zich niet op de eigen persoonvorming, zoals het vormingsparadigma, maar kan zich wel bezighouden met gedrag. Op een impliciete manier heeft het instrumentele paradigma overigens wel gevolgen voor de eigen persoonsvorming: omdat vrijheid en eigen verantwoordelijkheid minder worden benadrukt, ligt het voor de hand dat het onderwijs bijdraagt aan vorming waarin studenten leren voldoen aan verwachtingen van buiten. Die verwachtingen hebben meestal te maken met hoe studenten omgaan met leerstof, maar kunnen ook betrekking hebben op de manier waarop zij zich in sociale situaties behoren te gedragen.

Onlangs heeft de Universiteit van Amsterdam een meetinstrument ontworpen voor burgerschapsvorming (www.burgerschapmeten.nl). Met dit meetinstrument kunnen scholen zicht krijgen op de burgerschapscompetenties en het burgerschapsgedrag van hun leerlingen. Centraal staan vier sociale taken die kinderen en jongeren als burger moeten kunnen vervullen: democratisch handelen, maatschappelijk verantwoord handelen, omgaan met conflicten en omgaan met verschillen. Het meetinstrument bevat een onderdeel dat door de leerlingen ingevuld wordt en een onderdeel dat de leraar invult. Vanuit het vormingsparadigma zijn vraagtekens te plaatsen bij dit instrument maar het bestaan ervan toont aan dat het wel degelijk mogelijk is om sociale doelen vanuit een instrumenteel paradigma te benaderen.

Iets dergelijks geldt voor artistieke doelen. Laten we muziek als voorbeeld nemen. Het muziekonderwijs kan er helemaal op gericht zijn dat studenten door anderen geschreven muziek reproduceren waarbij ze zo precies mogelijk volgen hoe anderen vinden dat de muziek gespeeld moet worden. Iedere les gaat de leraar noot voor noot na of de student de partituur juist uitvoert. Anderzijds kan muziekonderwijs ook gericht zijn op leren improviseren en een eigen stijl vinden. Ook artistieke doelen kunnen dus vanuit beide paradigma's aan de orde komen.

Een school in Singapore – instrumenteel of vormingsgericht?

Als we het schoolplein oplopen, zien we hoe de waarden van de Singaporese staat prominent in zes punten op de gevel van de school staan: “We moeten ons nationaal erfgoed eerbiedigen en trots zijn op de manier waarop we onze unieke levensstijl vormgeven. We zijn één, ondanks raciale en godsdienstige verschillen. Naar de mate van capaciteiten en inzet heeft iedereen gelijke kansen. Voor onze welvaart zijn we zelf verantwoordelijk. We verdedigen ons land met trots. We hebben vertrouwen in onze toekomst.”

Als we de school binnenlopen, zien we dat het onderwijs niet alleen maar gaat om prestaties wat betreft kennis en vaardigheden. Overal op de gangen hangen plakaten over de waarden van de school. Bij alle waarden zijn drie prestatieniveaus beschreven. Deze niveaus gelden onder meer voor sociaal-emotionele vorming, voor esthetische vorming, morele vorming en lichamelijke vorming. Daarbij gaat het onder andere om verantwoordelijkheid nemen voor het eigen leerproces, besef van goed en kwaad, initiatief tonen, bereid zijn om risico te nemen, gericht zijn op hoge prestaties, conflicten oplossen, met anderen meeleven, in teams samenwerken en bijdragen aan de gemeenschap op school en in de klas. Alles is bedoeld om bij te dragen aan personen met zelfvertrouwen die kunnen samenwerken en verantwoordelijke burgers zijn in een globale samenleving.

We bestuderen de prestatieniveaus van de waarde ‘bijdragen aan de gemeenschap op school en in de klas’. Op het niveau ‘op weg om het niveau te halen’ staat het volgende: *“Meestal vriendelijk, als het gevraagd wordt zal hij anderen helpen of erbij betrekken. Is meestal bereid samen te werken, maar heeft wel aanmoediging en ondersteuning nodig.”*

Bij ‘haalt het niveau’ staat te lezen: *“Vriendelijk, hulpvaardig en overwogen. Is betrokken bij de gemeenschap en draagt eraan bij. Heeft soms aanmoediging en ondersteuning nodig.”*

Bij ‘presteert boven het niveau’ staat: *“Aardig en vriendelijk, zoekt vaak mogelijkheden om anderen te helpen of erbij te betrekken. Neemt vrijwillig verantwoordelijkheid in de klas. Ondersteunt anderen en moedigt hen aan.”*

Gespreksvraag: Werkt deze school volgens het vormingsparadigma, of pakt de school vorming op een instrumentele manier aan?



Afbeelding 2.4.

De paradigma's en de kenniseconomie

Het instrumentele paradigma wordt vaak in verband gebracht met een nadruk op kennis. Dat de overheid de toetsprestaties tegenwoordig zo nauwlettend in de gaten houdt, wordt vaak gemotiveerd vanuit economische motieven. Nederland wordt een kennissamenleving genoemd met een grote dienstensector die afhankelijk is van een hoog kennisniveau. Sinds een aantal jaren leeft de angst dat het kennisniveau op een te laag peil staat of raakt. Iedere keer weer verschijnen er lijsten waarop de prestaties van Nederlandse leerlingen vergeleken worden met die in andere landen. Daarom legt de overheid steeds meer nadruk op reken- en taalprestaties (die door middel van de toetsen gemeten worden). Je kunt je dus afvragen of er de laatste tijd sprake is van een 'cognitieve switch' (Kaldewey en Bulterman-Bos, 2013). In de lerarenopleidingen is er eveneens de tendens om veel nadruk te leggen op de kennisbasis van aanstaande leraren. De aandacht voor parate kennis en goed getrainde vaardigheden past in het instrumentele paradigma omdat de nadruk valt op training van vooraf vastgestelde doelen. Maar kun je daarmee stellen dat er vanuit economische belangen vooral een stimulans uitgaat naar onderwijs volgens het instrumentele paradigma? Dat laatste is niet het geval. Een stimulans voor het vormingsparadigma kan evenzeer economisch gemotiveerd worden. De ontwikkeling van creativiteit en persoonlijke talenten, die binnen het vormingsparadigma wordt bepleit, past ook heel goed binnen de vaardigheden die de economie van de 21ste eeuw nodig heeft. Of het een goede zaak is dat onderwijs ondergeschikt gemaakt wordt aan eco-

nomische belangen, is een heel andere kwestie. Verder in dit boek dragen we bouwstenen aan om de lezer te helpen over deze vraag kritisch na te denken. Vooralsnog volstaat de conclusie dat het belang van beide paradigma's vanuit economische belangen kan worden gemotiveerd.

Een vergelijking van de twee paradigma's in vier punten

Als we de vormingsgerichte benadering systematisch tegenover de instrumentele benadering plaatsen, ontstaat het volgende overzicht.

INSTRUMENTELE BENADERING	VORMINGSBENADERING
1) De student wordt gevormd door het onderwijs volgens een vastgesteld traject.	1) De student vormt zichzelf in een sociaal-culturele omgeving waar het onderwijs deel van uitmaakt. De uitkomsten staan niet vast.
2) Het onderwijs is gericht op toetsbare kennis en uiterlijk waarneembaar gedrag.	2) Het onderwijs betreft de hele mens en is gericht op het volle leven.
3) Het onderwijs is gericht op objectieve kennis.	3) Het onderwijs is gericht op kritisch denken en leren oordelen.
4) In het onderwijs wordt op basis van objectieve criteria geselecteerd en gedifferentieerd.	4) In het onderwijs wordt niet geselecteerd, vorming is voor allen.

Slot

Nu we de verschillen en overeenkomsten op een rij hebben, zal de lezer zich mogelijk in het ene paradigma veel meer thuis voelen dan in het andere. Het kan ook zijn dat docenten zich *unheimisch* voelen wanneer ze een keus voor het een of het andere zouden moeten maken. Maar is het nodig om een keus te maken? Of is het mogelijk te schakelen tussen de twee paradigma's? Voordat we op die vraag ingaan, zullen we in de volgende hoofdstukken eerst de betekenis van de twee paradigma's voor de toetspraktijk belichten.

Gespreksvraag: Bij welk paradigma voelt u zich thuis en waarom? Deze vraag komt in dit boek vaker terug. Daarom kan het interessant zijn het antwoord op te schrijven en uw eigen ontwikkeling te volgen.

Hoofdstuk 3

Toetsen vanuit het instrumentele paradigma

Soms laten films zien hoe het hoger onderwijs er vroeger uit zag. In een klassieke collegezaal deliberaert een geleerde heer over zijn wijsheden. Jonge mensen luisteren met ernstige gezichten. Hier en daar klinkt een stem van een student. De docent stelt vragen; hij laat zich graag uitdagen. Vervolgens zie je een van de studenten door de stad fietsen. Bij een statig herenhuis belt ze aan. Ze wordt binnengelaten, treedt de studeerkamer van de docent binnen en wordt mondeling bevraagd over de tentamenstof. Meetlatten worden niet gehanteerd; het gaat om de kwaliteit van het gesprek. Uiteindelijk wordt er echter wel een oordeel uitgesproken. Meestal gaat het goed; in dit geval niet. De student worstelt en probeert haar visie nogmaals uiteen te zetten. Ze bereikt de docent niet. Mistroostig fietst ze na afloop van het tentamen door de stad terug naar haar kamer. Ze probeert op een rijtje te krijgen wat er mis is gegaan. Was kritisch ingaan op de stellingen van de docent de bedoeling geweest? Waren haar eigen opmerkingen gerechtvaardigd? Spelen persoonlijke factoren een rol? Of was het juist die éne zin in haar antwoord die de docent op het verkeerde been zette? Een antwoord komt er niet. De docent heeft gesproken.

Zo ging het vroeger. Vanuit het instrumentele paradigma zijn begrippen en modellen ontstaan om situaties zoals hier beschreven te verbeteren. Doelstellingen en objectieve toetsen scheppen kaders en geven richting die het onderwijs ten goede komen: studenten weten beter wat er van hen verwacht wordt, docenten hebben gronden om hun beslissingen te onderbouwen.

In dit hoofdstuk laten we de begrippen uit de instrumentele traditie de revue passeren. We zullen zien dat het instrumentele paradigma een antwoord geeft op een praktijk waarin de docent zich niet hoeft te verantwoorden. Aan het eind van dit hoofdstuk wordt echter duidelijk dat de oplossingen van het instrumentele paradigma weer nieuwe problemen oproepen, die op hun beurt weer vragen om een antwoord.

Evalueren vanuit het instrumentele paradigma

Het instrumentele paradigma hebben we in hoofdstuk 2 beschreven als een benadering waarbij afzonderlijke leerdoelstellingen getoetst worden door middel van instrumenten die objectief meten. Met objectief bedoelen we het streven om feiten (de werkelijkheid zelf) en meningen (de menselijke interpretatie van die werkelijkheid) te scheiden. Dit hoofdstuk begint met het nut van onderwijsdoelstellingen en gaat dieper in op de vraag waarop gelet moet worden bij het ontwikkelen van toetsinstrumenten. Daarbij passeren de begrippen betrouwbaarheid en validiteit de revue. Omdat dit boek bedoeld is voor visievorming – handleidingen om goede toetsen te maken zijn elders voorradig – gaan we vooral in op het ‘waarom’ van deze begrippen en verwijzen we voor de praktische uitwerking van toetsen maken naar publicaties die mbo- en hbo-instellingen doorgaans zelf in huis hebben.

Doelstellingen

Het instrumentele paradigma veronderstelt dat bij de leerstof doelstellingen worden geformuleerd. Vragen naar de doelstellingen kunnen lijken op vragen naar de bekende weg. Doelstellingen, vallen die niet samen met het doorwerken van de stof? Als docenten doelstellingen louter op het niveau van de stof formuleren, dus in termen van wat plaatsvindt in de les of colleges, dreigen doelstellingen bureaucratische blokken aan het been te worden. Want wat is de meerwaarde van een herhaling van onderwerp van de les?

Als het goed is, zijn doelstellingen als koppen bij een tekst. Ze geven aan waar een groter geheel over gaat. Doelstellingen formuleren nodigt de docent uit om na te denken over het kader waarin een afzonderlijk les staat. Inderdaad, in deze les zijn we bezig met deze leerstof maar waarom doen we dat? Met welke andere begrippen heeft de leerstof van vandaag te maken? Welk taalgebruik hoort niet alleen bij dit college maar ook bij andere? Hoe help ik mijn studenten zich grotere gehelen eigen te maken?

Uiteraard kan het zinvol zijn om onderscheid te maken tussen cursusdoelstellingen en college- of lesdoelstellingen. Cursusdoelstellingen betreffen de verwachte uitkomsten van de gehele cursus terwijl college- of lesdoelstellingen gaan over de verwachte uitkomsten van ieder afzonderlijk college of ieder afzonderlijke les. Het belang van het formuleren van doelstellingen ligt echter in beide gevallen op metaniveau. Er wordt richting gegeven aan de stof, of het nu gaat om grote of om kleine hoeveelheden. Doelstellingen delen de stof op in zinvolle gehelen, evenals kopjes in een tekst dat doen. Zo wordt de tekst leesbaar en de stof behapbaar.

Betrouwbaarheid en validiteit

Het instrumentele paradigma veronderstelt vervolgens dat vastgesteld wordt of de doelstellingen gehaald zijn. Daartoe worden instrumenten ontworpen in de vorm van toetsen die zo objectief mogelijk meten. Het moet niet uitmaken wie deze instrumenten afneemt en ook niet waar of wanneer deze toetsen worden afgenomen.

Toetsen moeten valide en betrouwbaar zijn. Eerst volgt een korte toelichting bij het begrip validiteit, daarna wordt het begrip betrouwbaarheid toegelicht. Validiteit gaat over de vraag of de test inderdaad meet wat hij behoort te meten. Gaat de toets over de afgesproken leerdoelen, komen de verschillende leerdoelen op een evenwichtige manier aan bod, klopt de puntentelling met de zwaarte van de leerdoelen?

Een docent economie is gewend om de toetsen af te sluiten met een aantal opgaven die verder gaan dan de vooraf vastgestelde doelstellingen in het boek en verder gaan dan wat hij besproken heeft. Aan de hand van deze opgaven wil hij erachter komen wie de slimme studenten zijn en wie niet. In de docentenvergadering wordt een van zijn toetsen besproken. Hij krijgt kritiek in verband met de validiteit van de toetsen. Een collega stelt dat de toets moet meten of de leerdoelstellingen zijn gehaald. Dat de toets gebruikt wordt om op intelligentie te selecteren, vindt hij niet terecht. De taak van de docent is niet om te selecteren, vindt hij. Het onderwijssysteem doet dit al genoeg. De taak van de docent is om studenten te helpen om de opleiding te halen.

Gespreksvragen: Deze docent gaat met zijn toets verder dan de doelstellingen van het boek. Vindt u dit terecht? Mogen docenten eigen doelstellingen toevoegen? Wanneer wel en wanneer niet?

In hoeverre meten uw toetsen wat ze (volgens u en volgens het opleidings- en toetsplan) behoren te meten?

Een toets moet betrouwbaar zijn zoals een weegschaal betrouwbaar moet zijn. Als je er om 11.05 uur op gaat staan, moet hij hetzelfde getal aangeven als wanneer je er om 11.10 uur op gaat staan mits je natuurlijk in deze paar minuten niets gegeten hebt. Zo wordt ook van een toets verwacht dat de uitkomsten

van vragen consistent zijn. Een tweede keer toetsen, of afname van een vergelijkbare toets, moet dezelfde resultaten opleveren.

In een docentenvergadering ligt de cijferlijst van Rafael op tafel. Zijn mentor beschrijft hem als een zwakke student. En inderdaad liggen zijn gemiddelden vrijwel allemaal onder de zes. Maar dan merkt een van de docenten op dat het maar de vraag is of deze metingen echt iets zeggen over de capaciteiten van Rafael. Het valt hem op dat zijn resultaten sterk wisselen. Het komt vaak voor dat zijn resultaten bij dezelfde vakken variëren tussen de twee en de acht. Een zwakke student haalt geen achten. Hoe vallen die achten te verklaren? Volgens zijn mentor gaat het hier vaak om toetsen waarvoor niet veel huiswerk gemaakt hoefde te worden waarop de andere docent opmerkt dat Rafael dan misschien wel juist een slimme student is. Hoe dan ook, de docenten zijn het onderling eens dat – als het waar is dat de cijfers alleen maar indicatoren zijn van de capaciteiten van Rafael – er veel minder variatie tussen de cijfers had moeten bestaan.

Gespreksvraag: Hoe is bij u de variatie tussen de verschillende scores van dezelfde student? Hoe verklaart u deze variatie?

Hoogleraar didactiek en toetsing van het taalvaardigheidsonderwijs Huub van den Bergh (2006) geeft een mooi voorbeeld waarin de begrippen betrouwbaarheid en validiteit verduidelijkt worden. In het voorbeeld, dat we hierna weergeven, laat hij ook zien dat het niet altijd eenvoudig is om aan die eisen te voldoen. We gaan in op zijn redenering om de discussie over betrouwbaarheid en validiteit tot leven te brengen en daarmee docenten het gevoel te geven dat het leuk is om over deze vragen na te denken.

Van den Bergh laat een groep leerlingen van de theoretische leerweg in het VMBO en een groep leerlingen van de basisgerichte leerweg drie tekstbegrip-toetsen maken. Omdat de leerlingen van de theoretische leerweg op een hoger niveau onderwijs volgen, zou je verwachten dat zij het hoogst scoren bij alle drie de toetsen. Voor de eerste toets klopt dit. Maar bij de tweede toets zijn de resultaten omgedraaid: de leerlingen van de basisgerichte leerweg hebben in vergelijking met de leerlingen van de theoretische leerweg gemiddeld veel meer vragen goed beantwoord. Gelukkig is er nog een derde toets. Op die toets zijn de verschillen minimaal. In dit experiment worden dus drie toetsen afgenomen met drie verschillende uitkomsten. Duidelijk is dat de toetsen er niet in slagen om de groep van de theoretische leerweg te onderscheiden van de groep



van de basisgerichte leerweg. Dit duidt op onbetrouwbare meetinstrumenten. Hoe komt het dat de metingen zo variëren en dus onbetrouwbaar zijn? Van den Bergh licht een tipje van de sluier op door zich af te vragen wat de toetsen feitelijk gemeten hebben. Nadere analyse leert dat de beroepsgerichte leerlingen koks in opleiding waren. De tweede toets, waarop zij tegen de verwachting in hoog scoorden, ging over een tekst met recepten. De derde toets, waarop zij even goed scoorden, betroffen uit het Oudnederlands vertaalde recepten. Kennelijk hadden de koks veel voordeel van hun grotere kennis van recepten en ervaring met het lezen van de manier waarop een recept wordt beschreven. Van den Bergh concludeert dat de toets niet alleen tekstbegrip meet, maar tegelijkertijd ook nog algemene (wereld)kennis. Kennelijk meet de test dus meer dan de bedoeling is, namelijk tekstbegrip. Dit duidt op een validiteitsprobleem, want het was de bedoeling om alleen maar tekstbegrip te meten. Maar Van den Bergh stelt dat het bij Nederlands bijna onvermijdelijk is dat de toets ook gaat over kennis van de wereld van leerlingen. Goede toetsen maken over tekstbegrip is heel moeilijk. In dit voorbeeld wordt trouwens ook duidelijk dat betrouwbaarheid en validiteit in de praktijk vaak hand in hand gaan. Het gebrek aan validiteit valt op te maken uit een gebrek aan betrouwbaarheid, die weer blijkt uit de inconsistente resultaten: de testresultaten van de laagste groep zijn niet altijd het laagst.

Betrouwbaarheid als instrumenteel begrip

De term ‘betrouwbaar’ roept veel op. Iets is betrouwbaar als je erop af kunt gaan zonder bang te hoeven zijn dat je bedrogen wordt. Mensen zijn betrouwbaar als ze eerlijk en geloofwaardig zijn. De term ‘betrouwbaarheid’ heeft in het dagelijks spraakgebruik vaak een morele betekenis die verwijst naar iets goeds, terwijl onbetrouwbaarheid verwijst naar wat niet goed is. Het instrumentele denken probeert echter zo waarde vrij mogelijk te zijn. De term betrouwbaarheid heeft daarbij geen morele maar een procedurele betekenis. Termen als ‘consistentie’ en ‘duidelijkheid’ zouden – omdat ze minder suggereren dat je op de test kunt afgaan zonder bedrogen uit te komen – een geschiktere aanduiding zijn dan de term betrouwbaarheid (Koretz, 2008). De term betrouwbaarheid is een technische term zonder morele betekenis.

Docent rekenen mbo: “Ik ben het op twee punten niet eens met de manier waarop het Cito rekentoetsen voor het mbo maakt. Om te beginnen worden er heel veel redactiesommen gegeven. In plaats van rekenen ben je meer bezig met het toetsen van begrijpend lezen. Sommige studenten kunnen eigenlijk geen Nederlands, terwijl ze wel sommen kunnen maken. En dan nog iets anders. Het Cito hanteert bij alle studenten afkomstig uit de afdelingen theoretische leerweg (tl), kaderberoepsgerichte leerweg (kb) en basisberoepsgerichte leerweg (bb) hetzelfde examen. Bij tl is de normering zwaarder. Dat vind ik dwaasheid. Voor een bb leerling zijn de formuleringen van het andere niveau veel te moeilijk. Je moet per doelgroep formuleren in een toetsitem! Als je dat niet doet, klopt er toch iets niet?”

Gespreksvraag: Op welke manier betreffen de twee punten van kritiek de betrouwbaarheid en de validiteit van de Cito-rekentoets?

Een toets is nooit volmaakt. Betrouwbaarheid en validiteit zijn altijd een kwestie van gradatie. Dit geldt ook voor officieel goedgekeurde toetsen. Het risico dat je afgaat op de uitkomsten en toch bedrogen uitkomt, valt nooit uit te sluiten. Daarom staat er in de *Standards for Educational and Psychological Testing* dat belangrijke beslissingen nooit uitsluitend gebaseerd mogen zijn op de uitkomsten van een enkele test (AERA, APA & NCME, 1999). In de praktijk wordt vaak tegen deze richtlijn gezondigd (Koretz, 2008).

De onbetrouwbaarheid van het eindexamen Nederlands

In de praktijk van het onderwijs lukt het ook officiële instanties niet altijd om aan de kwaliteitseisen te voldoen. Al jaren worden er door docenten vragen gesteld rondom de *manier waarop* de Nederlandse taalvaardigheid gemeten wordt. Onder betrokkenen bij dit vakgebied is de betrouwbaarheid van het eindexamen Nederlands berucht. Jaar in jaar uit is die betrouwbaarheid lager dan die van andere talen. De reeds eerder genoemde Van den Bergh (2006) analyseert deze kwestie. Tekstbegrip wordt zowel gemeten door leerlingen een traditionele tekstbegriptoets met meerkeuze vragen te laten maken als door hen een samenvatting te laten maken. Als deze twee meetmethoden afzonderlijk een betrouwbare meting van tekstbegrip zouden vormen, zouden de uitkomsten samen moeten hangen: wie hoog scoort op de meerkeuzevragen, zou ook hoog moeten scoren op de samenvatting maar dit is niet het geval. Er is maar weinig verband tussen de uitkomsten van beide meetmethoden terwijl ze allebei hetzelfde pretenderen te meten. Dat duidt op onbetrouwbaarheid. Als je op verschillende weegschalen weegt en ze zijn allebei betrouwbaar, zouden beide weegschalen immers dezelfde uitkomsten moeten geven.

Van den Bergh laat ook zien dat een half puntje meer of minder voor Nederlands wel degelijk uitmaakt voor het aantal geslaagden. Terwijl je – gezien de geringe betrouwbaarheid van het examen – op individueel niveau niet op een half puntje meer of minder zou moeten kijken. Dit onderstreept wat we al eerder geconcludeerd hebben, namelijk dat nooit in absolute zin aan de eisen van betrouwbaarheid en validiteit voldaan wordt. Het is belangrijk te blijven zoeken naar zo goed mogelijke meetmethoden, maar volmaakt zijn ze niet.

Gespreksvraag: Wat staat een volledige betrouwbaarheid of een volledige validiteit van toetsen in de weg?

Ongemerkt zijn we inmiddels overgeschakeld van toetsen die gemaakt worden door docenten naar toetsen die gemaakt worden door toetsspecialisten (psychometrici) of officiële instanties. Maar het is duidelijk dat als het maken van goede toetsen een uitdaging blijft voor mensen die de hele dag niet anders doen, dit zeker ook het geval is als leraren en docenten de toetsen zelf maken. Niettemin is het zinvol om de eigen toetsen kritisch te analyseren vanuit de eisen van betrouwbaarheid en validiteit. Het is logisch dat docenten die zich deze vragen stellen, het beter doen dan docenten die dat niet doen. Daarmee alleen al wordt de waarde van het instrumentele model duidelijk.

Toetsplannen

Doelstellingen zijn bedoeld om grotere hoeveelheden stof behapbaar te maken. In een opleidingsplan worden de doelstellingen in samenhang tot elkaar beschreven. Zo komt als het ware de landkaart van de opleiding tot stand, een overzicht van wat studenten moeten kennen en kunnen. Ook het toetsplan vormt een landkaart van de opleiding, maar benadert de opleiding vanuit een andere invalshoek. Tussen het opleidingsplan en het toetsplan is een nauwe relatie. Het zijn immers de doelstellingen uit het opleidingsplan die in het toetsplan getoetst worden. Toch vloeit het één niet automatisch uit het ander voort. De manier waarop getoetst wordt, beïnvloedt namelijk sterk datgene wat daadwerkelijk geleerd wordt. Onderzoek heeft aangetoond dat evaluatie – en dus ook de manier waarop getoetst wordt – de sterkst sturende factor is voor het leergedrag van leerlingen (Dochy & Janssens, 2003). Studenten bij wie tekstbegrip gemeten worden door middel van een tekst met meerkeuzevragen leren bijvoorbeeld hele andere dingen dan studenten bij wie tekstbegrip gemeten worden door middel van een samenvatting. Omdat de manier van evalueren zo bepalend is, is het nuttig om naast een opleidingsplan ook een toetsplan te hebben. Een toetsplan bevat een (impliciete) visie op toetsen waarmee nader gespecificeerd wordt wat de studenten dienen te kennen en kunnen als zij de opleiding afgerond hebben.

Riet is docent rekenwiskunde op de pabo die heel bewust van verschillende toetsvormen gebruik maakt. Digitale toetsen gebruikt ze om vast te stellen of studenten voldoende rekenvaardigheden hebben. Ze wil echter niet alleen dat studenten zelf kunnen rekenen, maar ook dat ze goed rekenonderwijs kunnen geven. Die twee vallen bepaald niet samen. Daarom vult ze de digitale toetsen aan met andere evaluatievormen. Studenten in het tweede jaar moeten een verantwoording maken bij een lessenserie. Ze moeten bijvoorbeeld aangeven hoe ze aansluiten bij hun observaties van de vorige les. Als niet alle lesdoelen bereikt waren, moeten ze daar in de huidige les op inspelen. In het derde jaar staat het individuele kind centraal. Studenten maken een groepsoverzicht en een individueel handelingsplan. Hierop geeft Riet feedback; ze beschouwt deze opdracht als een formatieve evaluatie. In de summatieve opdracht moeten studenten aangeven wat ze met feedback gedaan hebben. Riet heeft deze verschillende vormen van toetsen in haar toetsplan verantwoord.

Gespreksvraag: Maakt u gebruik van verschillende manieren van evalueren? Wat zijn uw ervaringen?

Toetsmatrijzen

Vanuit de helikopterblik die het toetsplan biedt, gaan we nu kijken naar afzonderlijke toetsen. We hebben al gezegd dat toetsen betrouwbaar en valide behoren te zijn. De stof is echter een groot geheel dat nooit in zijn volledigheid in de toets aan de orde kan komen. Wie een toets maakt, neemt een steekproef en kiest altijd bepaalde aspecten uit de leerstof terwijl andere aspecten blijven liggen. Iedereen weet dat een student het met deze keuze al dan niet kan treffen. Kiest de docent onderwerpen die de student toevallig weet (denk aan de koks in opleiding die een recept als tekst kregen), dan heeft de student geluk. Maar als er onderwerpen aan de orde komen die de student minder goed kent, heeft de student pech. Deze toevalsfactor in de keuze van vragen maakt het onvermijdelijk dat er ook een toevalsfactor is in de uitkomsten van de toets. Psychometrici duiden dit verschijnsel aan met *error*.

Error is onvermijdelijk maar er zijn wel hulpmiddelen bedacht om te bevorderen dat de toets representatief is voor de stof. Een toetsmatrijs is zo'n middel. Het maken van een toetsmatrijs of specificatietabel dwingt de docent om zich af te vragen of de verschillende onderdelen van de leerstof op een evenwichtige manier terugkomen in de toets en of de toetsvragen zich op het juiste niveau bevinden. Meestal bestaat de tabel uit twee onderdelen. Aan de ene kant is de inhoud van de leerstof aangegeven en aan de andere kant het beheersingsniveau zoals kennis, begrip of inzicht en toepassing. Soms ook wordt het relatieve belang van onderdelen in de tabel aangegeven. Er zijn ook toetsmatrijzen waarbij aan de ene kant de hoofdstukken uit de leerstof staan en aan de andere kant de doelstellingen. Al deze vormen zijn bedoeld om te bevorderen dat de onderdelen van de toets een representatieve steekproef vormen uit de totale leerstof.

Henri maakt op de lerarenopleiding basisonderwijs een toetsplan voor een minor waarin studenten zich specialiseren in Engels. Hij is enthousiast omdat op deze manier helder wordt wat belangrijk is binnen de minor. Dat schept duidelijkheid voor student en docent. Een voorwaarde is wel dat het docententeam een duidelijke en gedeelde visie heeft. In het toetsplan staan de hoofdonderwerpen die in de minor aan de orde komen en de manier waarop ze getoetst worden. In de toetsmatrijs wordt de stap gemaakt naar concrete toets-items die in een opgavenbank staan. Henri benadrukt dat niet alles in een toetsplan gevat kan worden. Plezier in taal en nieuwsgierigheid om te leren vind hij wel belangrijk, maar deze waarden staan niet in het toetsplan. Het toetsplan moet niet alles volledig vastleggen, vindt hij. Als je merkt dat studenten afhaken, moet er ruimte zijn om op die situatie in te spelen.

Discussievraag: Hoe zien uw toetsen eruit? Uit welke dimensies bestaan ze?
Hoe verdeelt u de doelstellingen over de stof?

Taxonomieën

Het woord taxonomie slaat op ordening of classificatie. Vaak gaat het daarbij om een hiërarchische ordening. De meest bekende, en in de onderwijskunde gezaghebbende ordening, is de taxonomie van Bloom. Jaarlijks vonden op Blooms faculteit uitgebreide examens plaats. Bloom maakte een grote verzameling toetsvragen om daarmee te meten of de doelstellingen werden gehaald. Hij ging er vanuit dat zijn vragenbank de examinatoren tijd zou besparen en zou bijdragen aan een gezamenlijke taal over toetsen in het onderwijs. Op grond van zijn onderzoek met deze vragenbank construeerde hij zijn indeling. De taxonomie gaf het niveau van de vragen aan (Bloom, Engelhart, Furst, Hill & Krathwohl, 1956). Later is zijn taxonomie gereviseerd door Krathwohl. De oorspronkelijke taxonomie van Bloom staat hieronder afgedrukt.

Oorspronkelijke taxonomie van Bloom (1956)
kennen
begrijpen
toepassen
analyseren
synthetiseren
evalueren

Krathwohl (2002) verbond de kenniscomponent met inzichten uit de cognitiewetenschap. Zo kwam hij tot de onderstaande taxonomie.

Taxonomie van Krathwohl voor het cognitieve domein (2002)
feitenkennis
conceptuele kennis (kennis van de manier waarop basale elementen in gehelen passen)
procedurele kennis (kennis van technieken en methoden en weten wanneer je ze moet toepassen)
metacognitieve kennis (kennis over cognitie in het algemeen en bewust zijn van en kennis van de eigen cognitie).

Krathwohl plaatste de verschillende begrippen in een matrix maar merkte daarbij wel op dat niet alle vakjes op een zinvolle manier te vullen zijn.

Aangepaste taxonomie van Bloom gecombineerd met die van Krathwohl (2002)						
	onthouden	begrijpen	toepassen	analyseren	evalueren	creëren
feitenkennis						
conceptuele kennis						
procedurele kennis						
metacognitieve kennis						

De vaardigheden links(boven) worden aangeduid met *lower order skills*; hoe meer we naar rechts bewegen, des te meer gaat het om *higher order skills*. Het zijn juist deze *hogere orde vaardigheden* die een indicatie vormen van zelfstandig en academisch denken wat karakteristiek voor het hoger onderwijs zou moeten zijn. Tegelijkertijd gaat het ook in het kleuteronderwijs op een bepaalde manier om hogere orde vaardigheden.

In hoofdstuk 2 hebben we gezien dat het instrumentele paradigma ook toegepast kan worden op het niet-cognitieve domein. Het is goed mogelijk om op een instrumentele manier om te gaan met sociale of affectieve doelstellingen. Krathwohl maakte ook een taxonomie voor het affectieve domein.

Taxonomie van Krathwohl (2002) voor het affectieve domein
ontvangen = bewust zijn van het bestaan van bepaalde ideeën.
reageren = blij van erkenning geven aan de ideeën in de vorm van een (minimale) respons.
waarderen = openlijk bereid zijn de ideeën als waardevol te beschouwen.
organiseren = bereid zijn de ideeën te koppelen aan bestaande persoonlijke ideeën/filosofieën tot een harmonieus geheel
karakteriseren = persoon gedraagt zich consequent conform de geïnternaliseerde ideeën

Gespreksvraag: Werkt u met beheersingsniveaus? Zo ja, welke?

Ryan en Deci (2000) maakten een taxonomie voor motivatie waarbij op het laagste niveau sprake is van extrinsieke motivatie en op het hoogste niveau van intrinsieke motivatie.



Taxonomie van Ryan en Deci (2000) voor motivatie

niet gemotiveerd, geen inzet voor de taak
doet de taak onder invloed van externe beloning en straf
doet de taak doordat beloning en straf verinnerlijkt zijn
doet de taak omdat de waarde van de taak wordt ingezien
doet de taak vanuit persoonlijke betrokkenheid
doet de taak vanuit persoonlijke betrokkenheid, interesse en plezier.

Gespreksvraag: Hoe zouden de cijfers uitpakken als u uw studenten voorname-lijk zou beoordelen volgens de taxonomie van Krathwohl of van Deci en Ryan?

Weinig taxonomieën zullen de gebruiker het gevoel geven compleet dekkend te zijn. Wie ermee aan het werk gaat, zal vroeg of laat vragen hebben. Is het wel zo terecht is dat het ene niveau ‘moeilijker’ is dan het andere? Hoe kan ik dit begrip in mijn context toepassen en is dit wel zinvol? Taxonomieën zijn geen middelen waardoor de docent niet meer hoeft na te denken. Integendeel, het zijn middelen die kritisch denken over onderwijs en toetsing *bevorderen*. Ze helpen bij het maken van een toetsmatrijs en bevorderen daarmee dat de doelstellingen op een evenwichtige manier geëvalueerd worden.

Vaststellen van de resultaten

Als de toets gemaakt is en de toetsmatrijs aangeeft dat met de nodige factoren rekening gehouden is, kan de toets worden afgenomen. Toetsen met gesloten antwoordcategorieën, bijvoorbeeld meerkeuzetoetsen, kunnen door een computer nagekeken worden. Terwijl enerzijds het maken van goede meerkeuzevragen de docent veel tijd kost, kan de computer de docent anderzijds tijd besparen door het nakijkwerk over te nemen.

Meerkeuzevragen kunnen nakijkwerk besparen, maar werken reproductief en oppervlakkig leren in de hand (Struyven, Dochy, & Janssens, 2003). Er hoeft namelijk maar vrij weinig uit de studenten zelf te komen; ze hoeven het juiste antwoord alleen maar te herkennen. Bij essayvragen moeten studenten zelf hun antwoorden formuleren. Deze antwoorden moeten echter door docenten beoordeeld worden, met als risico dat de ene docent een antwoord anders inschat dan de andere of bij de ene student strenger is dan bij de andere.

Een nakijkmodel is bedacht om de kwaliteit van de beoordeling (van met name essayvragen) te bevorderen. Daarin staat welke aspecten de student in

het antwoord moet betrekken. Het maken van een nakijkmodel kan ook helpen om de kwaliteit van toetsvragen te verbeteren. Wie bijvoorbeeld bepaalde zaken van een student wil horen, moet die zaken expliciet in de vraag aan de orde laten komen. Wie drie redenen wil horen, kan beter om drie redenen vragen dan globaal de waarom-vraag stellen.

Hoewel meerkeuzevragen geslotener zijn dan essay-vragen, gaat het instrumentele model altijd uit van een zekere geslotenheid. De bedoeling is immers dat de leerdoelstellingen getoetst worden. De leerdoelstellingen stonden van tevoren vast; van studenten wordt gevraagd dat zij er blijk van geven dat zij met dit arsenaal uit de voeten kunnen. Alles draait om 'wat er al was'. Meestal gaat het daarbij om kennis op reproductief of begripsniveau. Vanuit het instrumentele paradigma ontstaat geen stimulans om met iets nieuws of origineels aan te komen.

Waar is de toets voor bedoeld?

Gestandaardiseerde toetsen behoren een omschreven doel te hebben. Tegenwoordig bestaan er voor diverse vakken (taal en rekenen) referentieniveaus. Een toets kan bedoeld zijn om het referentieniveau vast te stellen. Wanneer uit deze toetsen blijkt dat vrijwel alle studenten de toets halen, is er reden voor feest. De bedoeling was immers dat iedereen dit streefniveau zou halen. De Cito-eindtoets is niet bedoeld om een niveau te meten dat alle leerlingen behoren te halen, maar is juist bedoeld om onderscheid te maken tussen leerlingen die een bepaald niveau wél halen en leerlingen die dat niet halen. De Cito-eindtoets moet immers een advies geven voor het vervolgonderwijs. Vragen die door een groot deel van de leerlingen goed gemaakt worden, zijn voor referentieniveau-toetsen wél geschikt, maar voor de Cito-eindtoets ongeschikt. Voor de Cito-eindtoets gaan toetsmakers doelbewust op zoek naar vragen die alleen door een deel van de leerlingen goed gemaakt worden. De Cito-eindtoets is – zo heet dit in vaktermen – bedoeld om te discrimineren. Dit woord betekent in dit geval niet 'het bewust benadelen' van studenten, maar het scherp onderscheid maken, in dit geval tussen leerlingen die het vwo, de havo of het vmbo aankunnen. Niet alle toetsen hebben dus hetzelfde doel. Het is belangrijk om dit doel in de gaten te houden.

Ook de toetsen van de leerlingvolgsystemen die in het basisonderwijs worden gebruikt zijn met een eigen bedoeling ontworpen. Volgens de wetten van de psychometrie mag een test niet voor iets anders gebruikt worden dan waarvoor hij bedoeld is. Resultaten van toetsen die bedoeld zijn om

informatie te verschaffen over individuele leerlingen, mogen niet zomaar gebruikt worden om uitspraken te doen over de kwaliteit van de school of de leerkracht (Koretz, 2008). In de praktijk wordt vaak tegen dit beginsel gezondigd.

Gespreksvraag: Welke toetsen in uw opleiding zijn bedoeld om te 'discrimineren'? Kent u voorbeelden in uw opleiding van toetsen die voor iets anders gebruikt worden dan waarvoor ze bedoeld zijn?

Kenmerken van instrumentele manier van toetsen

We schetsen samenvattend de kenmerken van de instrumentele manier van toetsen. Het instrumentele paradigma gaat uit van vooraf gedefinieerde leerdoelstellingen en meet zo objectief mogelijk of de doelstellingen zijn gehaald. De validiteit en de betrouwbaarheid zeggen iets over de kwaliteit van de toets. Het liefst maakt het instrumentele paradigma gebruik van gestandaardiseerde toetsen. De standaardisatie wijst erop dat aan bepaalde minimumkwaliteitseisen wordt voldaan, zowel als het gaat om de aard van de toets als om de normering ervan. Gestandaardiseerde toetsen zijn methode-overstijgend in die zin dat ze proberen aan te sluiten bij meerdere uitvoeringen van het onderwijs. Dat plaatst deze toetsen wel op afstand van het daadwerkelijk gegeven onderwijs. Als de docent de toets zelf maakt, bestaat de mogelijkheid om de toets af te stemmen op wat is onderwezen. Instrumentele toetsen – of ze nu gestandaardiseerd zijn of niet – staan ook los van het onderwijs in die zin dat er aparte onderwijstijd voor ingeruimd wordt. Toetsen maken het mogelijk om studenten met elkaar te vergelijken. Het kan zijn dat de resultaten onverwachte patronen laten zien zijn. De resultaten kunnen dan gebruikt worden om de kwaliteit van de toetsen te verbeteren.

Toetsen geven informatie over de kwaliteit van het leerproces maar het is de vraag in hoeverre ze een houvast bieden om de instructie te verbeteren. Op grond van toetsen is wel vast te stellen *dat* een leerling iets niet beheerst, maar het wordt minder duidelijk *waarom* een leerling iets niet beheerst. In het instrumentele model ligt de nadruk meer op summatieve toetsing (die plaatsvindt aan het einde van een deel van het leerproces) dan op formatieve toetsing (die plaatsvindt tijdens het leerproces en bedoeld is om de leerling feedback te geven). Meestal gaan die toetsen over kennis op reproductief of begripsniveau. De evaluatie vindt uitsluitend door de leraar plaats; de leerling zelf maakt de toets maar speelt geen rol in de beoordeling ervan.

Deze kenmerken komen overeen met wat Dochy en Janssens (2003) aanduiden met de testcultuur (zie kader hieronder).

Kenmerken van de testcultuur (Dochy & Janssens, 2003)

1. Toetsen staan los van het onderwijs
2. Klemtoon op summatieve evaluatie
3. Evaluatie uitsluitend door de leraar
4. Het toetsen van vooral kennis
5. Gestandaardiseerde toetsen
6. Leerlingen met elkaar vergelijken

Gespreksvraag: Hoe staat u tegenover de instrumentele visie op toetsen?

Slot

In de onderwijskunde heeft het instrumentele paradigma lange tijd gedomineerd. Eind jaren tachtig drong het inzicht door dat de instrumentele visie op het vaststellen van de kwaliteit van leren een keerzijde heeft. De toetscultuur bleek een eigen onderwijswerkelijkheid te scheppen, waarbij juist *higher order skills*, de denkvaardigheden die bij uitstek kenmerkend zijn voor kritisch denken op hoog niveau, het moesten ontgelden. Er klonk nog meer kritiek. Deze manier van onderwijs was niet erg motiverend. De leerstof werd parades van bovenaf bij de studenten geparachuteerd. De eigen leef- en denkwereld van studenten speelde geen rol en de afstand tot de beroepspraktijk of de praktijk van het leven was groot. Leerlingen leerden bovendien onvoldoende zelfstandig studeren. Daarom ontstond er een tegenbeweging. Daarin werd er gezocht naar meer authentieke vormen van toetsing. Deze beweging wordt in het volgende hoofdstuk geschetst.

Hoofdstuk 4

Van instrumentele toetsing naar assessment

De instrumentele benadering was bedoeld om het leren van studenten te structureren en zo het onderwijs te verbeteren. Gaandeweg werd echter duidelijk dat er van die verbetering niet genoeg terecht kwam (Wiliam, 2011). Het leek zo'n mooi ideaal: neutrale, objectieve middelen om vast te stellen of de onderwijsdoelstellingen gehaald waren. Maar de middelen bleken minder neutraal en objectief te zijn dan gedacht. Er werd duidelijk dat doelen en middelen nauw met elkaar verbonden zijn (Biesta, 2007). Toetsen bleken een eigen (onderwijs)werkelijkheid te scheppen die niet altijd wenselijk was.

Eerst laten we de bezwaren tegen de instrumentele onderwijscultuur de revue passeren. Onder de noemer 'evalueren om te leren' kwam een nieuwe beweging tot stand met nieuwe aanpakken om te evalueren. Deze waren ontleend aan het vormingsparadigma. We eindigen dit hoofdstuk met een voorbeeld waarin een cursus verbeterd werd door elementen uit het vormingsparadigma erin te verwerken.

Bezwaren tegen de instrumentele onderwijscultuur

In de loop van de tijd zijn allerlei nadelen van de instrumentele onderwijscultuur naar voren gekomen. Hier onder staan ze op een rijtje.

Weinig handvatten om het leerproces te bevorderen

Toetsing mocht dan 'objectieve' informatie leveren; erg veel aanknopingspunten om het leerproces te bevorderen boden ze niet. Veel tijd en aandacht gingen uit naar het meetproces. Beslissingen over selectie konden zo keurig verantwoord worden. Maar schoten studenten daar iets mee op? Verbeterde het onderwijs daardoor? Toetsen leverden slechts oppervlakkige informatie. Er kon wel aangegeven worden of student een doelstelling gehaald hadden

maar die informatie kwam te laat, namelijk na afloop van het leerproces. De instructie liep daardoor permanent achter de feiten aan. Wat er mis was gegaan als de doelstellingen niet waren gehaald, bleef onduidelijk. De *waarom*-vraag bleef onbeantwoord. De instrumentele benadering bood dus te weinig handvatten om het leerproces te bevorderen.

Gespreksvraag: In hoeverre geven toetsresultaten u handvatten om uw onderwijs af te stemmen op wat studenten nodig hebben?

Reductie tot wat makkelijk meetbaar is; weinig ruimte voor echte intellectuele vaardigheden en creativiteit

Niet alles wat in het onderwijs van waarde is, is gemakkelijk te meten. We zagen al dat het instrumentele paradigma de nadruk legt op 'wat er al was'. Studenten maken zich een van tevoren uitgestippeld programma eigen. Hun eigen vragen komen in dit verhaal niet voor. Wat studenten moeten laten zien, is vooraf al vastgelegd in een nakijk- of antwoordmodel. Originaliteit en creativiteit passen niet binnen dit kader. We zagen al dat evaluatie de sterkst sturende factor is voor het leergedrag (Dochy & Janssens, 2003). Het instrumentele model legt vooral de nadruk op de lagere niveaus in de taxonomie van Bloom, namelijk onthouden, begrijpen en toepassen. De hogere niveaus – synthetiseren en creëren – zijn niet in een antwoordmodel te vatten. Die raken dus systematisch op de achtergrond. Anders gezegd: reproductieve vaardigheden komen uit de verf, maar productieve vaardigheden niet. Dit heeft ernstige gevolgen voor de inhoud van het onderwijs.

In het dagelijks leven bestaan convergente en divergente problemen. Convergente problemen kennen verschillende oplossingsmethoden maar slechts één uitkomst. Een rekensom, bijvoorbeeld, kan opgelost worden door verschillende oplossingsmethoden toe te passen; de uitkomst van de som is echter altijd dezelfde. Een rekensom is daarom een convergent probleem. Bij divergente problemen kunnen heel veel verschillende antwoorden goed zijn. Wie een huis met vier slaapkamers wil bouwen, kan dit op verschillende manieren doen die allemaal aan de gestelde eisen voldoen. Een beschrijving van hoe het Paleis op de Dam eruit ziet kan op talloze verschillende manieren aan de eisen voldoen en is daarmee dus een divergent vraagstuk.

Om divergente problemen op te lossen, zijn persoonlijke betrokkenheid en creativiteit nodig. Het is echter niet mogelijk om al deze factoren in een

antwoordmodel te vatten, want de manieren om de opdracht uit te voeren zijn talrijk. Er is sprake van minder goed, goed, of beter, maar wat daarmee bedoeld wordt en hoe het onderscheid wordt bepaald, is niet eenvoudig aan te geven. Het instrumentele paradigma kan dus niet goed met divergente problemen uit de voeten.

De verleiding om divergente problemen te laten liggen – want het moet toch duidelijk zijn wat goed is en wat fout – is groot. Het onderwijs laat dan de doelstellingen die niet tot gemakkelijk meetbaar gedrag leiden links liggen, terwijl de onderwijsinhoud gereduceerd wordt tot doelstellingen die wel makkelijk meetbaar zijn. Leren om te leven raakt dan erg op de achtergrond. Immers past een groot deel van de levensechte problemen niet in de instrumentele cultuur (zie kader hieronder).

Vervolgens dreigt een vicieuze cirkel te ontstaan. Omdat bij de beoordeling voornamelijk op gemakkelijk meetbare factoren gelet wordt, is de instructie ook op deze makkelijk meetbare – en vaak platte – factoren gericht. De inhoud van het onderwijs wordt daarmee verengd tot wat een toets kan meten. Het is navrant dat de instrumentele poging om het onderwijs te verbeteren er juist in resulteert dat de *higher order skills*, de ‘echte intellectuele vaardigheden’, op de achtergrond raken.

Divergente problemen zijn moeilijk te beoordelen. De beoordelingscriteria werken soms sterk reducerend. Om tekstbegrip te meten, wordt in het eindexamen Nederlands tekstbegrip getoetst door studenten een samenvatting van een tekst te laten maken. Maar hoe moet die samenvatting beoordeeld worden? Het zou voor de hand liggen hier criteria voor te gebruiken die met tekstbegrip te maken hebben. Die zijn echter heel moeilijk in algemene correctievoorschriften te vatten. In de praktijk gebeurt er iets anders. Aspecten van de tekst, die gemakkelijk in objectieve criteria te vangen zijn, krijgen een hele belangrijke rol. Bij het eindexamen Nederlands wordt er bij de samenvatting – die bedoeld is om tekstbegrip te meten – gelet op het aantal woorden. Systematisch worden punten afgetrokken als leerlingen teveel woorden gebruiken. Dit brengt hoogleraar Van den Bergh (2006 – zie ook de kaders in het voorgaande hoofdstuk) tot de verzuchting dat het dwaasheid is om te veronderstellen dat naarmate leerlingen iets beter begrijpen, ze kortere antwoorden geven. Bij de beoordeling wordt ook gelet op spelling. Weer iets dat met tekstbegrip niets te maken heeft maar zich in feite op een ‘lager’ niveau van complexiteit bevindt.

Deze manier van beoordelen heeft vervolgens consequenties voor de inhoud van het onderwijs. Voordat leerlingen het eindexamen maken, moeten ze niet alleen getraind worden in samenvattingen maken maar vooral in samenvattingen maken *die voldoen aan de beoordelingscriteria*. Leerlingen die veel aandacht willen besteden aan een vlotte schrijfstijl, moeten weten dat ze hier maar weinig punten mee kunnen verdienen en dus beter kunnen letten op het aantal woorden en de spelling. Uiteindelijk drijft de manier van toetsing de studenten bij de doelstellingen vandaan. De aandacht gaat niet uit naar tekstbegrip en schrijfstijl maar naar eenvoudiger zaken als het aantal woorden, de spelling en het benoemen van een aantal hoofdpunten van de tekst.

Gespreksvraag: Op welke manier gaat het in uw vak om vragen waarbij de antwoorden niet zo gemakkelijk onderverdeeld kunnen worden in goed of fout? Hoe gaat u om met moeilijk meetbare doelstellingen?

Andersoortige leerprocessen komen niet uit de verf

Uit het bovenstaande blijkt dat de creatieve aspecten – die in vrijwel alle vakken een rol spelen – het onderspit dreigen te delven. Dit heeft betrekking op de creatieve aspecten van allerlei vakken zoals het vreemde talenonderwijs (denk aan literatuur) of de geschiedenis (denk aan het gezamenlijk puzzelen aan een vredesvraagstuk). Daarnaast zijn het natuurlijk ook de creatieve vakken zelf – muziek, tekenen, drama enzovoort – die op het tweede plan terechtkomen. Ook vakken zoals godsdienst, levensbeschouwing en maatschappijleer, die niet binnen de geldende meetkaders vallen, worden vakken achter de streep (zie kader op pag. 55).

Het instrumentele paradigma brengt een verborgen curriculum met zich mee. Daarmee bedoelen we dat er naast de expliciete doelstellingen ook een boodschap op meta-niveau aan leerlingen wordt meegegeven. Het verborgen curriculum bij de instrumentele benadering geeft leerlingen de boodschap dat het in het onderwijs voornamelijk gaat om resultaten die uitgedrukt kunnen worden in cijfers. Deze meta-boodschap wringt echter met de eigen aard van vakken als godsdienst en maatschappijleer. Daar gaat het eerder om het leren stellen van de juiste vragen dan om het geven van de juiste antwoorden. Doordat het hele onderwijs is doordeesemd met de impliciete boodschap dat het gaat om ‘juiste antwoorden’ komen vakken die andere accenten leggen in een uitzonderingspositie terecht. Het verborgen curriculum zet de studenten voortdurend op een ander spoor. Zo kan het verborgen curriculum van

het instrumentele paradigma ertoe bijdragen dat andersoortige leerprocessen minder uit de verf komen. Dit verborgen curriculum versterkt het tweederangs karakter van vakken waarin het gaat om morele vorming, oordeelsvorming of artistieke vorming.

Ook dit kan weer een cirkel veroorzaken. Als leerprocessen minder goed uit de verf komen, moeten docenten actie ondernemen. Niet zelden gebeurt dit door gebruik te maken van het instrumentele paradigma. Men grijpt dan naar het aloude recept van kennis en vaardigheden. Voortaan draaien ook godsdienst en maatschappijleer om meetbare kennisdoelen. Zo is de cirkel rond en wordt het curriculum enger en enger.

Wat hier boven geïllustreerd is aan de hand van de vakken godsdienst en maatschappijleer, geldt op meerdere vlakken. We kunnen ook denken aan spelend leren dat voor kinderen een belangrijke manier is om zich te ontwikkelen, maar niet tot meetbare resultaten leidt en daarom in de verdrukking komt. Daarmee wordt onrecht gedaan aan de enorme educatieve waarde van spelen, niet alleen voor kleuters. Aan het eind van dit hoofdstuk volgt een voorbeeld van cursus in het hoger onderwijs die mede dankzij spelend leren verbeterd werd.

Ook op christelijke scholen is het vak godsdienst vaak een tweederangs vak. Om anderen te kunnen begrijpen is het ook nodig om te begrijpen wie jezelf bent. Voor een groot deel is de traditie waarin je bent opgegroeid een onbewuste filter die de blik op de ander kleurt. Zicht krijgen op de eigen traditie, deze traditie leren wegen in het licht van grote waarden enerzijds en mensen vanuit andere tradities in hun anders-zijn leren aanvaarden, hebben alles met elkaar te maken. Het gaat om oordeelsvorming en morele vorming. Het vak godsdienst gaat meer gepaard met vragen leren stellen dan met de juiste antwoorden kunnen geven. Deze visie op het vak past minder goed in een meetcultuur.

Gespreksvraag: Welke impliciete boodschappen op meta-niveau zouden in uw onderwijs aanwezig kunnen zijn? U kunt hierbij ook aan positieve factoren denken.

Docent rekenen pabo: “Leerlingen moeten natuurlijk bepaalde basisvaardigheden aanleren, maar als je alleen focust op basisvaardigheden, mis je de essentie van wiskunde. De essentie van wiskunde is niet het opdreunen van tafels,

maar herkennen dat in concrete situaties wiskundige problemen schuil gaan. Leerlingen moeten het zelfvertrouwen ontwikkelen om zo'n probleem op te halen en erover na te gaan denken. Ik wil liever dat ze nadenken dan dat ze het trucje opdreunen. Hier is ook kritiek op. 'Er wordt te weinig geoefend', zeggen ze dan. En natuurlijk moeten er momenten van oefenen zijn. Kinderen heerlijk sommetjes laten maken is ook prima, dat moet je niet gelijk af doen met kinderen bezig houden. We zijn daarom niet tegen oefenen, maar kinderen laten oefenen is didactisch niet zo'n kunst. Kinderen leren nadenken is veel moeilijker. Hier besteden we dus veel aandacht aan. Hoe lastig het is ook is om een interactieve rekenles te geven, het bevordert wel dat leerlingen leren nadenken."

Onderwijs negeert de eigen persoon en de talenten van de student

Een volgend bezwaar tegen het instrumentele paradigma heeft te maken met de rol van de student zelf. De onderzoeker Eisner (2004) wees al in 1976 op het destructieve effect van het instrumentele paradigma voor de creatieve vakken en voor de moeilijk meetbare aspecten van deze vakken. Hij benadrukte ook dat het gedrag van studenten niet zo gemakkelijk voorspelbaar is als in het instrumentele paradigma wordt aangenomen. In de alledaagse werkelijkheid zijn studenten vaak met andere dingen bezig dan met wat de doelstellingen voorschrijven. Goede leraren spelen daarop in. Vanuit de belevingswereld van de student maken ze een beweging om de studenten verder te helpen. Echter, het staat bepaald niet vast dat deze tocht precies bij de voorgeschreven doelstellingen uitkomt. Moet de docent daar toch per se aankomen, dan wordt hij of zij gedwongen om over de hoofden van de studenten heen te praten. Dat kan niet de bedoeling zijn.

Het instrumentele model dreigt de studenten zelf dus te vergeten. Wie ze zijn en hoe ze in het leven staan, komt in het instrumentele model niet voor. De eigen vragen van studenten komen er stiefmoederlijk vanaf. Studenten worden niet uitgenodigd om op zichzelf te reflecteren. Zelfkennis en eigen verantwoordelijkheid voor het leerproces worden weinig gestimuleerd. Volgens een belangrijke criticaster van het instrumentele paradigma, Alfie Kohn (2011), leren cijfers en toetsen studenten vooral dat ze aan verwachtingen van anderen moeten voldoen. Het leert hen niet zelf nadenken.

Gespreksvraag: Op welke manier speelt de leef- en denkwereld van studenten een rol in uw onderwijs?

Studenten leren niet omgaan met levensechte problemen

Een volgend bezwaar tegen het instrumentele paradigma heeft te maken met de bruikbaarheid van de kennis en vaardigheden die studenten op deze manier aanleren in het dagelijks leven. Onderwijsdoelstellingen worden opgedeeld in meetbare subdoelen die zo objectief mogelijk getoetst worden, bijvoorbeeld met een meerkeuzetoets. Het echte leven in de beroepspraktijk bestaat in zekere zin wellicht wel uit meerkeuzevragen, maar de antwoorden worden niet gevormd door vier opties waarvan er één goed is. In het echte leven worden kennis en vaardigheden volop gebruikt, maar zelden getoetst door middel van examens.

Toch hecht het instrumentele paradigma veel waarde aan die examens. We hebben gezien dat opgaven op een toets altijd een steekproef vormen uit een groter geheel van leerstof. Als er echter grote druk op de toetsresultaten komt te staan, zal alle energie zich richten op de voorbereiding op het type toetsvragen dat de toets bevat. Er wordt volop geoefend met het type opgaven van de toets. Niet zozeer om vaardigheden onder de knie te krijgen die noodzakelijk zijn voor het dagelijks leven, maar om te begrijpen aan welke beoordelingscriteria voldaan moet worden en vertrouwd te raken met het type instinkers dat gebruikt wordt. Als studenten met die criteria vertrouwd zijn, scoren ze beter. Het is echter de vraag wat die resultaten over de kwaliteit van het leerproces zeggen (zie kader hieronder).

Duiden hogere toetsresultaten op beter onderwijs? Dat valt te bezien. Vrij snel nadat in het basisonderwijs Opbrengstgericht Werken (OGW) werd geïntroduceerd, kwamen er positieve berichten van het ministerie van OCW dat de rekenprestaties verbeterd waren (Ministerie OCW, 2012). Is het denkbaar dat het onderwijs binnen anderhalf jaar tijd zomaar beter is geworden? Er zijn meerdere manieren om deze resultaten te verklaren. Allereerst kan niet uitgesloten worden dat de kwaliteit van onderwijs inderdaad verbeterd was. Maar evenmin kan worden uitgesloten dat leraren zich er door de introductie van OGW meer bewust van geworden waren dat er op de resultaten gelet ging worden en daarom hun onderwijs – ten koste van andere aspecten – beter bij de toetsen hadden laten aansluiten. *Teaching-to-the-test* kan de oorzaak zijn dat de resultaten verbeterden. De vraag is of het onderwijs daar kwalitatief beter van wordt.

Zo ontstaat een soort pseudowerkelijkheid. Studenten ontwikkelen kennis en vaardigheden die vooral in het onderwijs van toepassing zijn maar dit onderwijs heeft weinig meer met het echte leven te maken. Onderwijs dreigt daarmee te verworden tot een zinloze kwalificatiemachine die statussymbolen aflevert

aan hen die daar gevoelig voor zijn. De groep leerlingen die er binnen dit systeem niet in slaagt om te scoren, dreigt af te haken of is in ieder geval dolblij na verloop van tijd van deze kwelling verlost te zijn. Vooral middelmatige en zwakke leerlingen worden door het systeem van constante vergelijking met betere leerlingen gedemotiveerd (Dochy & Janssen, 2003; Onderwijsraad 2013).

Een docent geschiedenis geeft een les in het vmbo over Alexander de Grote. Hij is enthousiast over dit onderwerp omdat hij aan de hand van strijdverhalen een lastig stuk geschiedenis kan duidelijk maken. Heldenverhalen spreken zijn leerlingen aan. Bij het maken van de toets heeft hij een paar open vragen maar ook een aantal meerkeuzevragen. Zoals de volgende:

Alexander van Macedonië heette ook:

- a. Alexander de Grote
- b. Alexander de Geweldige
- c. Alexander de Overwinnaar

Als deze docent met een paar collega's over de toets praat, worden vragen opgeworpen. Voegen de strijdverhalen over Alexander werkelijk iets toe aan hun begrip van de wereld? En heeft hij met de toetsitems over de naam en Alexanders daden relevante kennis getoetst? Ook al vinden de leerlingen strijdverhalen spannend, levert de aandacht hiervoor iets op voor hun eigen leven? Alexander was geen lieverdje. Zouden de houding en de moraal van Alexander niet interessanter zijn om aan de orde te stellen dan feitjes over zijn leven? Maar stel dat de docent een open gesprek met de leerlingen over Alexanders houding en moraal centraal zou stellen en een relatie zou leggen met de helden van de leerlingen, hoe zouden zijn toetsen er dan uit moeten zien?

Gespreksvraag: kunt u suggesties doen aan deze docent geschiedenis? Welke mogelijkheden en problemen doemen op?

Naar een nieuw paradigma

Over het instrumentele paradigma valt veel goeds te zeggen. Tegelijk is de kritiek ook niet mals. Het is niet verwonderlijk dat deze kritiek de slinger deed omslaan naar de andere kant.

Gespreksvraag: Hoe staat u nu tegenover het instrumentele paradigma?



Evaluatie binnen het vormingsparadigma

Wie zich de kritiek op het instrumentele paradigma aantrekt, wordt in de richting van het vormingsparadigma gemanoeuvreerd. In vorige hoofdstukken zijn kenmerken van het vormingsparadigma toegelicht. Vanuit deze kenmerken gaat deze paragraaf over de manier waarop binnen vormend onderwijs geëvalueerd kan worden. Kan evaluatie een rol spelen om het leerproces te bevorderen?

We hebben gezien dat het instrumentele paradigma weinig ruimte biedt voor de leef- en denkwereld van studenten en hun eigen vragen. Het vormingsparadigma benadrukt juist het belang van die eigen vragen en eigen talenten van studenten.

Het instrumentele paradigma creëert een eigen werkelijkheid van waarheden die eigenlijk alleen nog maar relevant is binnen de scholen terwijl het vormingsparadigma het belang benadrukt van leren voor het leven.

Met het instrumentele paradigma kwamen morele vorming, oordeelsvorming en artistieke vorming op de tweede rang terecht. Het vormingsparadigma plaatst deze leerprocessen juist op de eerste rang.

Ook divergente vaardigheden en hogere orde vaardigheden kwamen bij het

instrumentele paradigma op het tweede plan terecht omdat ze leiden tot uitkomsten die niet te vatten zijn in termen van goed en fout. Het vormingsparadigma benadrukt echter het belang van open leerprocessen.

Vorming is bedoeld voor alle studenten die op verschillende manieren en in verschillende tempo's hun talenten ontwikkelen. Vanuit het vormingsperspectief heeft evaluatie een informerende functie maar geen selecterende functie. Dit klinkt alsof het vormingparadigma hét antwoord is. Of dit ook zo is, zullen we verderop in dit boek ontdekken. Vooralsnog proberen we zo praktisch mogelijk te maken hoe elementen uit het vormingsparadigma het (instrumentele) onderwijs kunnen verrijken.

Evalueren om te leren

Tegenover evaluatie *van* het leerproces in de instrumentele benadering, staat evaluatie *om* het leerproces ofwel evalueren om te leren (Wiliam 2011; Casteleijns, Segers & Struyven, 2011). Daarbij is evaluatie niet zozeer een middel om vast te stellen in hoeverre de studenten de doelstellingen gehaald hebben, maar een middel om het leerproces in het algemeen te bevorderen. Instructie en evaluatie vormen daarbij twee kanten van dezelfde medaille. Docenten organiseren de instructie zodanig dat voortdurend duidelijk wordt hoe studenten vanuit hun eigen leef- en denkwereld met de stof bezig zijn. Dat kan soms op een simpele manier door langslopen en te zien hoe leerlingen werken. Voor het basisonderwijs zijn allerlei observatiesystemen ontwikkeld. Soms is het minder eenvoudig en moet de didactische aanpak gericht worden op het verhelderen van de concepties van studenten. Het gaat bij het verzamelen van deze informatie niet in eerste instantie om de vraag of een antwoord goed is of fout, maar vooral om diepgaande, kwalitatieve informatie over de concepties van studenten. Daar stemt de docent de instructie vervolgens op af (zie kader op pag. 61).

Interactie tussen docent en studenten geeft de docent inzicht in de manier waarop studenten de leerstof opvatten, maar is toch vaak beperkt. Vaak zijn het steeds dezelfde studenten die reageren terwijl anderen hun mond houden. Er zijn manieren denkbaar waarop de docent het onderwijs zodanig inricht dat de manier waarop studenten met de leerstof omgaan veel systematischer naar voren komt (zie het voorbeeld aan het einde van dit hoofdstuk en ook het experiment van Wiliam, kader op p. 79). Dit geeft de docent voortdurend de mogelijkheid om in te spelen op wat studenten inbrengen. Evaluatie is dus niet gescheiden van het leerproces maar vindt tijdens het leerproces plaats en is ingebed in een visie op onderwijs en onderdeel van het leerproces.

Deborah Ball is onderwijsonderzoeker in de Verenigde Staten, geeft wiskunde in het basisonderwijs en is expert in het omgaan met grote groepen. Op de universiteit staat ze doorgaans voor collegezalen met zo'n tweehonderd studenten. De opzet van haar colleges is helemaal gestructureerd rondom vragen. Ze heeft deze vragen van tevoren zorgvuldig uitgekiend. Ze zijn ontworpen om studenten te helpen om een bepaalde gedachtegang op te bouwen. Oogcontact vindt ze heel belangrijk. Ook in grote groepen hebben studenten goed door dat ze door oogcontact bij de les gehouden worden. Meestal vraagt ze de studenten eerst om haar vragen met hun burens te bespreken. Vervolgens vraagt ze naar groepsstandpunten. Dat is veiliger dan persoonlijke standpunten. Deborah heeft geleerd om studenten de tijd te geven om na te denken. Soms wacht ze heel lang. Als er geen vingers komen, doorbreekt ze de soms pijnlijke stiltes met de opmerking dat het goed is om na te denken. Bij de eerste vinger die dan toch komt, bedankt ze deze persoon en nodigt ze meer mensen uit om hun vinger op te steken door te zeggen dat ze nog wat meer bedenktijd geeft. De antwoorden van studenten geven Deborah informatie over hoe studenten denken. Vanuit die antwoorden beweegt ze naar de volgende vraag.

Op <http://www.youtube.com/watch?v=pPhCmvpPorU> is te zien hoe zij werkt.

Hoewel het vormingsparadigma meer ruimte vraagt voor de vrije invulling door de studenten, staan docenten niet werkloos aan de kant. Ook zij dragen verantwoordelijkheid voor het leerproces. Daarmee wijkt het vormingsparadigma af van het zelfontplooiingsparadigma. Het vormingsparadigma laat studenten zelf wel een belangrijke rol spelen, niet alleen door activerende werkvormen aan te bieden, maar ook door hen een actieve rol te geven bij de evaluatie van het leerproces. De bedoeling was immers ook dat studenten leren om een oordeel te vormen en dit oordeel te funderen. Studenten worden daarom zowel uitgenodigd om feedback te geven op elkaar als om te reflecteren op hun eigen leerproces. Niet alleen de docenten, maar ook de studenten moeten zicht hebben op het leerproces om te weten wat de volgende stap kan zijn. De eigen verantwoordelijkheid van studenten voor zichzelf en voor elkaar wordt dus nadrukkelijk benoemd. De docenten, op hun beurt, proberen de studenten te bereiken. Het is niet ongebruikelijk dat docenten hun doelstellingen tussentijds bijstellen. Wat er geleerd wordt, wordt samen bepaald en hangt mede af van de groep.

Ook in het vormingsmodel komt er een moment waarop het leerproces wordt afgerond of waarin vaardigheden gedemonstreerd moeten worden. Dit komt zoveel mogelijk overeen met wat de student 'in het echt' ook moet kunnen. Daarom wordt er gesproken van *authentic assessment*.

In de praktijk van het hbo en het mbo is men daarmee vertrouwd geraakt. Competenties worden in de praktijk getoetst. Studenten verzamelen zelf informatie en maken een portfolio. Deze vormen van evaluatie vinden plaats binnen de onderwijsvisie van competentiegericht opleiden, die verwant is aan het vormingsparadigma (De Muynck & Roeleveld, 2005).

Gespreksvraag: Bent u het eens met de stelling dat competentiegericht opleiden past binnen het vormingsparadigma?

Een voorbeeld: vormende elementen toevoegen

Het onderstaande voorbeeld laat zien hoe een traditionele cursus werd verbeterd door inspiratie te putten uit het vormingsparadigma. Het is een waar gebeurd verhaal; niet alle kenmerken van het vormingsparadigma komen erin voor.

Docenten van de cursus 'Omgaan met sociaal-culturele verschillen' zijn gewend om tijdens de colleges de literatuur te bespreken waarover studenten aan het eind van de cursus een tentamen maken. Ze proberen studenten zoveel mogelijk bij de stof te betrekken maar zijn niet tevreden. Niet alle studenten lopen warm voor de discussies die ze proberen los te maken. Als het gaat om de vraag hoe je met sociaal-culturele verschillen omgaat in de klas, verzandt de discussie in open deuren als 'ieder kind is gelijkwaardig'.

De docenten besluiten om de cursus opnieuw op te zetten. Ze willen aanstaande leraren leren om de grotere kaders te zien waarin het onderwijs plaatsvindt en besluiten daarom hen niet in de rol van leraar te plaatsen, maar in die van beleidsmakers onderwijs. Zo wordt de cursus een spel waarbij studenten de rol van beleidsmaker spelen en de leerstof een middel is om deze beroepsrol te vervullen. Vervolgens zoeken de docenten naar bredere visies die studenten kunnen helpen om de voor- en nadelen van hun eigen visie in perspectief te plaatsen. Twee centrale begrippen uit de literatuur vervullen deze rol: universalisme en relativisme (Procee, 1991). Vanuit het universalisme wordt benadrukt dat voor verschillende culturen dezelfde grondwaarden gelden of zouden moeten gelden. Vanuit het relativisme wordt benadrukt dat verschillende culturen juist verschillende waarden hebben en dat daarvoor ruimte moet zijn.

De contactmomenten bestaan uit twee onderdelen: een literatuurgedeelte en een toepassingsgedeelte. Het literatuurgedeelte gaat over de vraag 'begrijpen wij wat de schrijver bedoelt?' In het tweede gedeelte gaan studenten iets met de kennis doen in de rol van beleidsmakers. Ze bekijken bijvoorbeeld vanuit welke cultureel perspectief schoolboeken geschreven zijn. In een rollenspel waarin de studenten c.q. beleidsmakers van een grote schoolvereniging advies moeten uitbrengen aan schooldirecteuren, presenteren ze hun bevindingen. Tijdens een ander college wordt een situatie nagespeeld waarin schooldirecteuren samenkomen om te bespreken hoe ze onderwijsprestaties van allochtone leerlingen kunnen bevorderen. Studenten kruipen in de huid van de auteurs van wie ze literatuur gelezen hebben en geven vanuit het perspectief van die auteur een visie op de problematiek alsof de auteurs in levende lijve aanwezig zijn.

Daarnaast is er relatief vroeg in de cursus een wat grotere verwerkingsopdracht die formatief en summatief beoordeeld wordt. Studenten krijgen een fictieve casus met een dilemma voorgelegd aan de hand waarvan ze een politicus moeten adviseren. De casus gaat over een groep burgers, onder wie een aantal militante actievoerders, die van mening is dat er in het curriculum op Nederlandse scholen te weinig aandacht is voor de islam. De groep eist erkenning voor de belangrijke rol die de Arabische beschaving heeft gespeeld, ook voor de Europese cultuur als geheel en verwijst daarbij naar gezaghebbende bronnen. In de rol van beleidsmakers moeten de studenten in tweetallen een kort artikel schrijven. Vanuit een helikopterblik moeten ze bouwstenen aandragen waarmee een politicus op deze maatschappelijke situatie kan reageren. Dit advies moet beargumenteerd worden vanuit de literatuur die studenten gelezen hebben. Deze literatuur bevat echter vele perspectieven. In de wereld van relativisme en universalisme moeten studenten een eigen visie vormen en deze onderbouwen.

Deze opdracht verschaft de docenten een schat aan informatie. Kunnen studenten abstract denken? Welke indrukken laat hun advies achter bij de lezer? Maken ze op een zinnige manier gebruik van literatuur? Formuleren ze helder? In deze opdracht gaat het dus niet om de vraag of de student 'het juiste antwoord' geeft maar of de student een zelf gekozen invalshoek goed onderbouwt. De docenten geven eerst formatieve feedback. Ze leggen hun eigen beoordelingen aan elkaar voor en ontwikkelen zo al doende een gezamenlijk kader. Ze praten over vragen als: Wat stellen we ons voor bij een goede onderbouwing? Wat bedoelen we met een helikopterblik? Welke bewoordingen kiezen we om studenten verder te helpen? Tijdens de rest van de cursus kunnen de docenten inspelen op wat hen in de schrijfopdrachten is opgevallen.



In de nieuwe opzet gaat het niet alleen om de leerstof, maar ook om de vraag wie de studenten zelf zijn. Tegelijkertijd leren studenten hun gedachtegang goed te onderbouwen met behulp van diverse begrippen waaronder relativisme en universalisme. De gesprekken tijdens de contactmomenten gaan voortdurend over levensbeschouwelijke vragen maar niet op een manier waarbij alles kan en alles mag. Doordat de studenten als persoon meer uit de verf komen, leren de docenten de studenten beter kennen. Dit bevordert de onderlinge relatie.

De cursus eindigt, als vanouds, met een tentamen bestaande uit essay-vragen. Dit tentamen is echter minder omvangrijk dan in de oude opzet. Studenten zijn bovendien gericht voorbereid doordat ze de stof al op diverse manieren hebben toegepast. Docenten hebben van tevoren op een rijtje gezet welke aspecten ten minste moeten voorkomen in een goed antwoord. Dit vergemakkelijkt het nakijken en bevordert de betrouwbaarheid van de beoordeling.

De waardering voor deze cursus stijgt significant. In een terugblik proberen de docenten 'het succes' te verklaren zodat zij ook voor andere cursussen van deze cursus kunnen leren. Ze komen tot de volgende punten:

1. Het is belangrijk om studenten in een rol te plaatsen waarin ze bepaalde kennis nodig hebben.
2. De literatuur moet gepaard gaan met open toepassingsopdrachten waaruit een eigen interpretatie van de studenten kan volgen.
3. Er dient een theoretisch kader te zijn om de eigen interpretatie te ordenen. Het gaat als het ware om een groot verhaal waarin hun kleine verhaal past.
4. Een gedeelte van de open opdrachten moet formatief geëvalueerd worden.
5. Voor formatieve en summatieve evaluatie van open opdrachten dient een beoordelingskader gemaakt te worden, maar daarin kan niet alles letterlijk vastgelegd worden.
6. Overleg tussen docenten onderling over deze beoordelingskaders is noodzakelijk en kost tijd maar is tegelijk leerzaam en inspirerend. Naarmate de docenten het beoordelingskader meer in de vingers hebben, is overleg minder nodig: docenten voelen elkaar beter aan.

Gespreksvragen:

Op welke manier kunt u de punten toepassen op uw eigen lessen of modules?
Welke kenmerken van de verschillende paradigma's herkent u in de nieuwe manier waarop de cursus is opgezet?

Kenmerken van ‘assessmentcultuur’ tegenover testcultuur

In het bovenstaande voorbeeld werd een cursus, die aanvankelijk alleen geëvalueerd werd via een schriftelijk tentamen met essayvragen, opnieuw opgezet. De cursus werd verrijkt met formatieve evaluatie. Er werden levenschte situaties nagespeeld en er werd gekeken hoe de studenten in die situaties presteerden. Dit wordt aangeduid met ‘authentiek’ assessment. Denk daarbij bijvoorbeeld aan het advies dat de studenten moeten schrijven voor de politicus. Zoiets was in het dagelijks leven ook mogelijk geweest. De kwaliteit van deze producten of presentaties kan niet met gestandaardiseerde instrumenten gemeten worden. Er kan wel degelijk iets gezegd worden over de kwaliteit, vandaar dat we spreken van ‘assessment’.

Zo komt onderwijs tot stand waarin ruimte is voor bredere vorming dan binnen het instrumentele paradigma. De testcultuur maakt plaats voor een assessmentcultuur (zie voor uitgebreidere informatie Dochy & Janssens, 2003).

Testcultuur	Assessmentcultuur
Toetsen staan los van het onderwijs	Integratie van toetsen en onderwijs
Klemtoon op summatieve evaluatie	Evenwicht tussen formatieve en summatieve evaluatie
Evaluatie uitsluitend door de leraar	Evaluatie zowel door de leraar als door de lerende
Het toetsen van vooral kennis	Het toetsen van kennis, toepassing van kennis, vaardigheden en competenties
Gestandaardiseerde toetsen	Gebruik van levensechte situaties in de vorm van problemen, casussen, simulatie etc.
Studenten met elkaar vergelijken	Studenten volgen en hun prestaties vergelijken met een absolute norm en met hun eigen prestaties in het verleden

Gespreksvraag: Welke manier van toetsen past het beste bij uw praktijk?

Slot

In dit hoofdstuk hebben we gezien hoe visies op toetsen samenhangen met visies op onderwijs. De instrumentele benadering, die op zichzelf goede kanten heeft, heeft ook beperkingen die de kwaliteit van onderwijs niet altijd ten

goede komen. Evaluatie kan echter niet alleen ingezet worden als middel om te selecteren maar ook als middel om te leren. De testcultuur maakt dan ruimte voor een assessmentcultuur.

Vorming is in principe een open proces en vraagt daarom ook om open vormen van evalueren. In deze redenering zit echter een spanningsveld: wie open staat voor iets unieks, moet blij zijn met wat er komt. Deze prestaties kunnen niet langs een meetlat gelegd worden. Het is daarmee de vraag of garanties naar buiten verstrekt kunnen worden waaruit blijkt dat aan de gestelde kwaliteitseisen is voldaan. Als de functie van persoonsvorming op de voorgrond komt, lijkt de kwalificerende functie van het onderwijs naar de achtergrond te verdwijnen (Biesta, 2012). Snijdt het vormingsparadigma daarmee in eigen vlees, waardoor het instrumentele paradigma weer opgeroepen wordt? In dit hoofdstuk zijn we ervan uitgegaan dat beide paradigma's elkaar kunnen verrijken. In het volgende hoofdstuk zien we dat er – toen in het onderwijsbeleid aandacht gevraagd werd voor elementen uit het vormingsparadigma – een enorme discussie ontstond die een reactie in de richting van het instrumentele paradigma veroorzaakte.

Hoofdstuk 5

Slingeren tussen instrumenteel en vormend paradigma

Toen de beperkingen van het instrumentele paradigma duidelijk werden, sloeg de slinger door naar de andere kant. ‘Het nieuwe leren’, dat een aantal overeenkomsten met het vormingsparadigma had, pretendeerde een alternatief te bieden. Deze beweging veroorzaakte echter vrij snel een tegenreactie die alles te maken had met zorgen over het niveau. ‘Het nieuwe leren’ leek ondanks goede bedoelingen ten koste te gaan van de breedte en diepgang van kennisbeheersing. Zo kwamen toetsen en de nadruk op kennisverwerving weer terug. Terwijl scholen nog bezig zijn om deze instrumentele golf te verwerken, lijkt er een nieuwe golf aan te komen die benadrukt dat alleen toetsbare kennis toch heel beperkt is. Deze slingerbeweging zegt iets over de aard van het onderwijs en de manier waarop geprobeerd wordt om daar grip op te krijgen. In het zesde hoofdstuk wordt dit verder uitgewerkt. In dit hoofdstuk gaan we dieper in op het onderwijsbeleid en op daar achterliggende opinies. Het hart van het onderwijs ligt echter niet bij onderwijsbeleid, maar in de klas. Om die reden sluiten we dit hoofdstuk af met de beschrijving van een onderwijsexperiment waaruit bleek dat het onderwijs verbeterd kon worden door cijfers af te schaffen en formatieve evaluatie te bevorderen.

Nieuw leren en competentiegericht leren

De bezwaren tegen instrumentele manieren van toetsen en evalueren leidden tot het inzicht dat er breder gekeken moest worden naar de ontwikkeling van studenten. Visies op toetsen hangen echter samen met visies op onderwijs. Geen wonder dus, dat ook visies op onderwijs verschoven. In Nederland hebben we in de jaren negentig van de vorige eeuw en daarna kennisgemaakt met de verschuiving in het onderwijsbeleid onder de noemers ‘het nieuwe leren’ en ‘competentiegericht leren’. In het voorgezet onderwijs speelde vooral de zorg dat het traditionele onderwijs studenten onvoldoende zelfstandig maak-

te, waardoor de overgang naar het hoger onderwijs te groot was. Daar kreeg 'het nieuwe leren' vorm in Het Studiehuis. Het competentiegerichte leren kreeg vooral in het mbo en hbo ingang. Daar probeerde men het onderwijs beter aan te laten sluiten bij het echte leven.

Voordat het competentiegerichte leren werd geïntroduceerd, leek het beroeps- onderwijs sterk op het voortgezet onderwijs. Er was leerstof die didactisch geanalyseerd werd in vakken en toetsen. De schoolse cultuur stond soms ver af van de praktijk waarvoor het beroepsonderwijs opleidde. De student als persoon en diens denk- en leefwereld kwamen in de stage wel naar voren, maar of dat in de rest van het onderwijs ook het geval was, was de vraag. Eén van de idealen van het competentiegerichte onderwijs was om de student actiever bij het onderwijs te betrekken en daarmee de motivatie te bevorderen. Onder competentie werd verstaan *'de bekwaamheid om kennis, vaardigheden en houdingen geïntegreerd en bewust in te zetten in verschillende praktijksituaties volgens een erkende kwaliteitsstandaard'* (De Boer, Van der Burg, Hoek, Kerkhoffs, Kuhlemeier, Raanhuis & Willemse, 2005, p. 19). Niet alleen de stages maar het onderwijs als geheel zou zich op deze competenties moeten richten. Voor allerlei beroepsopleidingen werden competentieprofielen opgesteld. Scholen gingen naarstig op zoek naar nieuwe evaluatievormen die pasten bij ervaringsleren in praktijksituaties binnen en buiten de school. Zo kwamen assessments in zwang. Termen als een persoonlijk ontwikkelingsplan (pop) en reflectie op eigen handelen werden overal gehoord. Diverse (hoge)scholen richtten zich op probleemgestuurd leren, vraaggestuurd leren of zelfregulerend leren. Op veel scholen werd benadrukt dat de studenten eigenaar moesten zijn van hun eigen leerproces. Daarbij was het geenszins de bedoeling om studenten aan hun lot over te laten. De school moest 'een rijke leeromgeving' bieden die studenten zou ondersteunen bij het vaststellen en doorlopen van een persoonlijke leerroute. Die route zou aansluiten bij de belangstelling en de talenten van de studenten en hen een optimale voorbereiding geven op een bepaalde plek op de arbeidsmarkt of doorstroming bevorderen naar hoger onderwijs. Docenten werden begeleiders of coaches. Vakken werden geïntegreerd. Er kwamen kwalificatiedossiers en portfolio's. Studenten moesten zelf bewijzen verzamelen dat ze over de benodigde competenties beschikten (De Mynck & Roeleveld, 2005).

Competentiegerichte leren was één van de vormen die onder de term 'het nieuwe leren' werd gevat. Van Aalst & Kok (2004) en De Boer e.a. (2005) noemden daarvoor de volgende kenmerken in vergelijking met wat 'het oude leren' werd genoemd.

Van het 'Oude leren'	naar het 'Nieuwe leren'
Aanbodsturing: consumeren van door anderen gevonden kennis in een door anderen bedachte volgorde	Vraagsturing: actief construeren van kennis, voortbouwend op aanwezige kennis en denkschema's
Onderwijzen in een klassieke benadering; alle studenten op hetzelfde moment dezelfde leerstof	Leren in gevarieerde arrangementen en met een variëteit van sturingsmogelijkheden
Van buiten leren in een één-op-één-relatie met een bron van objectieve kennis; overdrachtsmodel	Van binnen leren in een voortdurende interactie binnen een gegeven sociale en culturele context; subjectieve kennis; metacognitieve vaardigheden (leren leren) om het eigen leerproces te reguleren
Standaardisering	Maatwerk
De docent is de expert, verteller en beoordelaar; de student is een passieve luisteraar die kennis reproduceert; weinig mogelijkheden voor interactie	De docent is ook coach en begeleider; de lerende is actief en construeert zelf kennis; kleinschalige leeromgeving, met veel interactie
Lineaire leerstofplanning; opdeling in deelvaardigheden en hapklare brokken	Circulair leerstofaanbod; duurzaam, flexibel, functioneel geïntegreerd en betekenisvol
Leren voor de toets; de toets dient om de docent te informeren of de student iets heeft geleerd om daarmee vervolgens tot ranking en selectie te kunnen overgaan	Toetsen wat er geleerd is; de toets dient ook om student en docent te informeren over het leerproces om op basis daarvan vervolgens het arrangement eventueel te verrijken
Schoolse abstracties; leren is losgemaakt van de context; leren van feiten en geïsoleerde vaardigheden; weinig aansluitend op reeds aanwezige kennis en vaardigheden	Authentieke contexten; situatiegebonden leren; verwerven van competenties voortbouwend op reeds aanwezige en elders verworven competenties
Ordering in disciplines (vakken) en daarbinnen de structuur van het vak volgend	Ordering naar problemen en belangstellingsgebieden; het begrijpen van de structuur van het vak is zelf een deel van de inhoud
Beperkt aantal bronnen (de docent en het boek)	Een rijke variatie aan leerbronnen en hulpmiddelen (onder andere ict)
Leren vindt alleen plaats op school; geen verbinding tussen het buitenschoolse en het binnenschoolse leren	Leren vindt de hele dag door plaats; het leren op school en het leren daarbuiten zijn verbonden, vervlochten en gestapeld

Afbeelding 5.1. Kenmerken van 'het nieuwe leren', naar Van Aalst en Kok (2004) en De Boer e.a. (2005)

Hoewel er enkele duidelijke verschillen zijn, kost het weinig moeite om elementen van het vormingsparadigma te herkennen in deze kenmerken van 'het nieuwe leren'. Studenten kregen zelf een veel actievere rol. Door reflectie werden ze gestimuleerd om hun eigen werk mede te beoordelen. Objectieve kennis kreeg een minder centrale plaats. Het onderwijs was meer gericht op leren handelen in de beroepspraktijk. Daarbij was ruimte voor Erkenning van Verworven Competenties (EVC's), dus ook het buitenschoolse leren kreeg een plaats. Studenten hoefden dus niet meer allemaal dezelfde leerweg te doorlopen. Instrumentele vormen van toetsing met hun strenge selectieve werking maakten plaats voor evaluatie in levensechte contexten.

Gespreksvraag: Op welke manier hebt u de hier beschreven ontwikkelingen meegemaakt? Instemmend of kritisch? Hoe staat u op dit moment tegenover de twee paradigma's?

Bezwaren tegen het vormingsparadigma

Hoewel het competentiegericht onderwijs door een deel van de docenten met enthousiasme werd onthaald, met name omdat zij er een mogelijkheid in zagen om hun onderwijs betekenisvoller te maken, heeft de invoering van competentiegericht onderwijs tot felle discussies geleid. 'Het nieuwe leren' benadrukte dat studenten informatie op hun eigen manier interpreteren waarbij hun persoonlijke vaardigheden, voorkeuren en hun eigen culturele achtergrond een belangrijke rol spelen (zie het kader hieronder). Sommige aanhangers van 'het nieuwe leren' gingen zo ver dat ze vonden dat ware kennis niet bestaat. Wat als 'waar' wordt beschouwd, hangt samen met de cultuur waarin je opgroeit. Europeanen zeggen dat Columbus Amerika heeft ontdekt; Indianen zullen zeggen dat Columbus Amerika heeft bezet. Waar in het traditionele onderwijs een strikt onderscheid werd gemaakt tussen wat 'goed' en 'fout' is, werd dat verschil binnen 'het nieuwe leren' (en het daarmee verwante sociaal-constructivisme) gerelativeerd.

Een kleuterjuffrouw neemt een toets af om de ontwikkeling van haar leerlingen in kaart te brengen. Bij een plaatje van een step, een bus, een brommer en een fiets luidt de vraag: wat gaat het snelst? De juf werkt midden in de stad waar het verkeer voortdurend verstopt is. Wat antwoorden kinderen? De meeste kinderen vinden dat niet de bus maar de brommer het snelst is. Volgens de opstellers

van de toets is dit antwoord fout. De juf begrijpt echter dat leerlingen kennis construeren vanuit hun sociale omgeving. In de stad is de brommer sneller dan de bus, wát de opstellers van de toets ook vinden!

De relativering van wat ‘ware kennis’ is, werd verbonden met de opvatting dat kennis niet zomaar overgedragen kan worden alsof studenten holle vaten zijn die gevuld moeten worden. In plaats daarvan werd benadrukt dat studenten altijd al met bepaalde opvattingen de school binnenkomen, die een rol speelt bij wat ze daadwerkelijk leren. Kennis wordt door studenten zelf geconstrueerd, zo werd benadrukt. Dit werd in veel gevallen te letterlijk opgevat. Studenten moesten zelf op zoek gaan naar informatie, maar al gauw rees de vraag of het aldus van het internet bij elkaar geraapte knip-en-plakwerk werkelijk tot beter onderwijs zou leiden.

Vragen of de student wel in staat is om zelf kennis te construeren, werden prangend omdat ook benadrukt werd dat de rol van de docent moest veranderen naar een begeleider van het proces waarin studenten hun leerproces zelf reguleren. Inspirerende docenten, die studenten wisten te motiveren voor hun vak, voelden zich aan de zijlijn gezet. Intussen beloofde de theorie van ‘het nieuwe leren’ dat studenten gemotiveerder zouden worden maar de praktijk van het onderwijs bleek weerbarstig.

Docent rekenen pabo: "Ik heb meegewerkt aan het Surfproject waarbij we een leerlijn van mbo naar hbo ontwikkeld hebben met bijbehorende digitale toetsen. Met deze toetsen kunnen studenten zelfstandig oefenen. Ze geven studenten informatie over hun eigen leerproces. Ik vind het belangrijk dat studenten zicht krijgen op hun eigen ontwikkeling zodat ze weten aan welke onderdelen ze nog moeten werken. Als studenten er lerend in staan, werkt dit erg goed. Dan kunnen ze zelfs hun lerende houding verbeteren met digitale toetsen. Maar als ze er niet lerend in staan, werkt het niet! Je kunt deze manier van zelfstudie eigenlijk niet opleggen. Ze moeten er zelf voor kiezen. En dat is voor veel jongeren best veel gevraagd. Digitale toetsen worden in het rooster ingepland en dus aan studenten opgelegd. Waarschijnlijk zal een deel van de studenten deze uitdaging oppakken en werken aan het eigen leerproces. Maar voor een deel van de studenten zal dit niet werken. De resultaten zijn dan veel minder."

Gespreksvraag: Er zijn docenten die met iedere slingerbeweging hun voordeel doen. Ze grijpen de beweging als het ware aan om hun eigen repertoire uit te breiden. Herkent u daar iets van?

Met haar rede ‘Leren in het studiehuis’ gooide hoogleraar Greetje van der Werf (2005) olie op het vuur. Ze stelde dat de effectiviteit van ‘het nieuwe leren’ niet met empirisch onderzoek kon worden onderbouwd en in bepaalde gevallen zelfs schadelijk is. Van der Werf pleitte voor een interpretatie van het directe-instructiemodel dat verwant is aan het model Didactische Analyse, zoals uitgelegd in hoofdstuk 2. Ze verwees naar onderzoek waaruit bleek dat dit effectiever was dan constructivistisch onderwijs. Hoogleraar Instructietechnologie Ton de Jong (2006), die met haar de degen kruiste, kon echter allerlei andere publicaties noemen waaruit bleek dat leren in contexten, onderzoekend leren en samenwerkend leren er in vergelijkend onderzoek wél beter uitgekomen was. Beide hoogleraren beriepen zich op wetenschappelijk onderzoek, maar de resultaten daarvan waren inconsistent. Intussen had Van der Werf met haar oratie stem gegeven aan een gevoel dat leefde onder een deel van de docenten dat het oude systeem veel beter was. De Jong kon op grond van andere onderzoeken laten zien dat er ook reden was om ‘het nieuwe leren’ te ondersteunen.

Protesten tegen ‘het nieuwe leren’ werden gecombineerd met zorgen over niveauverlaging. De resultaten van het ‘levensechte’ onderwijs werden niet gemeten met gestandaardiseerde toetsen. Er bestond ook geen traditie om te beoordelen of de werkstukken en competentiedossiers aan de benodigde kwaliteitseisen voldeden. Het niveau van het onderwijs kon dus niet gegarandeerd worden.

De bezwaren tegen het vormingsparadigma werden duidelijk: Hoe kan gegarandeerd worden dat de brede vorming werkelijk ergens over gaat als er geen meetinstrumenten zijn die dat vaststellen? Met welk recht kwalificeert het onderwijs studenten voor een diploma als externe, objectiveerbare normen ontbreken? De diploma’s die uitgegeven worden, moeten wel wat voorstellen! Ook hier is de theorie van Biesta (2012) over de functies van onderwijs van toepassing: als de persoonvormende functie veel nadruk krijgt, wordt de kwalificerende functie naar de achtergrond gedrukt.

Een nieuwe instrumentele golf

In 2007 besloot de Tweede Kamer om een parlementaire onderzoekscommissie in te stellen onder leiding van Jeroen Dijsselbloem, die zich voornamelijk richtte op het voortgezet onderwijs. Het rapport van deze commissie, *Tijd voor Onderwijs*, bepleitte ogenschijnlijk dat de overheid zich minder inhoudelijk met het onderwijs zou moeten bemoeien. De overheid zou zich voortaan alleen maar bezig moeten houden met het *wat*, dus de inhoud van de leerstof, terwijl het *hoe*, de didactiek, overgelaten moest worden aan het onderwijsveld.

Intussen deed de commissie voorstellen die de ontwikkelingen binnen het onderwijs diepgaand beïnvloedden. Ze betroffen lang niet alleen het 'wat'. De commissie concludeerde dat onderwijsvernieuwing voortaan beter wetenschappelijk onderbouwd diende te zijn. Voor het basisonderwijs stelde de commissie voor om in groep drie een begintoets in te voeren. Er moest een Nationale Monitor van het onderwijs komen. De commissie dacht weer in termen van afzonderlijke vakken, waarvoor een canon moest worden opgesteld in de vorm van een kennisbasis met leerstandaarden. Een vak als Nederlands, zo schreef de commissie, zou niet meer met andere vakken geïntegreerd moeten worden. Ook de leerdoelen moesten worden geconcretiseerd. Zo zou voor iedereen helder worden wat leerlingen aan het eind van de schoolloopbaan moesten kennen en kunnen. Op deze manier dacht de commissie ook de positie van de leraar te versterken. Voor wat betreft de lerarenopleidingen vond de commissie dat de overheid heldere onderwijsdoelen moest formuleren en die via centrale examinering moest invoeren.

Uit deze aanbevelingen blijkt dat de commissie een eigen visie op onderwijs had, namelijk één waarin doelstellingen en toetsing centraal stonden (Dijsselbloem, 2008). Sindsdien hebben we gezien dat het instrumentele paradigma weer terug is van weggeweest. Deze beweging treft het hele onderwijsveld, zowel het primaire, het secundaire onderwijs als het beroepsonderwijs. Voortdurend wordt de discussie gevoed door de angst voor niveaudaling (zie het kader over rekenonderwijs in hoofdstuk 6 op p. 85 en 87/88).

De examens in het voortgezet onderwijs die vanouds uit twee delen bestonden, namelijk een centraal schriftelijk eindexamen en een schoolexamen, kregen een andere functie. De oorspronkelijke bedoeling van het schoolexamen was om resultaten te evalueren die minder makkelijk meetbaar waren in landelijke eindexamens. Bij het vak tekenen bijvoorbeeld, leren leerlingen in eerste instantie tekenen. Daarnaast krijgen ze kunstgeschiedenis.

Het laatste wordt landelijk getoetst maar het eerste niet. Heel bewust was het examen zo opgesteld dat recht werd gedaan aan verschillende aspecten binnen het vak. Het sprak vanzelf dat leerlingen niet even goed hoefden te zijn in verschillende aspecten van het vak. Wie goed is in tekenen, hoeft niet per se goed te zijn in kunstgeschiedenis; daarom waren er juist twee examens. Een verschil tussen schoolexamen en centraal examen wees erop dat beide examens andere dingen evalueerden en allebei dus bestaansrecht hadden. Totdat de nieuwe instrumentele golf toesloeg. De gedachte dat er twee examens bestonden om verschillende zaken te meten, maakte plaats voor de strenge technocratische eis van betrouwbaarheid. Als de examens zuiver meten, moet het ene examen tot dezelfde resultaten leiden als het andere – met als gevolg dat hoe mooier de tekeningen van leerlingen worden, des te meer tekendocenten zich zorgen beginnen te maken over de kloof die ontstaat met het centraal examen en de negatieve beoordeling van de school die daar het gevolg van is.

Gespreksvraag: Een meting van een leerling-vaardigheid wordt betrouwbaar genoemd als twee of meer verschillende metingen tot ongeveer hetzelfde resultaat leiden. Het voorbeeld van tekenen laat zien dat binnen hetzelfde vakgebied sprake kan zijn van verschillende vaardigheden waar dezelfde leerling verschillend op kan scoren. Is dit in uw vak ook het geval?

De instrumentele golf deed ook het basisonderwijs aan, onder andere omdat het daar verplicht werd om met gestandaardiseerde toetsen de leerling-resultaten te volgen. Rekenen en taal werden vakken voor de streep, de andere vakken kwamen achter de streep terecht. We verkeren nu in een situatie waarin de overheid de toetsresultaten nauwlettend in de gaten houdt, precies zoals Dijsselbloem had voorgesteld, terwijl zijn onderzoekscommissie zich niet eens met het basisonderwijs had beziggehouden. De commissie Dijsselbloem had de overheid verweten dat zij teveel in de vrijheid van scholen was getreden. Met haar eigen instrumentele visie op onderwijs trad Dijsselbloem echter evenzeer in de vrijheid van scholen.

Ook het beroepsonderwijs, met name het hoger beroepsonderwijs, kreeg een instrumentele golf te verwerken. Als eerste hadden lerarenopleidingen hiermee te maken. Er kwamen verontrustende verhalen over leraren die niet konden rekenen en spellen. Zo werd een verplichte reken- en taaltoets ingevoerd. Niet lang daarna verschenen zogenoemde kennisbasissen voor verschillende vakken. Intussen kwam het hele hbo onder kritiek te staan omdat opleidingen

niet aan het vereiste niveau voldeden en diploma's waren afgegeven voor werk dat beneden de maat was. De toetsing deugde dus niet. Dergelijke berichten brachten de overheid steeds meer op de gedachte dat er centrale examinering zou moeten plaatsvinden, ook in het hbo.

Vreemde ogen dwingen in het hbo

De diploma-drama's in het hbo vroegen om maatregelen. Landelijke examens geven een objectief oordeel dat voor iedereen hetzelfde is, zo was de aanname. Tegelijkertijd was ook duidelijk dat alleen een beperkt deel van de competenties die in het beroepsonderwijs van waarde zijn, met landelijke toetsen meetbaar zijn. De Vereniging Hogescholen liet een studie verrichten over hoe de toetsing in het hbo, waar beroepscompetenties centraal staan, zou kunnen verbeteren. Dat heeft geleid tot het advies 'Vreemde ogen dwingen' (Commissie externe validering examenkwaliteit hoger beroepsonderwijs, 2012). De voordelen van zo'n 'klassiek eindexamen', zoals in het voortgezet onderwijs, waren de commissie wel duidelijk, maar de nadelen ook. Zo'n eindexamen zou ingrijpende gevolgen hebben voor het onderwijs en de invoering zou uiteindelijk niet leiden tot die kwaliteiten die van een hbo-student verwacht worden. Hbo-docenten zouden hun professionele ruimte verliezen en als het ware uitvoerders worden van maatregelen van bovenaf. De commissie wilde de hbo-docenten actief bij toetsing betrekken. Zo is het kernwoord in de aanbevelingen 'externe validering' geworden. Daarmee geeft de commissie aan dat de kwaliteit van de studenten niet alleen door de hogeschool zelf erkend moet worden maar ook door anderen buiten de hogeschool. Toetsing hoeft dus niet landelijk, maar hogescholen kunnen elkaar helpen. Docenten moeten zelf beoordelingscriteria voor werkstukken opstellen. Ze moeten geschoold worden om betere toetsen te kunnen maken. Bij het beoordelen van het werk van studenten zouden ze meer kennis moeten delen.

In dit verband wordt ook wel gesproken van het vier-ogen principe, wat duidt op de verplichting dat bij het maken van een toets of bij een beoordeling van een student steeds twee docenten hun mening geven. Ten slotte moet iedere opleiding een toetsbeleid voeren waarin externe validering een plaats krijgt.

Gespreksvraag: Speelt externe validering een rol in uw praktijk? Maakt u samen met collega's toetsen? Beoordeelt u wel eens werk van studenten samen met anderen, binnen of buiten de onderwijsinstelling?

Een smalle kijk op onderwijskwaliteit

Het onderwijs is nog volop bezig met het verwerken van de instrumentele golf. De bezwaren tegen de instrumentele benadering zijn er echter niet anders op geworden, daarvoor kunnen we hoofdstuk 3 er nog eens op naslaan. Het is dus niet verwonderlijk dat de eerste voorbode van een nieuwe slingerbeweging aangekomen is. Onlangs verscheen het rapport 'Een smalle kijk op onderwijskwaliteit' van de Onderwijsraad (2013). De afgelopen periode was de aandacht van overheid en inspectie te eenzijdig gericht op meetbare doelen. Dat geldt in het bijzonder voor taal- en rekenprestaties, zo stelt het rapport. De Onderwijsraad maakt zich onder andere zorgen over de eigenwaarde van leerlingen die niet goed presteren op de basisvaardigheden en vindt daarom dat het onderwijs de brede kwaliteit van leerlingen inzichtelijk moet maken. Daarbij vraagt de raad aandacht voor burgerschapsvorming en vakoverstijgende vaardigheden. De samenleving heeft ook behoefte aan creativiteit, probleemoplossend vermogen, samenwerking, culturele en morele sensitiviteit, zorgzaamheid en vakmanschap. Er moet meer persoonlijke differentiatie mogelijk zijn in onderwijsprogramma's en leerlingen zouden elkaar vanuit verschillende achtergronden in het onderwijs moeten treffen. Voor burgerschapsvorming wil de onderwijsraad getalsmatige indicatoren instellen. Het vormingsparadigma wordt dus op een instrumentele manier op de agenda gezet. De vraag is of we daarmee weer terugkomen bij het vormingsparadigma.

Een reactie uit het mbo: "Het instrumentele paradigma heeft erin geresulteerd dat we nu schilderwerkstukken van studenten beoordelen met een aanvinklijst. Vanwege de objectiviteit! Daar hebben de muzisch-creatieve vakken de laatste jaren behoorlijk onder geleden en hebben sterk aan kracht ingeboet..."

Betere resultaten door cijfers af te schaffen?

Onze aandacht is in dit hoofdstuk verschoven van het handelen van docenten op microniveau naar het onderwijsbeleid op macroniveau. Het hart van het onderwijs bevindt zich echter altijd daar waar docenten en studenten elkaar ontmoeten. In dit hoofdstuk hebben we de commissie Dijsselbloem genoemd die het mislukken van 'het nieuwe leren' heeft onderzocht. We hebben echter nog weinig aandacht besteed aan wat dit paradigma kan bijdragen aan onderwijs. Daarom eindigen we dit hoofdstuk toch weer op microniveau met een experiment waarin het onderwijs verbeterd werd door het afschaffen van cijfers. Het uitproberen van een andere manier van onderwijzen was niet gemakkelijk maar een waar gevecht. Het experiment vond plaats op een school in Engeland



onder leiding van de onderzoeker Dylan Wiliam. In twee afleveringen van een uur heeft de BBC dit experiment op zeer boeiende wijze vastgelegd. Wie deze twee uur bekijkt, maakt het experiment als het ware zelf mee.

- <http://www.youtube.com/watch?v=J25d9aC1GZA>;
- <http://www.youtube.com/watch?v=1iD6Zadhg4M>

Dylan Wiliam kan geplaatst worden binnen de beweging die het grote belang van formatieve evaluatie benadrukt, ‘evalueren om te leren’ genaamd. Feedback geven blijkt een van de meest bepalende factoren die de onderwijsprestaties ten goede komen (Hattie, 2009; zie ook Casteleijns, Segers & Struyven, 2011), maar de ene soort feedback levert meer op dan de andere. De beste resultaten konden worden vastgesteld wanneer docenten ingingen op de manier waarop studenten de opdracht inhoudelijk hadden uitgevoerd (Hattie & Timperley, 2007). Het onderstaande verhaal toont hoe docenten dit gegeven in hun praktijk hebben toegepast.

Gespreksvraag: Hoe ziet de feedback die u uw leerlingen geeft eruit?

Het begint allemaal met een klas met tieners waarin slechts een handjevol leerlingen meedoet. Verreweg de meerderheid zit onderuit gezakt en interesseert zich niet voor de les. Er wordt veel gekletst. Zelfs goedwillende leerlingen verklaren weinig tot niets te leren. Deze klas is op de school overigens geen uitzondering.

Bij wijze van experiment stelt Wiliam de docenten voor om alleen bij deze klas een aantal maatregelen uit te proberen. Voortaan worden er geen cijfers meer gegeven. In plaats daarvan komt inhoudelijk commentaar, dat de leerlingen aan het denken moet zetten. Zo wil Wiliam dat leerlingen hun eigen leerproces gaan sturen. Cijfers zitten dat eigenaarschap in de weg, vindt hij. Een voldoende betekent dat een leerling 'genoeg' gedaan heeft, dat stimuleert niet om te blijven leren. Wiliam wil ook niet dat er vingers opgestoken worden. Hij geeft leraren een bekertje met ijsstokjes waarop de namen van de leerlingen staan. Bij iedere vraag kan de leraar een stokje trekken en er zo voor zorgen dat iedere leerling een gelijke kans heeft op een beurt. Daarnaast verzint hij maatregelen waardoor de leraar inzicht krijgt in hoe de leerlingen denken en wat ze nodig hebben. Het zijn maatregelen die formatieve evaluatie mogelijk maken. Op de tafels van studenten komen plastic bekertjes in verschillende kleuren, waarbij rood aangeeft dat er hulp nodig is. Via kleine whiteboards die iedereen krijgt uitgereikt, kunnen leerlingen hun antwoorden in één handbeweging publiek tonen. Alles is bedoeld om actief leren te bevorderen en leraren te informeren over hoe het leerproces verloopt zodat ze pas overstappen op een ander onderdeel als ze zeker weten dat het voorafgaande onderdeel is begrepen. Ten slotte krijgt de klas iedere dag gymnastiek.

Rondom de beurten via de ijsstokjes voltrekt zich een klein drama. Uit het bekertje van de leraren blijkt een stokje verdwenen te zijn. Wat is hier aan de hand? Een goede leerling heeft haar stokje stiekem mee naar huis genomen. Het nieuwe systeem bevalt haar helemaal niet. Ze vond het heerlijk om met goede antwoorden te scoren en stak dus – heel slim – alleen haar vinger op als ze wist dat haar antwoord juist was. Nu moet ze haar mond houden als ze het antwoord weet en krijgt zomaar een beurt, ook als ze het antwoord niet weet. Haar glorieuze reputatie als topstudent gaat naar de knoppen! Het groepje leerlingen dat voorheen goed meedeed, dreigt volledig af te haken. Wiliam probeert met hen te praten. Het gaat er toch niet om hoe goed je bent, maar om wat je leert? Dit is toch een school? Als je alleen maar vragen beantwoordt waarvan je het antwoord al weet, dan leer je toch niets? Maar de dames topstudenten hebben met hun schoolcarrière kennelijk een andere bedoeling.

Intussen staan boven het werk van de leerlingen geen cijfers meer. Leraren geven inhoudelijk commentaar. Dat doen ze soms via opdrachtjes ter verbetering, soms via complimenten of door naar de mening van leerlingen te vragen. Ook dat roept weerstand op. De behoefte om zichzelf met anderen te vergelijken is onuitroeibaar. Leerlingen gaan complimenten tellen. Ze proberen toch uit te vinden wie er excellent is, wie gewoon goed is of matig.

De geschiedenislerares merkt wel dat de klas anders reageert dan vroeger. Toen ze nog cijfers gaf, werd haar commentaar nauwelijks gelezen. Zodra het cijfer bekend was, verdween al haar commentaar linea recta de prullenbak in. Nu besteden de leerlingen er tenminste aandacht aan. Maar toch heeft ze vragen bij het nieuwe systeem. Cijfers geven toch informatie of leerlingen de doelstellingen gehaald hebben? Dat helpt toch bij het sturen van het leerproces? Is daar iets op tegen?

Na een week of drie lijkt het erop dat de maatregel om cijfers af te schaffen zijn langste tijd gehad heeft. De afwezigheid van cijfers leidt keer op keer tot teleurstelling onder de leerlingen, alsof ze hun werk voor niets doen. Dat het inhoudelijke commentaar belangrijker is dan de cijfers, overtuigt hen niet. Goede leerlingen zitten op de achterkant van hun boek te tekenen of spelen demonstratief met hun haar. Als ze niet op de éne manier in het middelpunt van de belangstelling kunnen staan, dan maar op de andere. Leraren worden het gezeur ook zat. Ze begrijpen maar al te goed dat leerlingen cijfers belangrijk vinden. Vroeger waren ze ook zo. Zélf zijn ze ook onderdeel van de lat-cultuur waarin datgene wat niet gemeten wordt, geen waarde heeft.

Er wordt maar weinig vooruitgang geboekt. Toch wordt fouten maken geleidelijk aan minder erg gevonden. Reputatie lijkt niet meer het enige te zijn dat telt. De werkhouding van de leerlingen is echter nog niet om over naar huis te schrijven. Wiliam vraagt leraren een laatste truc uit te halen. Iedere dag kiezen de leraren een leerling uit die door goed gedrag voor de hele klas een punt kan verdienen. Per dag zijn er vijf lessen. Welke leerling die dag gekozen is, blijft voor de leerlingen geheim. Goed gedrag in minstens drie van de vijf lessen van de dag levert een punt op voor de hele klas. Het behalen van een bepaald aantal punten binnen vier weken wordt beloond met een dagje pretpark. Zo worden groepsdruk en beloning ingezet om de werksfeer te verbeteren.

Na de eerste dag heeft de klas geen punt verdiend. Dat is jammer, maar het laat wel zien dat het de leraren menens is dat de sfeer fundamenteel moet ver-

anderen. Het beloningssysteem is wennen. Leerlingen maken nogal eens lawaai zonder dat ze zich daar al te zeer van bewust zijn. Dat de leraren hen regelmatig aan 'de geheime leerling' herinneren, blijkt geen overbodige luxe. Ook deze aanpak is geen wondermiddel. Toch wordt gaandeweg duidelijk dat de verantwoordelijkheid voor de sfeer in de groep niet langer alleen maar bij de leraar ligt. Een leerling roept: "Wees stil, want misschien ben jij vandaag de geheime leerling!" Waarop iemand anders zegt: "Hou zelf je mond, jij kunt het ook zijn!"

Een week voor de zomervakantie moet de klas nog twee punten verdienen. De eerste dag van die week wordt geen punt verkregen. De klas wordt streng toegesproken. Op het laatste moment wordt het benodigde aantal punten nog n t binnengehaald. De veranderde sfeer in de klas blijkt tijdens het bezoek aan het pretpark. Er is meer onderling contact tussen de verschillende groepjes; de klas vormt meer een geheel. Na afloop van de twaalf weken maakt Wiliam de balans op door de leerresultaten van deze klas te vergelijken met die van klassen die niet aan het experiment meegedaan hebben. Bij natuurwetenschappen verschillen ze niet; bij wiskunde en Engels zijn ze significant hoger.

Het experiment maakt pijnlijk duidelijk hoe een instrumentele benadering het onderwijs kan doen onttaarden in een reputatiespel en hoeveel aanleidingen er zijn om de focus te verleggen. Maar eenvoudig is dat niet. Zowel de leerlingen als de leraren hebben de neiging om zich vast te klampen aan wat vertrouwd is. Wiliam herhaalt zijn boodschap in vele toonaarden: evaluatie van het leerproces mag nooit een doel in zichzelf zijn. Evaluatie moet   n centraal doel dienen, namelijk informatie verzamelen over het leerproces om vervolgens beter te kunnen begrijpen welke benadering leerlingen nodig hebben.

De BBC-afleveringen laten zien dat onderwijsvernieuwing met spanningen gepaard gaat. Verandering van spijs doet lang niet altijd eten. Gewoontes, ook gewoontes die het leerproces niet bevorderen, gaan leerlingen en leraren in het bloed zitten. Ze zijn niet in een handomdraai af te schaffen. Het verborgen curriculum van het instrumentele paradigma heeft zijn werk gedaan. Toch had de poging om het anders aan te pakken resultaat. De winst is uiteraard niet alleen toe te schrijven aan het feit dat leerlingen kwalitatieve feedback kregen in plaats van een cijfer. Juist de brede benadering van pedagogiek en didactiek hebben tot een klimaat geleid waarin de leerresultaten en de sfeer verbeterden. Opmerkelijk in deze casus is dat het verbeterde klimaat niet spontaan ontstond. Wiliam gebruikte aan het einde zelfs een klassieke manier van beloning, wat zelfs opgevat zou kunnen worden als een pressiemiddel. Dat neemt niet

weg dat deze casus de kracht van het vormingsparadigma laat zien, zelfs in een situatie waarin het instrumentele paradigma in ieders bloed zit.

Gespreksvraag: Hebt u wel eens meegemaakt dat u iets in het onderwijs wilde veranderen, maar bij leerlingen op oude gewoonten stuitte die het gevolg waren van het verborgen curriculum van de oude benadering?

Slot

In dit hoofdstuk hebben we laten zien dat het instrumentele paradigma plaats maakte voor het vormingsparadigma, wat vervolgens weer tot een instrumentele reactie leidde die het onderwijsveld momenteel nog aan het verwerken is. Tegelijkertijd is de voorbode van een nieuwe slingerbeweging zichtbaar. In de volgende hoofdstukken proberen we te begrijpen waar de voortdurende slingerbeweging uit voortkomt en proberen we een manier te vinden om deze te overstijgen.

Hoofdstuk 6

Normatieve professionalisering

In dit hoofdstuk proberen we bij het hart van goed onderwijs te komen. De commissie Dijsselbloem (2008) concludeerde dat onderwijsvernieuwing nog meer dan voorheen gebaseerd moet zijn op resultaten van wetenschappelijk onderzoek. Aanpakken voor onderwijsvernieuwing moesten *evidence-based* zijn. Aan de hand van de resultaten van ruim dertig jaar effectiviteitsonderzoek laten we zien dat we als docenten niet terechtkomen bij één of andere aanpak, maar bij onszelf. We moeten zélf bepalen wat we onder goed onderwijs verstaan en dit zélf vormgeven in telkens weer unieke situaties. We zullen in dit hoofdstuk zien waarom goed onderwijs niet geheel in systemen te vangen is. Onderwijs, inclusief evaluatie en toetsing, is een normatieve praktijk die vraagt om innerlijke ontwikkeling van mensen.

Wat is effectief onderwijs? Een normatieve vraag!

Welke *evidence* is er voor de effectiviteit van scholen die zich op het vormingsparadigma oriënteren versus scholen die zich op het instrumentele paradigma oriënteren? Kan wetenschappelijk onderzoek ons verder helpen?

Vrije Scholen geven onderwijs vanuit de antroposofische uitgangspunten van Rudolf Steiner. Ze streven een brede ontwikkeling van hoofd, hart en handen na en leggen minder de nadruk op cognitieve prestaties. Reguliere scholen vinden cognitieve prestaties doorgaans wel belangrijk. Toetsing op Vrije Scholen is minder selectief dan op reguliere scholen. Wie doet het beter? De onderzoeker Hilde Steenbergen (2009) probeerde een tipje van de sluier op te lichten door de toegevoegde waarde van dertien Vrije scholen voor voortgezet onderwijs te vergelijken met die van reguliere scholen. Ze keek daarbij naar de ontwikkeling van leerlingen op cognitief en niet-cognitief niveau en corrigeerde voor hun beginniveau. Ze lette ook op wat er

gebeurde met de verschillen tussen leerlingen binnen de groepen gedurende de schoolperiode.

Uit de analyse blijkt dat de cognitieve prestaties negatief uitpakken voor de Vrije Scholen. Leerlingen op reguliere scholen boeken meer leerwinst. Intelligentieverschillen tussen leerlingen blijken gedurende de schoolloopbaan op Vrije Scholen groter te worden dan op reguliere scholen. Het academisch zelfbeeld van de leerlingen (het beeld dat ze hebben van hun eigen kennen en kunnen op school) is echter hoger op Vrije scholen, waarbij de intelligentere leerlingen niet noodzakelijkerwijs een hoger zelfbeeld hebben dan minder intelligente leerlingen. Op Vrije scholen blijkt nauwelijks samenhang tussen intelligentie en academisch zelfbeeld te bestaan terwijl het op reguliere scholen duidelijk is dat de intelligentste leerlingen ook het hoogst scoren op academisch zelfbeeld. De relatie van leerlingen met docenten blijkt op Vrije scholen beter te zijn. De prestatiemotivatie neemt op reguliere scholen gedurende de schoolloopbaan behoorlijk af, terwijl dit op Vrije Scholen minder het geval is. Intelligentere leerlingen op Vrije scholen zijn meer gemotiveerd om te presteren dan intelligente leerlingen op reguliere scholen. Bovendien maken leerlingen op Vrije Scholen vaker gebruik van leerstrategieën die gunstig zijn voor toekomstig leren.

Dit onderzoek levert dus een genuanceerd beeld op. Beide benaderingen blijken hun voordelen te hebben. Die voordelen liggen op verschillende vlakken: het reguliere onderwijs scoort duidelijk beter op cognitieve leerprestaties, terwijl Vrije Scholen op niet-cognitief gebied voordelen hebben. Een vraag die daarbij ook gesteld kan worden is in hoeverre de gebruikte instrumenten in staat zijn geweest de cognitieve en niet-cognitieve eigenschappen van leerlingen in al zijn volledigheid in kaart te brengen. Die vraag kan en moet bij ieder instrument gesteld worden. Het antwoord is ook duidelijk: instrumenten zijn altijd tot op zekere hoogte oppervlakkig. Toch biedt dit onderzoek wel degelijk verheldering. Op Vrije Scholen hebben leerlingen meer plezier in leren. Dat is van groot belang in een maatschappij die levenslang leren centraal stelt. Op reguliere scholen leren ze meer.

Het was fijn geweest als onderzoek voor ons had uitgewezen welke benadering het beste was. Maar dit onderzoek speelt de vraag terug naar onszelf: wat vinden we belangrijk? Wie hogere cognitieve prestaties belangrijk vindt, moet niet bij de Vrije School zijn. Wie andere factoren centraal stelt, juist weer wel. Dit onderzoek laat goed zien dat de vraag ‘wat is goed onderwijs?’ een normatieve vraag is. Onderzoek kan de discussie daarover verdiepen (Van

IJzendoorn, 1997), dat bewijst dit onderzoek. Maar juist omdat de voor- en nadelen van de benaderingen verschillende kanten van dezelfde medailles zijn, komen we uiteindelijk toch weer uit bij de vraag waaraan we zelf de voorkeur geven. Het antwoord op die vraag hangt samen met de manier waarop wij in het leven staan, met de levensbeschouwing die we hebben.

Gespreksvraag: Wat vindt u belangrijker: hoge prestaties of met plezier naar school gaan? Of is dit geen tegenstelling?

Zelfs bij rekenwiskunde is de vraag 'wat is goed onderwijs' een normatieve vraag. De discussie over 'realistisch rekenen' versus 'koopmansrekenen' laat dit zien. Realistisch rekenen legt de nadruk op het toepassen van rekenvaardigheden in levensechte contexten. Het focust niet alleen op juiste antwoorden, maar benadrukt dat voor leerlingen mathematisch denken een vorm is van kritisch denken (zie het kader in hoofdstuk 2 over leerling Christiaan). Bij 'koopmansrekenen' wordt gehecht aan technisch correct kunnen uitvoeren van optellen, aftrekken, delen enzovoort. In een tijd waarin iedereen met een of meerdere rekenmachines op zak loopt, wordt in het 'realistisch rekenen' meer waarde gehecht aan wat we noemen 'schattend rekenen' dan aan 'koopmansrekenen'. Angsten voor niveaudaling leidden echter tot een situatie waarin de Koninklijke Akademie van Wetenschappen (KNAW) eraan te pas moest komen om de vete tussen de voor-en tegenstanders van het realistisch rekenen te beslechten (KNAW, 2009).

De Akademie onderzoekt eerst of de prestaties inderdaad gedaald zijn. Die vraag is niet zomaar met ja of nee te beantwoorden. Door de tijd heen zijn de doelstellingen van het rekenonderwijs veranderd. 'Schattend rekenen' is belangrijker geworden terwijl 'koopmansrekenen' minder belangrijk is geworden. De Akademie laat zien dat de leerling-prestaties een mooie weerspiegeling vormen van deze verschuivende doelstelling. In de loop van de tijd zijn leerlingen beter in 'schattend rekenen' geworden en minder goed in 'koopmansrekenen'. Er is geen sprake van niveauverlaging maar van verandering van doelstellingen. Of die verandering gewenst is, hangt samen met wat de maatschappij belangrijk vindt. Ook hier blijkt dus dat de vraag 'wat is goed onderwijs' een normatieve vraag is.

De KNAW (2009) had op basis hiervan kunnen concluderen dat er alleen sprake is van een verschuiving in prioriteitstelling maar de KNAW durfde niet in

tevredenheid te eindigen. “Zeker als je je realiseert hoe er tegenwoordig door onze concurrenten in Azië wordt gerekend, zou tevredenheid dodelijk zijn”, schrijft de KNAW. In het persbericht bij het rapport en in de kranten NRC (2009) en Trouw (2009) werd vervolgens toch gesteld dat ‘het’ niveau gedaald was. Dat tegenover een achteruitgang in ‘koopmansrekenen’ een vooruitgang in ‘schat-tend rekenen’ stond, werd niet vermeld. De complexere werkelijkheid, die de normatieve vraag opwierp wat in het rekenonderwijs van de eenentwintigste eeuw prioriteit verdient, raakte zo op de achtergrond.

Dertig jaar effectiviteitsonderzoek

Toen de commissie Dijsselbloem in 2008 het belang van *evidence-based* onderzoek benadrukte, leek het alsof daarmee een nieuwe benadering gepromoot werd. In werkelijkheid bestond deze benadering al enkele decennia. De Amerikaan Thomas L. Good was één van de onderzoekers die aan de basis van deze onderzoekstraditie stond. Ruim veertig jaar deed hij onderzoek naar effectieve leerprestaties. Onlangs blikte hij terug op wat er bereikt is (Good, 2010).

Voor wat betreft een aantal factoren is uit het onderzoek gebleken dat ze een gunstig effect op de leerprestaties hebben. Leraren die inhoudelijke vakkennis hebben, leerlingen helpen om iets echt te snappen, op een warme en proactieve manier orde houden, helder zijn en formatieve feedback geven, bereiken bijvoorbeeld meer dan leraren die dat allemaal niet hebben, zijn of doen. Zo zijn er vele andere factoren te noemen die dankzij onderzoek verhelderd zijn. Good vindt dit belangrijke informatie. Tegelijkertijd is hij teleurgesteld. In al die jaren is veel minder bereikt dan hij had gehoopt.

Om uit te leggen waar dit 'm in zit, vergelijkt hij leraren met koks in een restaurant. Als op het menu kip met salade staat, een glaasje wijn en een stuk appeltaart als toetje, dan begrijpt iedereen onmiddellijk dat dit gerecht in een kwaliteitsrestaurant niet hetzelfde smaakt als in een snackbar. Alle koks gebruiken dezelfde basis-ingrediënten maar de kwaliteit van hun gerechten verschilt. Zo is het ook in het onderwijs. Het *evidence-based* onderzoek heeft niet geleid tot duidelijke uitspraken over – om in de beeldspraak te blijven – wat er op het menu hoort te staan. Goede en minder effectieve aanpakken blijken uit dezelfde basis-ingrediënten te bestaan, maar goede leraren gebruiken diezelfde basiselementen op een betere manier.

Het valt Good ook op dat kenmerken van onderwijs in de ene situatie onderdeel zijn van de oplossing terwijl diezelfde kenmerken in een andere situatie onderdeel van de kwaal lijken te zijn. Een goede balans vinden is bepaald niet eenvoudig, stelt hij, mede omdat leraren en leerlingen zo van elkaar verschillen. In plaats van dat onderzoek aan de slingerbewegingen tussen een instrumentele en een vormingsbenadering een einde maakt, ziet hij dat onderzoek deze slingerbewegingen volgt. Als in de onderwijspraktijk van vandaag de balans is doorgeslagen naar de ene kant, kan onderzoek dit laten zien. Dit effect is echter niet blijvend als in de onderwijspraktijk van morgen de slinger is doorgeslagen naar de andere kant.

Biesta (2007) brengt dit punt heel kernachtig onder woorden: *evidence-based* onderzoek laat niet zozeer zien wat effectief is, maar onderzoek laat zien wat effectief *is geweest*. Creemers en Slegers, twee Nederlandse onderzoekers die hun sporen verdiend hebben op het gebied van effectiviteitsonderzoek, zeggen eigenlijk hetzelfde. Kennis over wat effectief is geweest, is nog niet hetzelfde als kennis om effectief te worden (Creemers & Slegers, 2003).

Dit verschijnsel komt vaker voor, ook in de economie. Daar geldt het motto: behaalde resultaten in het verleden bieden geen garantie voor de toekomst. Dit motto is ook op onderwijs van toepassing. En zo zijn we weer terug bij af. De commissie Dijsselbloem probeerde om boven de partijen te gaan staan, maar was zelf onderdeel van een slingerbeweging. Ook onderwijsonderzoek biedt geen definitief antwoord. Het biedt wel verheldering, zoals we op heel veel plaatsen in dit boek laten zien. Maar de gegevens over 'wat werkt' zijn als ingrediënten die in de supermarkt te koop zijn: ze zijn noodzakelijk, maar om daarvan een smakelijke maaltijd te maken, zijn creativiteit en vakmanschap onmisbaar. Onderzoek over 'wat werkt' leidt wel tot kennis over onderwijs (begripkennis die in het hoofd zit), maar niet tot kennis van onderwijzen (de vaardigheid om in een complexe situatie te handelen). Voor dat laatste is het vakmanschap van de docent nodig (Bulterman-Bos, 2012).

Gespreksvraag: Worden er op uw school keuzes gemaakt met het argument dat de keuze *evidence based* is. Wat vindt u daarvan?

Anders dan een aantal kwaliteitskranten deed vermoeden, concludeerde de KNAW niet dat leraren de schuld waren van dalende rekenprestaties (KNAW,

2009). De KNAW zei dat er geen relatie gevonden kon worden tussen de gebruikte didactiek en prestaties van leerlingen. De didactici konden hun strijd bijl dus beter begraven. De leraar is belangrijker dan de gebruikte didactiek. De KNAW noemt alleen maar dat leraren beter opgeleid zouden moeten worden, maar stelt leraren niet expliciet aansprakelijk voor de dalende prestaties. Dat laatste hebben journalisten ervan gemaakt. Zo kwam er, uit naam van de Akademie, in de kwaliteitskranten Trouw (2009) en NRC (2009) te staan dat de schuld van de dalende prestaties bij leraren ligt, terwijl het rapport nog niet eens hard had gemaakt dat die daling werkelijk bestond (zie kader hierboven).

De cruciale betekenis van de docent

Onderwijs en opvoeding hebben onmiskenbaar levensbeschouwelijke elementen, zo bleek uit onze analyse. Het gaat niet zozeer om de vraag of een aanpak wordt gebruikt, maar hoe deze praktisch wordt ingevuld. Bovendien hebben aanpakken altijd voordelen op het ene vlak en nadelen op het andere vlak. Het gaat om het vinden van de juiste balans. Eenvoudig is dat niet. Beide paradigma's benadrukken facetten van de onderwijswerkelijkheid die waardevol, maar niet altijd even goed te verenigen zijn. Het vormende paradigma benadrukt hoe leerlingen van elkaar verschillen terwijl het instrumentele paradigma benadrukt hoe ze op elkaar lijken. Het vormende paradigma benadrukt de eigen talenten van de studenten, terwijl het instrumentele paradigma benadrukt dat standaarden juist helpen om die talenten tot ontwikkeling te brengen. Vorming wil ruimte scheppen voor eigen verantwoordelijkheid van de student terwijl de instrumentele benadering deze student stuurt in de richting van algemene, vaak zinvolle doelen. Logischerwijs zijn beide paradigma's paarden die voor dezelfde kar staan. Beide benadrukken zinvolle doelen maar de doelen liggen in totaal uiteenlopende richtingen. Zo ontstaat een voortdurend spanningsveld.

Sinds mensenheugenis gaan mensen al met dit spanningsveld om. Toen de pedagogiek ontstond aan het begin van de vorige eeuw, werd dit spanningsveld verwoord met de bekende tegenstelling *Führen oder Wachsen Lassen*. De schrijver van het gelijknamige boek, Theodor Litt, vertelt hoe er tijdens een conferentie in 1927 in Weimar twee groepen aanwezig waren. Een groep benadrukte het belang van sturing (*Führen*); een andere het belang van leren in vrijheid (*Wachsen Lassen*). Theodor Litt beschrijft hoe hij enorm enthousiast raakte toen de voorstanders van *Führen* aan het woord waren. Maar toen de voorstanders van *Wachsen Lassen* aan de beurt waren, raakte hij éven



enthousiast. Dit inspireerde hem tot een boek dat een standaardwerk in de pedagogiek zou worden: *Führen oder Wachsen Lassen* (1927/1965). De benaderingen mogen dan tegenstrijdig zijn aan elkaar, maar, zo concludeerde hij, ze zijn allebei waardevol.

Deze paradox bestaat al zolang er opvoeding bestaat. Omgaan met deze paradox is niet altijd gemakkelijk. Er ontstaat echter pas écht een probleem wanneer opvoeding en onderwijs als gevolg van het moderne wetenschappelijke denken louter rationeel worden benaderd en autoriteiten van buitenaf proberen om het onderwijs te verbeteren. In die wereld trekken de twee paarden de kar uit elkaar omdat er voldaan moet worden aan eisen die volgens de logica met elkaar in strijd zijn. Als we geluk hebben, wordt er op tijd ingegrepen en blijft de kar nog heel maar maakt wel een enorme slingerbeweging.

Toch zijn er ook in het wetenschappelijke domein steeds onderzoekers geweest die hebben benadrukt dat onderwijs de logica te boven kan gaan. Niet toevallig gebeurde dat door een onderzoeker, Magdalene Lampert (1985), die

haar onderzoekswerk aan universiteiten altijd heeft gecombineerd met werken op een basisschool. Ze laat zien hoe zij als lerares voortdurend te maken heeft met tegenstrijdige doelstellingen die wel allebei belangrijk zijn. Het is niet of-of maar en-en. Meisjes wil ze evenveel aandacht geven als jongens omdat ze ook voor meisjes excellentie in wiskunde nastreeft. Helaas, het zijn de jongens die steeds lastig zijn en dus haar aandacht opeisen en dat terwijl ze jongens niet meer aandacht wil geven dan meisjes. Zo ervaart ze ook een dilemma tussen vrijheid en standaardisatie, tussen de leerlingen pushen om te presteren en zorgen dat leerlingen met plezier naar school gaan. In dergelijke situaties bestaat geen 'juiste oplossing'. De tegenstrijdige doelen zorgen voor een voortdurende innerlijke spanning. Het is improviseren en de eindjes aan elkaar knopen. Ze noemt dit 'managen'. Ook Bruner (1996) onderkent dat het onderwijs vol is met dergelijke spanningsvelden waarbij logische oplossingen niet bestaan. Hij benadrukt echter wel dat sommige docenten erin slagen om praktische oplossingen te vinden.

De spanning tussen de pedagogische functie van het onderwijs die het belang van het leerproces benadrukt aan de ene kant, en de kwalificerende functie die selectie benadrukt aan de andere kant, ervaart docent rekenen Niek aan den lijve. Als studenten op de pabo aankomen, wil hij werken aan de houding ten aanzien van het vak. Vaak merkt hij dat studenten rekenen niet leuk vinden omdat ze (denken dat ze) het niet kunnen. Niek begint vanuit een positieve insteek. Iedereen kan met zijn voorkennis problemen oplossen, vaak kom je alleen daarmee al heel ver. Hij legt de nadruk op wat goed ging en niet teveel op wat ontbreekt. Echter in de eerste periode worden deze studenten keihard afgerekend bij de instaptoets. Bij een digitale toets is iets goed óf fout en niets bijna goed. Studenten moeten een score van 103 halen. Zo niet, dan zijn ze gezakt. Van dergelijke negatieve ervaringen worden hun rekenvaardigheden er niet beter op!

In het derde leerjaar van de pabo krijgen studenten een grote opdracht waarbij ze zelf leerdoelen moeten opstellen voor de stage. Doordat studenten zelf hun eigen leerdoelen mogen bepalen, wordt voorkomen dat zij theorie moeten toepassen die in hun specifieke stage niet van toepassing is. De studenten geven zelf aan met welke leerdoelen en met welke theorie ze aan het werk willen gaan. Het geheel is dus gericht op vorming en eigenaarschap van hun eigen leerproces. Dit slaagt bij een deel van de studenten maar docent Anita merkt dat het resultaat sterk afhankelijk is van de mate waarin studenten zich inzetten. De

student die actief hieraan werkt, leert veel en kan ze positief beoordelen. Een ander deel van de studenten toont onvoldoende inzet en doet maar wat. Zij leren weinig. Wel schrijven ze bladzijden vol. Als docent voelt zij het spanningsveld: voldoende of onvoldoende geven? En wat dan? Er moet hard afgerekend worden met een 'zacht' instrument. Uiteraard moeten studenten hun werk verbeteren, maar Anita merkt zelden dat dit tot echte kwaliteit leidt. Het doel is dan niet bereikt. Toch hebben de studenten in formele zin de opdracht correct uitgevoerd. Ze voelt dat ze bijna niet anders kan dan toch maar een voldoende geven.

Het pedagogische doel van onderwijs dat evalueren om te leren benadrukt, staat dus in een constante spanningsverhouding tot de kwalificerende functie van onderwijs die veronderstelt dat er gemeten, beoordeeld en een diploma uitgedeeld wordt (Bulterman-Bos, Verloop, Terwel en Wardekker, 2003; Biesta, 2012). De verhouding is paradoxaal. Meten kan de kwaliteit van onderwijs zowel frustreren als bevorderen. Iedere docent is het eens met de stelling dat evalueren het leerproces moet ondersteunen. Iedere docent wil graag zo eerlijk mogelijk beoordelen en voelt daarom verwantschap met het instrumentele paradigma. Toch weet iedere docent ook dat beoordelingen studenten zowel kunnen aanmoedigen als ontmoedigen. De 'hardheid' van sommige instrumenten is voor bepaalde studenten 'te hard', terwijl de 'zachtheid' van andere instrumenten voor bepaalde studenten 'te zacht' is. Standaarden kunnen studenten zowel uitdagen als vervreemden van zichzelf. De kwalificerende functie van onderwijs, die serieus genomen wordt door het instrumentele paradigma, kan soms heftig strijden met die andere functie die persoonsvorming benadrukt. Beide zijn belangrijk, maar het is vaak improviseren om ervoor te zorgen dat de twee elkaar niet in de weg zitten. Er zijn terreinen in het leven waar wetenschap kant en klare oplossingen kan bieden. Maar er zijn ook terreinen waar oplossingen van buitenaf niet volstaan en professionals zelf in hun eigen praktijk naar oplossingen moeten zoeken. Het (beroeps)onderwijs is zo'n terrein.

Gespreksvraag: In dit gedeelte wordt terugverwezen naar de drie functies van onderwijs, zoals uitgelegd in hoofdstuk 1 (de kwalificatie, persoonsvormende en de socialiserende functie). Hebt u wel eens te maken met tegenstrijdige doelstellingen die u allebei belangrijk vindt? Gaat dit gepaard met innerlijke spanningen?

Omgaan met paradoxen: voorbeelden

Hierboven hebben we gezien dat docenten in de praktijk soms dingen voor elkaar krijgen die op papier niet kloppen. Ze weten dilemma's te managen en tegenstellingen te verenigen. Kant-en-klare oplossingen zijn er niet. Dat maakt het beroep aan de ene kant zwaar en geeft haar aan de andere kant een wonderlijk mooi karakter. Als het lukt om tegenstellingen te overbruggen, ervaren docenten een grote vreugde in hun werk. Lukt het niet, dan kan dit het gevoel geven pijnlijk door de grond te gaan. De dilemma's van docenten liggen op allerlei vlakken. De uitdaging is om de spanning op een creatieve manier aan te gaan. Hieronder volgen enkele voorbeelden van docenten die dat doen en zo tegenstellingen overbruggen.

Boukje geeft Engels aan een eerste klas van het voortgezet onderwijs waarin het niveau uiteen loopt van vmbo tot vwo. Gemakkelijk is dit niet: de vwo'ers kunnen alles veel beter. Als ze niet uitkijkt, zitten de vmbo-leerlingen er voor spek en bonen bij en er is maar weinig voor nodig om hen het gevoel te geven dat ze niets voorstellen. Sterke en zwakke leerlingen zijn zoveel mogelijk naast elkaar gezet. Vooral bij mondelinge dialogen is dat handig. Al vrij vroeg in het jaar geeft Boukje huiswerk op twee niveaus. Niet alle leerlingen hoeven alles te maken. Toetsen zijn ook op twee niveaus. Gelukkig is Boukje dol op de vmbo-leerlingen. Ze zijn heel lichamelijk ingesteld en spontaan. Bij dit soort kinderen wil ze ook zelf nog wel eens een vriendschappelijke stomp uitdelen. Dan kan zo'n kind er ook weer tegenaan. De sfeer in de klas is vrolijk. Ze is altijd optimistisch over wat leerlingen kunnen leren. Sommige vmbo-leerlingen maken meer huiswerk dan strikt noodzakelijk is. Als ze boven hun niveau scoren, maakt Boukje daar een act van. Ze vindt het heerlijk om de hele klas, inclusief de zwakkere leerlingen, mee te slepen naar een hoger niveau. Al haar creativiteit en persoonlijke vaardigheden zet ze in om dit voor elkaar te krijgen.

In het geval van Boukje lijken tegenstellingen te verdwijnen. Terwijl het instrumentele paradigma kinderen onmiskenbaar opdeelt naar niveau, wil zij alle kinderen accepteren zoals ze zijn. Logisch gezien gaan deze twee niet samen maar Boukje weet een sfeer te scheppen waarin de kinderen zich zowel geaccepteerd voelen als hun niveau verhogen. Ze doet dit met haar persoonlijkheid (Lampert, 1985). Bijzonder is ook de vanzelfsprekendheid waarmee dit gebeurt, alsof het zo hoort. Het hoort helemaal bij Boukje.

Boukje maakt gebruik van de theorie over differentiëren, maar overstijgt deze theorie ook. Ze differentieert, maar doet ook het omgekeerde: ze integreert

alle kinderen in dezelfde groep (Bulterman-Bos, 2007). Daarvoor bestaat geen standaardtechniek. Door de manier waarop ze inspeelt op allerlei ‘gewone’ situaties in de klas, hebben zwakkere leerlingen niet het gevoel dat ze minder zijn dan andere leerlingen. Sterker nog, ze hebben het gevoel dat ze met de sterken mee mogen doen. Terwijl Boukje alle leerlingen wil accepteren zoals ze zijn, ontstaat er geen sfeer dat alleen je persoonlijkheid telt terwijl de schoolprestaties er niet toe doen. De spanning tussen de paradigma’s wordt een creatieve spanning die energie voortbrengt.

Spanningsvelden bevinden zich ook op het niveau van de leerstof. Studenten vinden de voorgeschreven leerstof lang niet altijd interessant. Toch slagen goede docenten erin om hun studenten te bereiken. De situatie kantelt: wat niet interessant leek, kleurt naar wél interessant. Het is geen techniek die docenten toepassen. Er gebeurt niet alleen iets op het niveau van de docent, maar er gaat ook iets uit van de studenten. Er komt iets tot stand, er gaat iets stromen, er gebeurt iets. De docent en de studenten krijgen er beide energie van. Hier bewijst zich het wezen van goed onderwijs.

Op basisschool De Morgenster worstelen de leraren met bijbelverhalen. Het bestuur heeft bepaald dat de christelijke identiteit gehandhaafd moet blijven maar niet iedereen weet daar raad mee. Weinig kinderen hebben van huis uit iets met godsdienst. Veel kinderen vinden de bijbel saai. De wonderlijke verhalen passen niet in hun leefwereld. Ook veel leerkrachten weten niet wat ze met die vreemde verhalen aan moeten vangen. Ze pakken liever een gewoon voorleesboek. In de praktijk komt van de godsdienstlessen dus niet altijd evenveel terecht.

Toch, juf Judith heeft onlangs volop Pasen gevierd met de kinderen. Het begon met een uitzending van schooltv waarin getoond werd hoe de thematiek ‘nieuw leven’ in veel godsdiensten voorkomt. Ook schilderijen van de kruisiging kwamen daarin voor. “Dat vind ik zo eng!”, klonk er. “Ja, dat zijn hele nare beelden”, bevestigde de juf, om vervolgens de kinderen te bepalen bij het verdriet onder de vrienden van Jezus. Van daaruit navigeerde ze naar de zorgzaamheid waarmee ze het lichaam van Jezus van het kruis haalden en neerlegden in een mooie tuin waar zich in de rotsen een kamertje bevond waarvoor ze een grote steen neerlegden.

Het verhaal ging over de ontzetting van Maria, die zag dat de steen was weggerold en dat het graf leeg was. Het verhaal ging ook over het ongeloof dat

Maria ontmoette toen ze de andere vrienden vertelde dat ze Jezus in de tuin was tegengekomen terwijl ze hem door haar tranen niet eens had herkend.

“Jullie weten wel”, zei juf Judith tegen de kinderen, “dat deze school het héél belangrijk vindt dat álle kinderen zich thuis voelen. Of je nu gelooft of niet. Sommige mensen geloven in God, anderen niet. Maar dit verhaal was zélfs voor de vrienden van Jezus moeilijk te geloven!”

De juf dacht terug aan de Matthäus Passion, die ze net die week gehoord had en ging nog even door op de vrienden van Jezus: “Ze konden het heel moeilijk geloven! Maar toen ze ervan gingen zingen, konden ze de blijdschap toch meevoelen.”

Daarop klonken er twee liedjes die de kinderen weldra meezongen. Dankzij hun eenvoud en hun aanstekelijke melodie sloegen ze in als een bom. De bedrukte sfeer rondom de beelden van de kruisiging en het verdriet van de vrienden was ineens veranderd. Even was er nog discussie onder de kinderen over de vraag welke mooier was: het liedje over Maria die een leeg graf vindt of het liedje over de lente. Beide gingen in de herhaling en werden nogmaals gedraaid tijdens het fruit eten. Het was volop feest in de klas!

Na afloop van de les vroeg Judith zich af wat ze nu eigenlijk met dit soort lessen wil. Ze realiseerde zich hoe goedgegelovig kinderen uit groep drie nog zijn. Werkelijkheid en fantasie lopen voor deze leeftijdsgroep makkelijk in elkaar over. In die zin is het gemakkelijk scoren.

Ze beoogt niet dat kinderen hun verstand uitschakelen maar ze is erop uit dat verhalen als deze iets voor de kinderen zullen betekenen als het zélf moeilijk hebben. Ze wil dat haar leerlingen zich dan niet zullen aansluiten bij het leger der ontevreden burgers, maar hoop zullen koesteren en uitstralen. Natuurlijk weet ze wel dat daarvoor meer nodig is dan dit ene verhaal. Toch is het haar streven dat dit verhaal op een niet-moralistische manier bijdraagt aan de vorming van vonkjes geestkracht. Opdat de leerlingen niet zwichten voor moeilijkheden, óók onder omstandigheden waarin dit rationeel gesproken onrealistisch is.

Ze kijkt nog eens terug op het paasfeest. Veel stelde het allemaal niet voor. Er waren geen cadeautjes, geen lekkernijen en nauwelijks versiering. Toch was het paasfeest in groep drie. Het ware feest raakt altijd een diepere laag!

Juf Judith overbrugt de tegenstelling die ontstaat doordat ze met leerstof aan de slag moet waar de leerlingen niet om gevraagd hebben en weinig gevoel voor hebben. Ze wil dit doen zonder de leerlingen te forceren. Ze neemt zowel het paasverhaal als de beleavingswereld van de kinderen serieus. Die twee weet ze bij elkaar te brengen door persoonlijke kwaliteiten in te zetten. Ze kan mooi vertellen waardoor het verhaal tot leven komt. Muziek versterkt de gevoelswaarde. De droevige stemming slaat om in een feeststemming. Judith heeft voor deze middelen gekozen vanuit visie op het vak en vanuit een professionele houding waarin ze als vanzelfsprekend op haar eigen praktijk reflecteert. In eerste instantie overbrugt Judith de tegenstelling dus doordat ze een vakvrouw is.

Toch is ook duidelijk dat juf Judith dit paasfeest niet zelf fabriceert. Net als haar leerlingen ontvangt ze het als een geschenk. In de klas wordt iets aange-reikt vanuit een hoger niveau waarop tegenstellingen verdwijnen. Dit hogere inspireert zowel leerlingen als leraren. Nieuwe energie gaat stromen. Heel vaak is dit het moment bij uitstek waarop docenten weer begrijpen waarom ze voor dit vak gekozen hebben.

Het voorbeeld van juf Judith gaat over een godsdienstles, maar dat is niet belangrijk. Bij alle vakken kan tussen docent en student een vonk ontstaan waardoor de stof gaat spreken en beide nieuwe energie ontvangen. Het vindt evenzeer plaats op (hoge)scholen van niet-religieuze signatuur als op religieus georiënteerde scholen. Deze 'hogere dimensie' is niet op bestelling leverbaar maar manifesteert zich op allerlei plaatsen: tussen studenten en docenten, tussen docenten onderling of in contacten met ouders. Vrijwel iedere docent herkent dit.

Uit een reflectieverslag van een pabo-student: "Niet voor niets had ik de hele week concentratie verzameld om de teugels strak in handen te houden. Al direct bij het eerste onderdeel had ik relatief veel gevraagd qua discipline: geen geklets, geen gemier, geen gedoe met elkaar, rechtop zitten, naar het bord kijken. Als kinderen van het rechte pad afweken, was ik daar meteen opgedoken. En ziedaar: de strenge hand leek heilzaam te zijn. Toen ik met mijn leesgroepjes klaar was, voelde ik de heerlijke sfeer van concentratie, van samen hard werken, alsof ik mij te midden van nijvere bijen bevond die in alle vrijheid rondvliegen maar intussen ook werken aan raten die druipen van de honing. Geen kind dat geforceerd werd, geen kind dat niet gemotiveerd was, geen kind dat ongelukkig leek. Iedereen had zijn eigen leerweg, er werd door de klas gelopen en er werd geleerd. Toch weer die oude waarheid: gezag schept vrijheid?"



Gespreksvraag: Kunt u voorbeelden geven van momenten waarop u energie van uw studenten ontving?

Innerlijke ontwikkeling

Binnen een rationeel wereldbeeld is geen ruimte voor de innerlijke spanning van docenten die tegenstrijdige doelstellingen willen verenigen. Het gaat alleen maar om uiterlijk waarneembaar gedrag, waarbij uitgegaan wordt van de rationele wet dat iets en zijn tegendeel niet tegelijkertijd waar kunnen zijn. Toch zijn het de goede docenten die innerlijke spanningen aangaan en tegenstellingen overbruggen.

Het belang van de innerlijke groei van docenten is heel beeldend beschreven door Parker Palmer. Zijn boek *The courage to teach* (1998) is in het Nederlands vertaald onder de titel *Leraar met hart en ziel* (2005). In dit boek gaat Palmer in op de innerlijke spanningsvelden van de docent. Heel eerlijk vertelt Palmer hoe moeilijk hij het soms vindt om studenten te motiveren. In de loop der jaren heeft hij ontdekt hoezeer hijzelf het instrument is dat hem daarbij tot beschikking staat. “We geven les in wie we zijn.” Docenten die niet zichzelf zijn, bieden hun studenten een slap aftreksel van andermans persoonlijkheid. Dat werkt niet. Docent zijn is voor hem niet zomaar werk maar heeft alles te maken met zijn eigen persoonlijke groei. Pijnlijke confrontaties met zichzelf zijn onvermijdelijk. Hij leerde inzien hoe hij vaak bezig was met de vraag hoe hij overkwam op studenten in plaats van met de vraag hoe hij hun leerproces bevorderde en beschrijft hoe studenten die hij niet kan bereiken zijn ogen openen voor een grotere wereld dan hij voorheen zag. Lesgeven is voor hem omgaan met dilemma’s. In dit verband verwijst hij graag naar een uitspraak van de natuurkundige en Nobelprijswinnaar Niels Bohr: “Het tegenovergestelde van ‘waar’ is ‘onwaar’, maar het tegenovergestelde van een diepe waarheid kan een andere diepe waarheid zijn.” Opvoeding en onderwijs vragen van ons dat we ons bewegen naar het niveau van diepere waarheden. Deze diepere waarheden doen niet zozeer een beroep op ons begripsvermogen, maar ze doen een beroep op onszelf. Dit legt Palmer uit door naar Schumacher te verwijzen:

“Ons leven lang worden wij geconfronteerd met het probleem om tegenstellingen met elkaar te verzoenen die logischerwijs niet verzoend kunnen worden (...) Hoe kunnen we de eis van vrijheid en die van discipline in het onderwijs samenbrengen? Talloze moeders en leraren doen dat in feite, maar niemand kan de oplossing onder woorden brengen. Ze doen het door in de situatie een

kracht in te schakelen die van een hogere orde is, waar tegenstellingen worden opgeheven – de kracht van de liefde (...) Problemen die ons in verschillende richtingen sturen, noodzaken ons als het ware ons in te spannen om boven onszelf uit te stijgen; zij vragen om krachten van een hoger niveau, roepen die daarmee ook op en brengen daarmee liefde, schoonheid en waarheid in ons leven. Alleen met behulp van deze hogere krachten kunnen de tegenstellingen worden verzoend in de bestaande situatie” (Palmer, 2005, p. 74 met verwijzing naar Schumacher, 1973).

Een techniek die hierbij de juiste oplossing levert, is niet voor handen. Aansluitend bij Schumacher schrijft Parker Palmer: *“Lesgeven en leren vereist een hogere graad van bewustzijn dan waar wij gewoonlijk in verkeren – en bewustzijn is altijd groter in een toestand van creatieve spanning. Een andere naam voor die spanning is paradox, een manier om tegengestelde polen bij elkaar te houden zodat er een elektrische lading ontstaat die ons wakker houdt. Het is niet zo dat alle goede leraren dezelfde techniek gebruiken, maar welke techniek ze ook toepassen, ze vinden altijd een manier om die creatieve spanning op te wekken” (Palmer, 2005, p. 66).*

Goed onderwijs vraagt om goede leraren. Zonder dat we de waarde van systematiek willen afdoen, is het duidelijk dat goed onderwijs niet in een systeem te vangen is. Zowel bij het instrumentele paradigma als bij het vormingsparadigma kunnen belangrijke vragen gesteld worden. Bij het instrumentele model is er de vraag: ‘Welke leerstof moet de student kennen?’ Bij het vormingsmodel is er de vraag: ‘Wie is de student?’ De vraag: ‘Wie is de leerkracht?’ overstijgt deze vragen en is wezenlijk bij het aangaan van de spanning die tussen de beide paradigma’s bestaat.

Gespreksvraag: Palmer zegt dat docenten hun studenten alleen kunnen bereiken vanuit hun eigen identiteit en integriteit. “We geven les in wie wij zijn.” Neem een moment in gedachten waarop u heel veel plezier had in het lesgeven. Wat vertelt dit moment over wie u bent?

‘De docent doet ertoe’ als nieuwe onvrijheid

De docent doet ertoe! Deze wijsheid is tegenwoordig overal te horen omdat onderzoek een waarheid bevestigt die al sinds mensenheugenis bekend is. Dit kan een erkenning betekenen voor docenten die noodzakelijk is nadat de betekenis van docenten lange tijd is genegeerd.

Tegelijkertijd kan het ook betekenen dat alles wat er eventueel mis is aan het onderwijs op het bordje van docenten geschoven wordt. De conclusie dat leraren ertoe doen, wordt dan ingekapseld in een wereldbeeld waarin kwaliteitseisen almaar toenemen en faciliteiten afnemen. Zo kan de conclusie dat de docent ertoe doet, leiden tot een nieuwe manier om docenten hun professionele ruimte af te nemen. De basisingrediënten voor goed onderwijs, beschreven onder anderen door Thomas Good (2010), Marzano (2003) en Hattie (2009), worden bijvoorbeeld gebruikt om docenten alsnog lastig te vallen met voorschriften en checklisten van buitenaf. Opnieuw worden docenten dan ingekapseld in een systeem. Alsnog moeten docenten voldoen aan anonieme blauwdrukken. De innerlijke kwaliteiten van docenten worden genegeerd en aan innerlijke spanningen wordt voorbijgegaan.

De conclusie dat de docent ertoe doet, betekent juist dat systemen hun beperkingen hebben. De lessen die docenten geven, zijn spiegels van hun ziel. Het beroep van docent is een creatief beroep in meerdere opzichten. Enerzijds is er creativiteit nodig om tegenstellingen te overbruggen wat laat zien dat lesgeven een kunst is en niet alleen een ambacht. Anderzijds koesteren docenten mensen-in-ontwikkeling. Studenten zijn complexer dan het leven zelf. Om studenten te kunnen begrijpen moeten docenten ook voortdurend in ontwikkeling zijn. Als een student iets geleerd heeft of als de docent iets over de student geleerd heeft, is een nieuwe werkelijkheid tot stand gekomen die tot voor kort nog niet bestond. Docenten en studenten zijn met huid en haar in dit vernieuwingsproces betrokken. Geen systeem dat dit kan omvatten.

Normatieve professionaliteit

Hierboven zijn positieve voorbeelden genoemd waarin het docenten lukte om tegenstellingen te overbruggen. Dat dit niet gemakkelijk is, zal iedere docent beamen. Studenten zijn soms moeilijk te bereiken. Hun hoofd staat soms naar totaal andere dingen dan de docent in de aanbieding heeft. Ze leveren werk in waarvan docenten op hun klompen aanvoelen dat er veel meer uitgehaald had kunnen worden maar ze kunnen niets doen omdat studenten formeel aan de criteria voldoen. Docenten weten van worstelingen van maandenlang werken aan een positief zelfbeeld bij hun studenten dat met één pennenstreek door toetsresultaten onderuitgehaald wordt. Officieel zouden docenten hier niet van wakker te hoeven liggen. Ze doen hun werk, houden zich aan de afspraken en dat moet het maar zijn. Toch zitten veel docenten zo niet in elkaar. In dergelijke situaties ervaren ze dat werken in het onderwijs een enorm beroep

doet op hun persoonlijkheid. De innerlijke spanning gaan ze actief aan. Zij voelen zich verantwoordelijk voor de toekomst van hun studenten.

Met hun ethische besef tonen docenten zich wat Kunneman noemt ‘normatieve professionals’ (Kunneman, 2005; Van Houten, 2008). Daarmee wordt bedoeld dat professionals (in dit geval docenten) te midden van voorschriften en regeldruk zich bewust zijn van hun morele verantwoordelijkheid voor de mensen waarvoor ze werken (in dit geval leerlingen of studenten). Leraren ontwikkelen een liefdevolle verhouding tot de innerlijke complexiteit in hun werk. Ze lopen niet weg voor de innerlijke spanningen, en voor wat Kunneman de ‘plekken der moeite’ noemt maar gebruiken die als het ware als een laboratorium voor vernieuwing van de praktijk en van zichzelf. De kwaliteiten die ze aldus ontwikkelen valt niet in systemen uit te drukken. Dit proces duidt Kunneman (2005) aan met ‘normatieve professionalisering’. Normatieve professionalisering verwijst naar een voortdurende activiteit, een nooit afgerond proces waarin professionals vanuit waarden als zorg, vriendschap, compassie, burgerschap, rechtvaardigheid en solidariteit de systemen in hun praktijk verrijken (Van Houten, 2008, p. 72-74).

Dit begrip benadrukt dat bepaalde praktijken niet van buitenaf verbeterd kunnen worden maar vragen om de authentieke en innerlijke betrokkenheid van de professional zelf. Ook beleidsmakers – hoezeer ze ook hun best doen door goede wetten te maken – moeten het uiteindelijk hebben van de professionals die vanuit hun innerlijke motivatie verantwoordelijk willen zijn voor de praktijk.

Gespreksvraag: Wat roept de zin bij u op: “Docenten lopen niet weg voor de innerlijke spanningen, en voor wat Kunneman de ‘plekken der moeite’ noemt maar gebruiken die als het ware als een laboratorium voor vernieuwing van de praktijk en van zichzelf?”

Slot

Onderwijs is een normatieve praktijk. Ook evaluatie en toetsing zijn normatieve praktijken. Effectiviteitsonderzoek heeft geleid tot belangrijke aandachtspunten waar leraren hun voordeel mee kunnen doen. Dat neemt niet weg dat het uiteindelijk toch aankomt op de praktische invulling van de leraren zelf. Met de manier waarop ze hun lessen invullen, brengen ze zichzelf tot uitdrukking. Ze staan vaak voor dilemma’s waarin er geen oplossingen van

buitenaf gegeven kunnen worden. Ze moeten die oplossingen zelf zoeken. Dit kan gepaard gaan met innerlijke spanningen en dit doet een enorm beroep op de persoonlijkheid van docenten.

De slingerbeweging in het onderwijsbeleid is een gevolg van een poging om van buitenaf grip te krijgen op het onderwijs om de kwaliteit ervan te verbeteren. Dit gebeurt binnen een wereldbeeld waarin aangenomen wordt dat goed onderwijs geven bestaat uit het toepassen van de juiste technieken. We hebben echter juist gezien dat die technieken hun beperkingen hebben. Voor docenten maakt het er dat niet makkelijker op.

In het volgende hoofdstuk gaan we in op de vraag waar dit wereldbeeld vandaan komt en wat voor alternatieven er zijn. Een wereldbeeld waarin innerlijke groei een plaats heeft, is ook nodig om passende vormen van evaluatie en toetsing een plaats te geven.

Hoofdstuk 7

Wereldbeelden, maakbaarheid en toetsing

Kwaliteit van onderwijs is niet in een systeem te vangen, zo concludeerden we. Het onderwijs moet het hebben van de professional die van binnenuit werkt aan kwaliteit. Het beroep van docent vraagt om spelen met doelstellingen die volgens de logica met elkaar in strijd zijn. Je moet studenten sturen en vrijlaten, ze uitdagen maar ook een gevoel geven van competentie opbouwen. Telkens bevinden professionals zich in unieke situaties waar ze zelf richting aan moeten geven. We noemden dit in het vorige hoofdstuk normatieve professionalisering. Voor docenten bevat onze conclusie niets nieuws. Zij worden echter voortdurend geconfronteerd met verwachtingen alsof er kasten vol ‘instant-professionaliteit’ bestaan, die er vanzelf uitrolt zodra docenten de moeite nemen om deze kasten te openen. In dit hoofdstuk gaan we in op de vraag waar deze verwachtingen vandaan komen. Ze hangen samen met een bepaalde visie op de werkelijkheid en op wat échte kennis is en hebben daarom ook gevolgen voor opvattingen over toetsing en evaluatie. Voor de uniciteit van mensen, voor menselijke situaties en daarmee ook voor de complexiteit van opvoeding en onderwijs biedt deze visie weinig ruimte. In dit hoofdstuk wordt een alternatieve visie op de werkelijkheid gepresenteerd. Van daaruit denken we opnieuw na over toetsing en evaluatie en over de manier waarop de slingerbeweging tussen beide paradigma's overstegen kan worden.

Hogere krachten of een plat wereldbeeld?

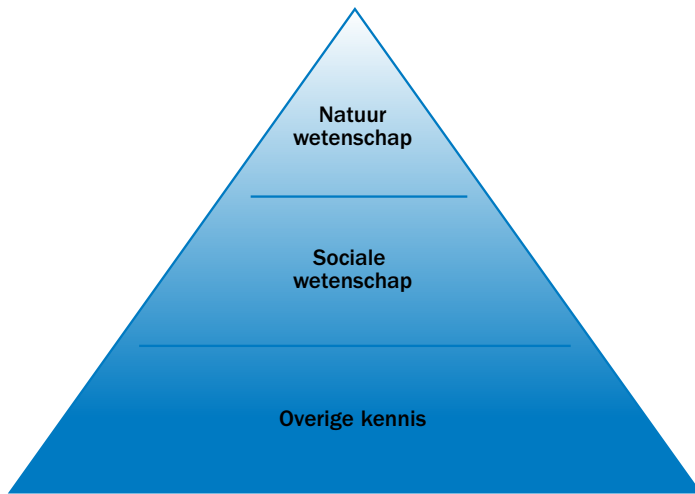
In het vorige hoofdstuk werd Schumacher geciteerd. Hij gaf aan dat veel ouders vrijheid en discipline met elkaar verenigen dankzij een kracht die zich op een hoger niveau bevindt. Een niveau waar tegenstellingen geen tegenstellingen meer zijn: de kracht van de liefde. Paradoxen dwingen ons om boven onszelf uit te stijgen, zei Schumacher. Ze maken krachten op een

hoger niveau los en zo stromen liefde, schoonheid, goedheid en waarheid ons leven binnen.

Wat bedoelt Schumacher met het 'hogere' niveau? Hij zet zich daarmee af tegen een plat, materialistisch wereldbeeld en een opvatting die niet per se samenvalt – maar wel veel gemeen heeft met – hoe Philipse het verlichtingsdenken omschrijft: “Verlichting slaat op de geesteshouding van zelfstandig en kritisch nadenken over welke bewering dan ook en op het moedige besluit om slechts geloof te hechten aan beweringen waarvoor toereikend bewijsmateriaal bestaat.” (Philipse, 2005, p. 12). Binnen dit wereldbeeld is objectieve, natuurwetenschappelijke kennis het hoogste wat er is. Dit wereldbeeld wordt materialistisch genoemd omdat men ervan uitgaat dat de wereld uiteindelijk verklaarbaar is vanuit atomen, moleculen en natuurwetten, dus materie. De werkelijkheid is in principe te vatten in een gesloten systeem: vanuit het ene aspect kan het andere voorspeld worden. Ook mensen worden beschouwd als complexe biochemische mechanismen, vergelijkbaar met computers. Tegenwoordig hebben we bijvoorbeeld vanuit het ‘breindenken’ vaak met dit wereldbeeld te maken: hoe mensen zijn, is terug te brengen tot de chemicaliën in hun hersenen, maar niet tot ‘henzelf’.

Als Schumacher spreekt over het 'hogere' niveau, wil hij daarmee aandacht vragen voor een ander werkelijkheidsbeeld. Eén waarin mensen juist geen machines zijn, maar waarin het voor ons allen een hele uitdaging is om onze weg door het leven te vinden. Dat wij met deze vraag worstelen, laat volgens Schumacher zien dat het niet om het éven is welke weg wij kiezen. Dat op zichzelf geeft al aan dat er zoiets bestaat als 'hoger' of 'lager'.

Bij het platte wereldbeeld, waarin alleen geloof wordt gehecht aan redeneringen waar bewijsmateriaal voor bestaat, hoort een visie op kennis. Natuurwetenschappelijke kennis staat bovenaan in de hiërarchie omdat daarvoor de hardste bewijzen geleverd kunnen worden. Bewijzen in de sociale wetenschappen zijn minder hard, die staan dus lager. Het grootste deel van de veelal praktische kennis en ervaringskennis waar wij ons in onze dagelijks leven op baseren maar waarvoor geen wetenschappelijke bewijzen bestaan, staat onderaan in de hiërarchie. Belangrijk in dit model is ook dat de menselijke betrokkenheid van beneden naar boven afneemt. Hoe meer wij erin slagen om ons fysiek en emotioneel te scheiden van wat wij leren kennen, hoe beter, zo veronderstelt deze visie op kennis. We noemen dit objectivisme (Palmer, 1998).



Afbeelding 7.1. Visie op kennis die behoort bij het platte materialistische wereldbeeld

Hogere krachten bestaan in dit objectivistische wereldbeeld uit (natuur)wetenschappelijke kennis. Een ‘hogere kracht’ waar onbaatzuchtige liefde volgens Schumacher een voorbeeld van is, heeft in dit model geen plaats. Vanuit dit wereldbeeld ligt het voor de hand dat het onderwijs zich moet richten op objectieve kennis en dat die kennis gemakkelijk toetsbaar is omdat het gaat om bewijsbare kennis.

Gespreksvraag: Welke betekenis heeft voor u ‘objectiviteit’? Hoe objectief vindt u de kennis die u zelf onderwijst?

De bijdrage van de verlichting

Voordat we kritische opmerkingen plaatsen bij het platte, materialistische wereldbeeld en het objectivistische kennisbeeld, is het zinvol om ook het goede te benadrukken dat het verlichtingsdenken betekent voor onze cultuur. Dankzij de verlichting heeft de medische wetenschap zich kunnen ontwikkelen: ziektes kunnen voorkomen en bestreden worden. De verlichting kwam voort uit waardevolle idealen. De grondleggers van het moderne rationele wereldbeeld, onder wie Descartes en Locke, leefden in een Europa dat verscheurd werd door godsdienstoorlogen. Ze waren naarstig op zoek naar manieren om een consensus over centrale maatschappelijke vragen mogelijk te maken en zo godsdiensttwisten te overstijgen. Daarom zochten ze naar bewijzen en onweerlegbare redeneringen. De wiskunde en de nieuwe, empirisch gerichte



Afbeelding 7.2. 'Wees ervan overtuigd dat ook in het nietigste de godheid aanwezig is'

De Waag, Amsterdam, plafond van het Theatrum Anatomicum

natuurwetenschappen boden daarvoor uitstekende modellen (Kunneman, 2005). De resultaten hiervan hebben inderdaad mensen geholpen om tot op de dag van vandaag op een zelfde manier naar de werkelijkheid te kijken.

Zorgvuldig kijken naar de werkelijkheid

Tegenwoordig leven velen in de veronderstelling dat de verlichting een poging was om godsdienst te vervangen door wetenschap. De redenering is dat mensen vroeger geen grip op de natuurlijke werkelijkheid hadden en daarom de hypothese 'God' nodig hadden. Dankzij de wetenschap werd die hypothese overbodig. Wetenschap en geloof zijn volgens deze redenering in strijd met elkaar. Charles Taylor (2007) duidt deze manier van denken aan met de 'aftreksom-theorie' en laat zien dat deze theorie onjuist is. Ook onder verlichtingswetenschappers waren overtuigd gelovigen (Stamperius, 2011). De relatie tussen geloof en wetenschap is historisch gezien veel complexer dan de aftreksom-theorie beschrijft. Belangrijke verlichtingswetenschappers vonden studie van de natuur essentieel, juist omdat zij op zoek waren naar God. Om die reden bestudeerde bijvoorbeeld Newton het licht (Shapin, 2005). De afbeelding hierboven laat dit zien. De foto is genomen in het anatomisch theater

van De Waag in Amsterdam waar chirurgijns en artsen samenkwamen om menselijke lichamen te bestuderen. De tekst op het plafond betekent: Wees ervan overtuigd dat ook in het nietigste de godheid aanwezig is. Het zorgvuldige kijken dat uit deze houding voortkwam, heeft heel veel betekend voor de ontwikkeling van de cultuur.

Kritisch denken

De verlichting heeft ook een belangrijke bijdrage geleverd aan de ontwikkeling van kritisch denken. Daarin was de verlichting overigens niet exclusief. Er hebben altijd dogmatische vormen van religie bestaan, waarbinnen kritiek niet geduld werd. Toch is kritisch denken ook een onderdeel van het christendom. De profeten in het Oude Testament leverden kritiek op de godsdienst van hun tijd. Jezus kwam in conflict met de gevestigde religieuze orde en Paulus geeft een nieuwe interpretatie aan het jodendom. Ook orthodox-protestantse kringen kennen het adagium 'Ecclesia reformata, semper reformanda': de hervormde kerk moet zich altijd blijven reformeren, hervormd (in de zin van vernieuwd) worden. Toch betekende de verlichting een nieuwe stap in de ontwikkeling van kritisch denken. Veel meer dan voorheen werden structuren ontworpen waardoor machthebbers zich moesten verantwoorden. Er werd beter nagedacht over waar beweringen op gebaseerd waren. Ook dat leverde een belangrijke bijdrage aan de ontwikkeling van de cultuur.

De prijs die betaald werd

Het wetenschappelijke denken bracht veel goeds maar er werd ook een prijs betaald (Toulmin, 2001; Kunneman, 2005). Wat sindsdien als 'rationeel' werd beschouwd, was wel heel erg smal. Wijsheden die in de loop van duizenden jaren waren ontstaan en generaties mensen verder hadden gebracht, moesten plotseling 'bewezen' worden. Zo ontstond een breuk met het verleden. Verstand en gevoel werden uit elkaar gehaald waarbij gevoelens tot een mindere categorie gingen behoren en intuïtie als een obscuur verschijnsel beschouwd werd. De persoonlijkheid van mensen was niet langer een bron die moest worden aangeboord maar een gevaar dat moest worden bezworen. Praktische kennis en vakmanschap kregen een lagere status ten gunste van theoretisch te rechtvaardigen, bewijsbare inzichten. Wat mensen in hun eigen leven door ervaring leerden, gold niet langer als voedingsbodem van werkelijk inzicht. Alleen algemene, abstracte principes golden als rationele kennis. De hele wereld werd gevat in universeel geldige inzichten en onveranderlijke

structuren. Kennis werd beperkt tot technische handelingsvoorschriften die van die onveranderlijke structuren afgeleid konden worden, als ware de hele werkelijkheid één groot mechanisme.

Door het objectivisme werd menselijke betrokkenheid gaandeweg beschouwd als een onbetrouwbare factor. Oprechte gevoelens werden afgedaan als sentiment. De verbeelding kwam als 'chaotisch' en 'onsystematisch' op de tweede rang terecht en verhalen kregen de bijklank van verzinsels. Muziek, beeldende kunst en dans werden minder belangrijk dan wetenschap. De levende werkelijkheid werd bestudeerd als ware zij levenloos (Palmer, 1998). De verlichting, die de mensheid van heerschappij probeerde te bevrijden, bracht haar eigen heerschappij voort, namelijk de instrumentele en technische heerschappij over de uiterlijke en innerlijke natuur (Hoogland, 2005). Schumacher (1977) is niet de enige die stelt dat de verlichting een eigen religie heeft geschapen, die zeker zo intolerant en totalitair is als de religies van voor de verlichting.

Gespreksvraag: Herkent u dat praktische kennis en vakmanschap (in sommige kringen) een lagere status hebben dan theoretisch te rechtvaardigen, bewijsbare inzichten? Wat vindt u hiervan?

Is intuïtie belangrijk in de onderwijspraktijk of moet een leraar zich baseren op 'bewijsbare' kennis? Josephine is een ongedurige leerling die moeilijk in het ga-reel te houden is. Haar leraar Willem zoekt naar een effectieve aanpak. Hij heeft een beloningssysteem ingesteld maar Josephine is niet onder de indruk van de sticker die ze krijgt. Wat nu? Behalve op theoretische kennis die Willem op de lerarenopleiding geleerd heeft, gaat hij af op zijn intuïtie. Hij kent Josephines ouders en schat in wat voor een gezin het is. Er is veel geld maar weinig tijd beschikbaar. Willem zoekt naar een beloning die aansluit bij Josephines behoeften en stelt voor dat als Josephine zich goed gedraagt, de leraar en zij in de pauze even gezellig met elkaar gaan kletsen. Dit blijkt beter te werken. Zijn aanpak komt tot stand door een combinatie van professionele kennis en intuïtie.

De tegenstelling tussen intuïtieve kennis en rationele kennis is minder groot dan sommigen denken. Willem voelt zich geen helderziende in die zin dat hij grip heeft op bovennatuurlijke krachten maar hij probeert 'helder te zien' door met kinderen een relatie aan te gaan. Hij kijkt goed naar kinderen, legt alle stukjes van de puzzel naast elkaar en vult de ontbrekende stukjes aan met zijn eigen ervaringskennis.

Verlichtingsmythes en maakbaarheid

Voordat we op zoek gaan naar een alternatief werkelijkheidsbeeld en bijbehorend kennisbeeld van waaruit we opnieuw kunnen nadenken over toetsing en evaluatie, geven we een voorbeeld van manieren waarop docenten op een impliciete manier te maken hebben met het objectiviteitsdenken. Docenten zijn overigens niet de enigen. Andere professionals die met mensen werken, zitten in hetzelfde schuitje.

Vanuit het verlichtingsdenken wordt aangenomen dat de werkelijkheid bestaat uit onveranderlijke structuren waarop de wetenschap grip kan krijgen. In het kielzog van deze aanname ontstaan hoge verwachtingen met betrekking tot goed onderwijs. We zien dit bijvoorbeeld als de staatssecretaris van onderwijs op zoek gaat naar een 'evidence based'-aanpak om pesten uit te bannen. Zijn goede bedoelingen zijn helder: pesten is een groot kwaad, dat te allen tijde tegengegaan moet worden. Zeer veel leraren zijn ook daarvan overtuigd. Hoe serieuzer leraren bezig zijn met pesten aanpakken, hoe meer ze waarschijnlijk geïnteresseerd zijn in de vraag welke mogelijkheden daarvoor bestaan en hoe hun praktijk daarvan kan profiteren. De staatssecretaris wil echter een stap verder gaan. Hij is op zoek naar manieren om pesten aan te pakken die vanuit Den Haag kunnen worden opgelegd, waarbij succes verzekerd is omdat wetenschappelijk bewijs bestaat voor hun werkzaamheid. Daarmee neemt hij aan dat pesten op dezelfde manier uitgeroeid kan worden als het gebruik van cfk's (die de ozonlaag aantasten) als drijfgas in spuitbussen. Een model dat van toepassing is op de levenloze natuur, wordt toegepast op de levende natuur. De staatssecretaris zoekt naar een technische oplossing vanuit Den Haag voor relationele problemen in de klas.

Als gevolg van een mechanistisch werkelijkheidsbeeld wordt het paard achter de wagen gespannen. Terwijl de menselijke betrokkenheid en verantwoordelijkheid van pester, de gepeste en hun opvoeders essentieel zijn om het probleem tegen te gaan, wordt gesuggereerd dat er een aanpak bestaat waarbij succes verzekerd is, louter doordat de professional standaardmaatregelen uitvoert. Dit werkt mythevorming in de hand, in dit geval de mythe dat leraren er via een paar trucjes voor kunnen zorgen dat leerlingen elkaar niet meer pesten. De aanname van het mechanistisch werkelijkheidsbeeld dat succes ook in sociale settingen te verzekeren is, maakt van docenten (maar ook van professionals in andere sectoren) machtige manipulators. Zo kan het werkelijke resultaat alleen nog maar tegenvallen. In plaats van in succes, resulteert deze benadering in wantrouwen en druk van omstanders: jullie docenten doen je werk niet goed! Door de toepassing van een werkelijkheidsbeeld dat

niet helemaal klopt, ontstaat een situatie waarin professionals verantwoordelijk gesteld worden voor een keur aan maatschappelijke problemen waar goed beschouwd maar weinigen een antwoord op hebben.

Rondom mbo- en hbo-professionals is in de loop der tijd een schil van academisch gevormde deskundigen ontstaan die leiding geeft aan professionals op de werkvloer en die managers en beleidsmakers adviseert. Vanuit het objectiviteitsdenken wordt verondersteld dat hun academische kennis van grotere waarde is dan de praktische ervaringskennis van de professional. Op basis van formeel verzamelde gegevens, brengen zij advies uit over beleidszaken of over hoe professionals in de praktijk (onder wie overigens ook academici) hun werk moeten doen. Ook als zij de praktijk zelf niet kennen, hebben zij meer recht van spreken dan de professionals die de praktijk wel kennen en door ervaring veel geleerd hebben. Hun academische achtergrond geeft hen macht want er wordt verondersteld dat deze academische kennis een superieur soort kennis is. Zo verliezen leraren niet alleen hun stem maar ook hun beroepstrots (Jansen, 2009; Bulterman-Bos, 2010). Intussen is het werk van de uitvoerende professionals er niet minder complex op geworden. De houding van beleidsmakers of leidinggevendenden (maar ook die van de studenten of cliënten) dat er kasten vol professionaliteit bestaan die alleen maar geopend hoeven te worden, kan bij de uitvoerende professionals leiden tot gevoelens van machteloosheid, ontgoocheling, frustratie, schuld, woede en angst. Kelchtermans (1999) brengt de kwetsbaarheid van leraren onder andere in verband met wat hij noemt 'de grenzen van eigen doelmatigheid': hun persoonlijke impact heeft grenzen omdat leerlingen niet alleen door de leraar beïnvloed worden en de professionele kennis van leraren geen technische garantie geeft voor succes.

Mythes over maakbaarheid maken het leven van docenten en andere professionals die met mensen werken niet makkelijker. Het is dus de moeite waard om te zoeken naar een beeld van de werkelijkheid waarin het specifiek menselijke een nadrukkelijke plaats heeft.

Gespreksvraag: Herkent u iets van kwetsbaarheid? Hoe verklaart u het ontstaan van de hierboven genoemde gevoelens?

Een multidimensionaal wereldbeeld

Tegenover het platte wereldbeeld waarin het objectiviteitsdenken centraal staat, plaatsen we een alternatieve visie op de wereld en op kennis. Die is

afkomstig van een vooraanstaand natuurwetenschapper, Michael Polanyi, die in 1891 in Boedapest werd geboren. Polanyi was gepokt en gemazeld in het natuurwetenschappelijke denken (Gelwick, 1977). Toch ging zijn belangstelling steeds meer uit naar de achterliggende ideeën van wetenschap. Om zich heen ontwikkelden zich twee maatschappelijke systemen, namelijk het communisme en het nazisme. Beide systemen hadden veel op met objectieve wetenschap en hadden grootse idealen voor een betere wereld. Toch liep het in beide systemen grondig mis. Dit stemde Polanyi tot nadenken, met name over de vraag hoe mensen zelf betrokken waren in het wetenschappelijk denken. Was het werkelijk waar dat de wetenschappelijke kennis die mensen ontwikkelden van een zodanig hoger niveau was dat ze zelf buiten spel kwamen te staan? Konden ze echt zeggen: 'het zit zo in elkaar', 'het is nu eenmaal zo' alsof ze er zelf niets mee te maken hadden? Is de hoogste vorm van kennis die soort van kennis die losstaat van wie wij zijn en hoe wij in de wereld staan?

Polanyi (1958) liet zien dat wetenschappelijke ontdekkingen juist gepaard gingen met een hoge mate van betrokkenheid van wetenschappers. Ze waren heel intensief met hun object van onderzoek bezig. Daardoor wisten ze vragen te stellen die de moeite waard waren of zagen ze verbanden die anderen niet zagen. Zo kwam Polanyi tot de conclusie dat alle kennis een persoonlijk element heeft. Objectieve kennis in de zin van kennis die zo volmaakt is dat een mens met deze kennis als het ware op de troon van God kan gaan zitten, bestaat niet. Waarheid is een relationeel begrip in de zin dat wie iets wil begrijpen van de werkelijkheid, een relatie met die werkelijkheid moet aangaan. Waarheid is daarmee echter geen persoonlijk verzinsel en ook geen volledig relatief begrip. Het is de uitdaging en plicht om zo goed mogelijk naar de werkelijkheid te kijken, recht te doen aan allerlei aspecten, zorgvuldig waar te nemen, goed te luisteren, proberen aan te voelen, verstand van zaken te krijgen. Zo ontstaan inzichten die mensen met recht voor hun rekening kunnen nemen. De woorden 'met recht' en 'voor hun rekening nemen' hebben een morele lading. Begrip voor iets opbouwen, of het nu gaat om begrip voor hoe een lichaamscel werkt of voor hoe een student leert, heeft volgens Polanyi alles te maken met verantwoordelijkheid nemen.

Hoe kunnen docenten dichter bij de waarheid komen? Docenten kunnen elkaar daarbij helpen.

“Lastige student, die Ruben”, verzucht docent Gerard tegen zijn jongere collega Angela op een toon alsof zij het ook moeilijk met Ruben heeft. “Altijd overal tegenin gaan, altijd de discussie frustreren met nonsens, altijd overal op- en aanmerkingen over maken!”

Angela weet niet hoe ze moet reageren. Is Ruben een lastige student? Ze kan zich geen enkel moment herinneren waarop Ruben haar colleges dwarsboomde. Ruben is anders dan vele andere studenten. Toen ze tijdens het eerste college de namen van de studenten leerde kennen, had ze tussen neus en lippen door gevraagd of hij familie was van een bekende schrijfster. “Ja, dat is een tante van mij”, had hij geantwoord en daar parmantig aan toegevoegd: “Maar wij spreken niet met deze tak!” Ruben is een jongen uit een gecompliceerd intellectueel en artistiek milieu, had Angela gedacht, wat haar eerder aan hem bond dan van hem scheidde. In de wandelgangen ging het gesprek al gauw over de kringen waarin Ruben verkeerde maar zich niet op zijn plaats voelde. Ook op deze lerarenopleiding voelde Ruben zich niet thuis. Omdat hij al diverse opleidingen geprobeerd had, had hij zich voorgenomen deze opleiding hoe dan ook af te maken.

Even voelt Angela hoe moeilijk het is om met collega's over studenten te praten. Alles wat je zegt over studenten, zegt ook iets over jezelf. Collega Gerard vindt ze geen type voor milieus vol intellectuele, artistieke cynici. Ze waagt toch een poging en vertelt wat ze weet over Rubens achtergrond. Ze heeft nooit last van hem, integendeel. Gerard blijft bedremmeld staan. Wat begon als een gesprek over een student, is een gesprek over zichzelf en zijn relatie met die student geworden. Dat was niet zijn bedoeling. Toch voelt hij dat Angela probeert hem iets aan te reiken waardoor hij anders naar Ruben kan kijken. Hij besluit eerlijk te zijn. “Ik denk dat ik een eenzijdig beeld van Ruben heb”, zegt hij. “Bedankt voor de informatie. Ik ga ermee aan het werk.”

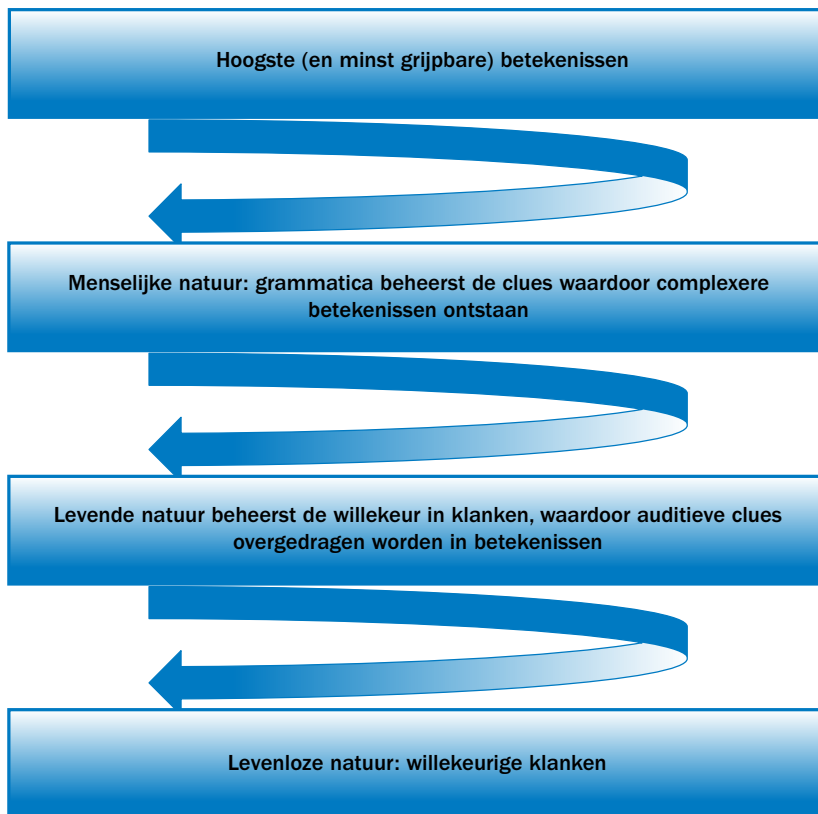
Polanyi's visie op de werkelijkheid is een aanklacht tegen reductionisme die het gevolg is van het objectiviteitsdenken. In het universum ziet Polanyi een hiërarchie van levenloze materie dat via opklimmende niveaus steeds meer menselijke betekenissen en menselijke verantwoordelijkheid noodzakelijk maakt. Ieder niveau komt op zijn eigen manier tot stand maar de operationele principes krijgen richting vanuit hoger gelegen niveaus. Om dit wereldbeeld te illustreren gebruikt Polanyi het beeld van de menselijke spraak (Gelwick, 1977; Grene 1969): Als je menselijke spraak op het laagste niveau bestudeert, worden klanken gemaakt die voldoen aan de natuurwetten. Op een iets hoger niveau wordt duidelijk dat het niet om willekeurige klanken gaat maar dat

slechts een bepaald soort klanken gebruikt wordt: dankzij de begrenzing van klanken klinken er woorden. Op een weer hoger niveau blijkt er tussen woorden een bepaalde relatie te bestaan. Dankzij de begrenzing van woorden via grammatica krijgen woorden in een zin betekenis. Op een weer hoger niveau krijgen zinnen richting vanuit hun onderlinge relatie. Dankzij die begrenzing kan er complexere betekenis ontstaan. Steeds is er sprake van hogere vormen van betekenis.

Polanyi benadrukt dat de niveaus niet tot elkaar gereduceerd kunnen worden. Ieder niveau heeft zijn eigen waarde en wetmatigheden en kan tot zijn recht komen dankzij de begrenzing vanuit een hoger niveau. Woordvorming zorgt ervoor dat klanken tot hun recht komen, zinsvorming zorgt ervoor dat woorden tot hun recht komen etc. Ieder niveau moet zowel vanuit de eigen wetmatigheden worden beoordeeld als vanuit het hoger gelegen niveau. Spelling en grammatica zijn bijvoorbeeld noodzakelijke ingrediënten van een goed opstel of van een literaire tekst. Maar een opstel of een literaire tekst beoordelen in termen van spelling of grammatica, terwijl de hogere betekenissen buiten beschouwing blijven, betekent een onacceptabele reductie. Bij de hogere niveaus gaat het om hogere niveaus van menselijke betrokkenheid. Daarbij kan het gaan om stijl maar ook om nog verder gaande betekenissen. Een boek kan stilistisch goed zijn en tegelijk op moreel niveau een gevaar voor de wereld vormen. Intelligentie is belangrijk om de wereld te leren kennen. Maar Polanyi benadrukt dat mensen met elkaar samenleven in sociale verbanden die uiteindelijk veel meer vragen dan intelligentie alleen. Ook intelligentie komt pas tot zijn recht als deze richting krijgt vanuit hogere niveaus.

De menselijke spraak is slechts een voorbeeld om het wereldbeeld uit te leggen. Elders gebruikt Polanyi het beeld van de stad die op het laagste niveau bestaat uit materie, op een hoger niveau uit gebouwen met bedoelingen en weer op een hoger niveau allerlei functies vervult die de menselijke samenleving mogelijk maken. Zo zijn de verschillende niveaus op vele terreinen in het leven te onderscheiden. Het is de kunst om recht te doen aan het juiste niveau en de daarbij behorende complexiteit te erkennen.

In het onderstaande schema hebben we een poging gedaan om de hiërarchie van Polanyi weer te geven. We duiden dit model aan met het begrip ‘multidimensionaal werkelijkheidsbeeld’.



Afbeelding 7.3. Het multidimensionale model naar aanleiding van Polanyi

Wat wordt concreet bedoeld met de uitspraak dat 'een lager niveau dankzij een hoger niveau tot zijn recht komt?' Op een mbo dat voorbereidt op economische en administratieve beroepen praten docenten economie en bedrijfsadministratie over hun werk. Maakt de christelijke identiteit van de school iets uit voor de manier waarop zij die vakken geven? Een docent bedrijfsadministratie vindt van niet: bedrijfsadministratie gaat in alle bedrijven volgens dezelfde regels die niets met levensbeschouwelijke waarden te maken hebben. Maar een collega ziet dat anders. Hij heeft zojuist de film *Hanna Arendt* gezien waarin historische beelden worden getoond van het proces tegen oorlogsmisdadiger Adolf Eichmann. Met grote precisie voerde Eichmann de administratie van de Jodenvervolgung. In de historische beelden komt Eichmann niet over als een moordenaar maar als een goed vakman. Als administrator had hij zijn werk goed gedaan. Vragen op een

niveau hoger had hij echter laten liggen, zoals: 'Wat administreer ik?' 'Waar toe administreer ik?' Bij deze vragen verwees Eichmann naar zijn superieuren. Hij had alleen maar gedaan wat hem was opgedragen. Polanyi zou zeggen dat het niveau van Eichmanns vakmanschap geen richting kreeg vanuit een hoger niveau waarop het ging om Eichmanns eigen verantwoordelijkheid voor de wereld.

Een multidimensionaal wereldbeeld, onderwijs en toetsing

Waarom hebben we dit multidimensionale wereldbeeld nodig om de voortdurende slingerbeweging tussen de twee paradigma's te overstijgen? Het is duidelijk dat het materialistische wereldbeeld, waarbij natuurwetenschappelijke kennis bovenaan in de hiërarchie staat, goed uit de voeten kan met de overdracht van al bestaande natuurwetenschappelijke kennis en goede procedures voortbrengt waarmee deze kennis getoetst kan worden. Leggen we dit model echter naast de taxonomie van Bloom (zie hoofdstuk 3), dan is ook duidelijk dat het hier voornamelijk gaat over de lagere niveaus van onthouden, begrijpen en toepassen en misschien analyseren. Voor de hogere niveaus, namelijk evalueren en creëren, is dit model aanmerkelijk minder geschikt. Op die hogere niveaus wordt bijvoorbeeld een beroep gedaan bij open opdrachten als een opstel, of wanneer studenten zelf onderzoek doen voor profielwerkstukken of afstudeerscripties. Ook omgaan met veel situaties in de beroepspraktijk is in feite een open opdracht. Daarvoor zijn geen externe standaarden aanwezig met meetinstrumenten. Studenten doen onverwachte ontdekkingen die niet in de doelstellingen waren opgenomen. Het materialistisch-natuurwetenschappelijke model geeft dan het gevoel dat het hier gaat om 'lagere' inhouden waarbij 'slechts subjectieve' beoordeling mogelijk is. Dat geeft een ongemakkelijk gevoel. De open opdracht dreigt naar de tweede of derde rang te verdwijnen.

Het multidimensionale wereldbeeld geeft aan dat er niveaus bestaan die met recht op een natuurwetenschappelijke manier bestudeerd worden, waarbij metingen een belangrijke rol spelen. Op hogere niveaus gaat het niet meer om zaken die meetbaar zijn met anonieme instrumenten. Het komt aan op beoordelen vanuit bekwaamheid en kennis van zaken. Het woord 'oordeel' in het 'beoordelen' maakt de morele connotatie duidelijk. Daarbij kan gedacht worden aan morele waarden zoals verantwoordelijkheid maar ook aan waarden die voortvloeien uit het vakmanschap. Het gaat daarbij om complexe processen waarbij soms recht gedaan moet worden aan tegenstrijdige doelstellingen.

Beoordelen op hogere niveaus kan niet plaatsvinden door anonieme computers of door willekeurige mensen. Wel door mensen die bepaalde vaardigheden, bepaalde innerlijke kwaliteiten ontwikkeld hebben. Van de beoordelaar wordt dus vakmanschap vereist.

In het multidimensionale wereldbeeld past een nieuwe manier om over onderwijs en toetsing na te denken. Eén die zich laat voeden door wetenschap én door praktische wijsheid. Eén die zich bezighoudt met het algemene én het unieke; die het belang van persoonlijke ontwikkeling erkent naast het belang van intellectuele en academische vorming, die erkent dat professionals, doordat ze de hele dag door met studenten omgaan, een gevoeligheid ontwikkelen die essentieel is om van hun beroepspraktijk als docent iets te begrijpen en om als beoordelaars op te treden bij de stof die ze onderwijzen (Bulterman-Bos, 2008a en b). Kortom, één die de waarde van logisch denken en systematiek erkent en tegelijk erkent dat mensen niet in systemen te vangen zijn. Deze manier van denken is niet alleen nodig voor de onderwijspraktijk maar ook in andere beroepspraktijken waarin omgaan met mensen centraal staat.

Slot

We hebben gezien dat de slingerbeweging, die het onderwijs dan weer eens de ene kant uitdrijft en dan weer eens de andere, samenhangt met een poging om het onderwijs volgens logische maatstaven in een systeem te vatten. Deze poging kan niet los gezien worden van het verlichtingsdenken, dat veel ruimte biedt voor logica maar geen raad weet met problemen waarvoor de oplossing meer vraagt dan logica alleen. Het multidimensionale wereldbeeld hebben we gepresenteerd als poging om de slingerbeweging te overstijgen. Het volgende hoofdstuk is een praktische uitwerking. Helpt het multidimensionale wereldbeeld om in toetsplannen aan het onmeetbare een volwaardige plaats te geven?

Hoofdstuk 8

Het onmeetbare in het toetsplan

Polanyi was geen onderwijskundige en heeft dus geen concrete handreikingen gedaan voor onderwijs en toetsing. In dit hoofdstuk maken wij gebruik van het multidimensionale wereldbeeld om opnieuw na te denken over evaluatie en toetsing. Daarbij gaan we in op het begrip ‘connaisseur’ en vragen we ons af of onmeetbare kwaliteiten een plaats kunnen krijgen in het toetsplan. In ieder geval geven we praktische voorbeelden van docenten die dit gedaan hebben.

Lagere niveaus als volwaardig niveaus

Op lagere niveaus in het multidimensionale model van Polanyi bestaat een zekere eenduidigheid waardoor meting met anonieme instrumenten op zijn plaats is. In het vorige hoofdstuk concludeerden we dat objectiviteit niet bestaat maar op de lagere niveaus lijkt die wel te bestaan. Dat komt doordat mensen hier gemakkelijk procedures kunnen ontwerpen waardoor ze tot een gezamenlijke visie op de werkelijkheid kunnen komen. Er bestaan bijvoorbeeld duidelijke afspraken voor wat een driehoek is, wat een centimeter en een graad betekent. Er zijn dan ook eenduidige antwoorden mogelijk, terwijl recht gedaan wordt aan de aard van een niveau. Vaak gaat het hier om technische situaties.

Technische kennis is belangrijk in het onderwijs. Heel veel vakgebieden kennen zo’n technische kant die berust op gegevens die iedereen voor waar aanneemt. Taal heeft een technische kant: leerlingen moeten letters kennen en van letters woorden kunnen vormen. Ze moeten spellingsregels kennen en iets begrijpen van grammatica. Zo hebben veel vakgebieden technische aspecten. Hoe belangrijk het ook is om kinderen zelf te leren nadenken en hun creativiteit te bevorderen; het heeft weinig zin om kinderen nieuwe spellingsregels te laten uitvinden. Het technische niveau is een volwaardig niveau met eigen wetten.

Vanuit het instrumentele paradigma worden middelen aangereikt om recht te doen aan technische vaardigheden. De hiërarchie maakt echter ook duidelijk dat de technische aspecten pas tot hun recht kunnen komen als ze richting krijgen vanuit een hoger niveau. Vertaald naar het onderwijs betekent dit dat het technische niveau voor studenten betekenis moet krijgen. Dat is een belangrijk principe in de didactiek. Het technische niveau heeft dus zijn eigen waarde maar komt pas tot zijn recht als het richting krijgt vanuit een hoger niveau van betekenis.

Hogere niveaus als volwaardige niveaus

Leren op hogere niveaus veronderstelt een meer persoonlijke invulling door de student dan leren op de lagere niveaus. Op hogere niveaus leren is niet voorbehouden aan het hoger onderwijs maar komt op alle terreinen binnen onderwijs voor. Een leerling die een opstel schrijft, geeft een persoonlijke invulling aan een thema, bedenkt zelf zinnen, een opbouw en een plot. Een kleuter die een rollenspel speelt, formuleert zelf zinnen, bouwt situaties op, gebruikt zijn verbeelding en drukt zichzelf uit. Deze kleuter creëert en komt daarmee op het hoogste niveau van de taxonomie van Bloom. Hoe hoger we komen in de hiërarchie, des te meer geldt dat vraagstukken van meerdere kanten bekeken kunnen worden en op verschillende manieren op een goede manier behandeld kunnen worden. Hogere niveaus zijn meerduidiger.

Op bepaalde hogere niveaus wordt de ziel verondersteld. Martha Nussbaum (2011, p. 21) omschrijft het begrip ziel als het denkvermogen en de verbeeldingskracht die beide ons tot mensen maken. Ze maken onze relatie met andere mensen tot rijke menselijke relaties, in plaats van relaties waarbij we elkaar alleen maar gebruiken en manipuleren. Binnen het multidimensionale wereldbeeld zou je verschillende vormen van menselijkheid kunnen onderscheiden. Er zijn mensen zoals Hitler en Stalin die hun macht gebruikten om tegenstanders uit te roeien. Er zijn ook mensen zoals Nelson Mandela en Corrie ten Boom, die na hun gevangenschap hun onderdrukkers met vergeving en verzoening tegemoet traden en zo iets betekenden voor de hele mensheid.

Een bekende uitspraak van Albert Einstein luidt: “Imagination is more important than knowledge.” Hogere niveaus hebben te maken met creativiteit, met kritisch denken. Hogere niveaus kunnen ook betrekking hebben op persoonlijke ontwikkeling of wezenlijke levensvragen. Waar het op lagere niveaus nog mag lijken dat objectiviteit bestaat omdat we het gemakkelijk eens

worden over hoe we naar de werkelijkheid kijken, zijn de hogere niveaus met waarden geladen.

Het vormingsparadigma vraagt aandacht voor deze hogere niveaus. We hebben in eerdere hoofdstukken gezien dat er weinig discussie was over de waarde van deze idealen. Het probleem was dat het zo moeilijk was om er in het onderwijs grip op te krijgen. Het is moeilijk te meten of de idealen gerealiseerd worden. Daardoor ontstond de vraag wat er van die mooie woorden terecht kwam.

De uitdaging is nu om vanuit het multidimensionale wereldbeeld te proberen om grip te krijgen op de hogere processen. De brede benadering die recht wil doen aan de vaardigheden die studenten nodig hebben om te participeren in de maatschappij, maakt het noodzakelijk om andere manieren van evaluatie toe te passen dan mogelijk zijn bij technische onderwerpen. Juist omdat opdrachten door studenten op een eigen manier kunnen worden ingevuld, is het onmogelijk om van tevoren volledig te omschrijven welk antwoord juist is. Polanyi zegt dat het op hogere niveaus meer gaat over beoordeling vanuit bekwaamheden en kennis van zaken. Hij gebruikt hier het begrip ‘connaisseur’. Vanuit deze begrippen proberen we iets te zeggen over het evalueren van onderwijsprestaties op hogere niveaus.

Gespreksvraag: Heeft uw vak technische aspecten? Welke zijn dit? Vanuit welke hogere niveaus krijgen deze technische aspecten richting?

Verschillen tussen lagere en hogere niveaus

- Lagere niveaus kunnen gemeten worden. Dit is een vorm van procedurele objectiviteit. Op hogere niveaus geldt de uitspraak van Einstein: “Not everything that counts, can be counted.”
- Op lagere niveaus geldt alleen de wet van de logica: de tegenstelling van waar is onwaar. Op hogere niveaus wordt de logica niet opgeheven maar wel overstegen. De tegenstelling van een diepere waarheid kan een andere diepe waarheid zijn.
- Op lagere niveaus is het bereiken van overeenstemming tussen mensen vanuit verschillende levensbeschouwingen en culturen gemakkelijk. Op een hoger niveau is dat moeilijker en toch kunnen zij vanuit hun gezamenlijke verantwoordelijkheid de dialoog aangaan en tot gezamenlijke conclusies komen. Culturen en levensbeschouwingen zijn in beweging.

- Tegelijkertijd zijn er zaken waar geen gezamenlijke conclusies getrokken kunnen worden en gaan verschillende groepen verschillende wegen.
- Van lagere zaken kun je bewijzen dat ze bestaan, maar bij hogere kan dat niet meer. Bestaat de liefde? Een liefdevol mens zal deze vraag instemmend beantwoorden maar een liefdeloos mens niet. Op een hoger niveau zeggen de antwoorden die mensen geven steeds meer over wie ze zelf zijn. Op een zeker niveau is de vraag niet meer 'hoe zit iets in elkaar?', maar 'hoe sta ik hier in?'.

Juf Ursula geeft rekenen in groep 8 waarbij ze zoveel mogelijk niveaus van het echte leven aan de orde wil laten komen. Ze doet dit aan de hand van het volgende probleem: Een kind heeft een schuurtje nodig waar haar paard in de winter kan schuilen. De tekening ligt al klaar. Moet het schuurtje geschilderd of gebeitst worden of kan het gewoon van onbehandeld hout blijven? Ursula kondigt aan dat de klas een aantal lessen zo levensecht mogelijk met dit probleem aan de slag gaat.

Juf Ursula onderscheidt allereerst het technische niveau: de leerlingen moeten in staat zijn te berekenen wat het kost aan materialen om het hout al dan niet te bewerken. Ook moeten ze in staat zijn om informatie te verzamelen over de vraag wat de gevolgen zijn voor het al dan niet bewerken van het hout voor de duurzaamheid van het schuurtje.

In de eerste les zijn de leerlingen voornamelijk met deze vragen bezig. Ze zetten de feiten op een rijtje. Wat valt er te zeggen over de bewerking van hout in relatie tot de duurzaamheid van het schuurtje? Hoeveel kosten die verschillende bewerkingen? Ursula kan het werk van leerlingen gemakkelijk beoordelen.

In de tweede les komt een hoger niveau in de hiërarchie aan de orde. Het gaat om de keuze die gemaakt wordt. De leerlingen gaan in groepjes in gesprek. Om beurten noemen ze eerst de voor- en nadelen van drie opties. Het gaat om een afweging van factoren: de duurzaamheid van het schuurtje, de belasting van de bewerking voor het milieu, de kosten van de bewerking en wat de leerlingen mooi vinden. Daarna neemt iedere leerling voor zichzelf een beslissing. Individueel schrijven ze op voor welke bewerking ze kiezen en waarom. Uiteraard zijn op dit niveau verschillende interpretaties mogelijk.

Op het volgende niveau gaat het om de vraag hoe leerlingen deze keuzes beoordelen in het licht van hogere waarden. In de volgende les presenteert juf Ursula alle keuzen die leerlingen gemaakt hebben met bijbehorende argumenten en stelt de vraag: Als je weet wie deze keuze gemaakt heeft, weet je dan ook nog iets over wie deze leerling is en wat hij of zij belangrijk vindt? De leerlingen

praten daarover in groepen en concluderen dat sommige leerlingen het milieu centraal stellen en anderen de kosten.

Vervolgens tilt juf Ursula de discussie op nog weer een hoger niveau: Hoe moeten we omgaan met deze verschillende standpunten? Leerlingen krijgen hun eigen blaadje met argumenten weer terug. In tweetallen stellen ze elkaar verduidelijkende vragen. Vervolgens is het de vraag of ze zich iets kunnen voorstellen bij elkaars afweging. Ten slotte worden ze het met elkaar eens of besluiten ze dat ze het niet met elkaar eens zijn.

Zo gaat de rekenles niet alleen meer over technische vaardigheden maar over democratische vaardigheden, levensbeschouwelijke vorming en esthetische en ethische vragen.

Hoe gaat juf Ursula om met toetsing als het technische niveau gepasseerd is? Ze observeert en registreert. Een aantal leerlingen werkt prima mee zo lang het om technische vragen gaat. Zodra het gaat om eigen interpretatie ziet ze dat ze quasi-ongeïnteresseerd achterover gaan hangen. Ze loopt langs en vraagt de leerlingen persoonlijk: "Best wel een lastige vraag hè? Wat vind jij er moeilijk aan?" Sommige leerlingen verdwijnen in hun branie, anderen vinden het fijn dat de juf hen begrijpt. Toch is het duidelijk dat niet iedereen dol is op dit soort lessen. Temeer omdat de leerlingen het niet gewend zijn om op deze manier te werken. Kort noteert de juf iets over het gedrag van sommige leerlingen.

Na afloop van de les legt ze het uit. Deze benadering is nieuw, maar gaat wel over de échte vragen. Sommige leerlingen kunnen hier al prima mee uit de voeten; voor anderen is het nog wennen. Binnenkort krijgt iedereen weer een kans. De tweede keer dat er zo'n les is, let ze goed op de kinderen die de vorige keer moeite hadden met het gesprek. Iedere keer als ze iets goeds bij hen merkt, geeft ze een knipoog: je leert het wel!

Zo laat juf Ursula zien dat onderwijzen en formatieve evaluatie samen gaan. Uiteindelijk leren leerlingen in haar klas veel méér dan de Cito-toetsen laten zien.

Wat zij de kinderen meegeeft op het brede terrein van levenskunst is niet altijd meetbaar, maar wel merkbaar.

Connaisseurs, een geoefende blik en vakmanschap

Helpt het multidimensionale wereldbeeld om grip te krijgen op processen op hoger niveau? Polanyi zegt dat hier onder andere connaisseurs bij nodig zijn. Connaisseurs zijn gevoelig geworden voor een bepaald object omdat ze er veel mee zijn omgegaan. Op de illustratie (afbeelding 8.1.) beoordeelt een

kunstkenner een voorwerp. Hij werpt er een blik op en herkent een groot aantal details die een leek niet opvallen. De innerlijke kwaliteiten die dit herkennen mogelijk maken, zijn ontstaan doordat de kunstkenner participeert in een gemeenschap van mensen die zich veel met kunst bezighouden en daardoor een geoefende blik hebben ontwikkeld. Dankzij de geoefende blik is het oordeel van een connaisseur betrouwbaarder dan dat van een leek. Deze vorm van betrouwbaarheid ontstaat dus niet op grond van afstandelijke objectiviteit maar dankzij jarenlange betrokkenheid bij kunstobjecten waardoor kennis van zaken is opgebouwd. Beide zijn volgens Polanyi belangrijk. Niet alleen voor vakmanschap, maar ook voor wetenschap.



Afbeelding 8.1.: De Connaisseur van H.J. Scheeres (Nederlands schilder, 1892-1864).
Met dank aan Simonis en Buunk, Kunsthandel Ede.

Hoe hoger in de hiërarchie, des te meer is evaluatie geen kwestie meer van meten maar van beoordelen vanuit kennis van zaken en een geoefende blik. Voor een deel kunnen connaisseurs in woorden uitleggen waar hun oordeel op gebaseerd is. Ze verwijzen bijvoorbeeld naar stijlkenmerken van een bepaalde periode. Toch kijkt iemand met kennis van die stijlkenmerken zonder de geoefende blik niet op dezelfde manier als connaisseurs. De innerlijke kwaliteiten – de geoefende blik – kunnen niet door conceptuele kennis vervangen worden. De bekwaamheden van connaisseurs zijn dus niet

volledig in conceptuele criteria te vatten. Behalve conceptuele kennis hebben connaisseurs dus ook kennis die niet onder woorden te brengen is. Polanyi noemt dit *tacit knowledge*. Deze vorm van kennis is niet conceptueel, maar perceptueel (Kessels & Korthagen, 1996). Hij zit niet zozeer alleen in het hoofd van de kunstkenner maar in zijn of haar hele manier van kijken, van zijn. Tegelijk is de taal die connaisseurs onderling spreken essentieel om samen verder te komen. Maar de begrippen die ze gebruiken zijn aanduidingen. Het zijn hulpmiddelen die een goede verstaander nodig hebben. Het beoordelen van kunst berust dus zowel op conceptuele kennis als op innerlijke vaardigheden van de beoordelaar. Connaisseurs handelen op grond van conceptuele kennis én op gevoel.

Wat betekent dit verhaal over connaisseurs voor het beoordelen van onderwijsprestaties op hogere niveaus? Op deze niveaus is de vraag of iets goed of fout is in eerste instantie niet aan de orde. Vaak gaat het om de vraag of iets waardevol en waardevol genoeg is. Deze vraag kan niet vanuit een neutrale houding beantwoordt worden. Dit vraagt om docenten die het vak dat ze onderwijzen goed kennen en ook ervaring hebben met de manier waarop men met dat vakgebied omgaan. Het komt aan op docenten die zowel pedagogisch-didactisch als vakinhoudelijk connaisseurs zijn en vandaaruit feedback kunnen geven (formatieve evaluatie) en kunnen vaststellen of de student aan de eisen voldoet (summatieve evaluatie). Het oordeel dat uiteindelijk geveld wordt is niet 'objectief'. Ook niet 'subjectief'. Het komt voort uit het vakmanschap van de docent.

Désanne interesseert zich voor creatief schrijven. Ze verdiept zich in hoe een plot opgebouwd wordt, in manieren om een verhaal spannend te maken, in het gebruik van beelden waardoor de tekst tot leven komt. Dankzij deze kennis kan Désanne haar studenten enorm goed helpen om hun schrijfvaardigheden te ontwikkelen. Als Désanne opstellen beoordeelt, staat er nooit zomaar een cijfer onder. Ze geeft tips en complimenten en maakt leerlingen op een speelse manier bewust van het effect van hun taal op de lezer. Aan het eind van het jaar kijkt ze terug op een aantal oude opstellen van leerlingen. Hun ontwikkeling is niet meetbaar. Wel merkbaar.

Dit vakmanschap wordt aangeleerd door het vak te beoefenen. Een docent die voor het eerst een open opdracht (een werkstuk) beoordeelt, heeft waarschijnlijk nog een minder gedifferentieerd beeld van wat er van studenten op dit niveau verwacht behoort te worden. Daarom is het goed als deze docent

in een gemeenschap van collega's verkeert om zijn of haar eigen oordeel te ijken. Overleg met collega's over hoe een bepaalde casus gewaardeerd wordt, is nodig om die indrukken uit te wisselen en tot een gemeenschappelijk verstaan te komen (zie de casus aan het einde van hoofdstuk 4). Dit kader is onmogelijk helemaal op te schrijven, maar juist door het proces van samenwerking kan een onzichtbaar web van beoordelingsnormen en meta-criteria tussen de teamleden gevormd worden. Uiteraard kost dit overleg tijd en die moet ergens vandaan komen. Tegelijk is het goed te realiseren dat tijd die geïnvesteerd wordt in dergelijk overleg, zich ook op andere fronten terugbetaalt. Samen concrete casussen van werk van studenten beoordelen kan helpen om gezamenlijk vakmanschap te ontwikkelen. Belangrijk is zich te realiseren dat het beoordelen van werk van studenten een vaardigheid is die geleerd wordt door daar in een gemeenschap van vakgenoten actief en kritisch mee bezig te zijn (Bulterman-Bos, Terwel, Verloop & Wardekker, 2002).

De laatste jaren heeft het vier-ogen-principe ingang gevonden in het hbo. Dit betekent dat meerdere mensen naar hetzelfde werk kijken. Idealiter gaan ze daarbij enerzijds af op een formeel beoordelingskader maar hun vakmanschap en ervaring als docent wordt ook actief ingeschakeld. Doordat ze al heel veel werk van studenten gezien hebben en vertrouwd zijn met het beoordelingskader, kunnen ze ook omgaan met factoren die niet letterlijk in het beoordelingskader staan.

Gespreksvraag: Op welke manier krijgen vaardigheden op hoger niveau in uw onderwijspraktijk een plaats en hoe worden ze beoordeeld?

Externe validering

De dominantie van het objectiviteitsdenken zorgt er voor dat het onderwijs gereduceerd dreigt te worden tot technische procedures en instrumentele vaardigheden (Nussbaum, 2011). Het heeft bovendien beoordelingen door docenten vanuit vakmanschap naar de achtergrond verdrongen. Misschien waren daar docenten zoals in hoofdstuk 3, die wel macht hadden om studenten te beoordelen maar zich daar niet voor hoefden te verantwoorden, wel de aanleiding toe. Hoe dan ook, in de ontwikkeling van een referentiekader van waaruit studenten feedback krijgen, ligt een belangrijke kern van wat docenten tot goede docenten maakt. Idealen vanuit het vormingsparadigma kunnen beter tot hun recht komen wanneer docenten samen referentiekaders en een werkwijze ontwikkelen waarmee ze grip krijgen op de



ontwikkeling van studenten op hogere niveaus. Het gaat daarbij voortdurend om waarde-geladen zaken. Maar anders dan het objectiviteitsdenken veronderstelt, is dat eerder een kwaliteitskenmerk dan een probleem. De persoonlijke betrokkenheid van docenten en de normatieve professionalisering die daaruit voortvloeit, is juist een belangrijk ingrediënt van goed onderwijs. Daarbij is het van belang dat docenten elkaar scherp houden en met elkaar gezamenlijke kaders op meta-niveau ontwikkelen.

Eerder hebben we echter gezegd dat het formele beoordelingskader nooit helemaal zal voldoen omdat een goede beoordeling ook gebaseerd is op *tacit knowledge*. Tegelijk blijft het de moeite waard om beoordelingskaders zoveel mogelijk aan te blijven scherpen en de waarden waarom het gaat te verhelderen en daarover goed te communiceren met studenten.

Veel docenten zijn al gewend om het werk of de prestaties van hun leerlingen als connaisseurs te beoordelen. In het middelbaar beroepsonderwijs beoordelen docenten bijvoorbeeld het werk van studenten aan de hand van kwalificatiedossiers. Ook bij deze beoordeling moet de vraag naar de validiteit gesteld worden: wordt er werkelijk beoordeeld wat er beoordeeld moet worden? Zegt een dossier over de stage niet meer over de manier waarop studenten hun praktijkervaringen op schrift kunnen zetten dan over hun praktische vaardigheden? Op grond waarvan wordt een beoordeling uitgesproken? Zijn er grote verschillen tussen docenten? Als vormen van onderlinge afstemming achterwege blijven, is de kans aanwezig dat het onderwijs in vaagheden en subjectiviteit vervalt en er weer dezelfde kritiek klinkt als toen het vormingsparadigma in de mode was. Een reactie vanuit het instrumentele paradigma ligt dan weer op de loer.

In het advies van de hbo-raad *Vreemde ogen dwingen* worden externe partijen ingezet om het niveau van de opleiding te garanderen. Daarbij wordt ook gedacht aan mensen van buiten de instelling. Deze externe validering wordt onder andere aanbevolen bij open opdrachten, bijvoorbeeld bij de afstudeeronderzoeken waarin het heel erg aankomt op de eigen invulling van de student. Wanneer het hier gaat om de vreemde ogen van bureaucraten die louter beoordelen op grond van formele criteria, komt het objectiviteitsdenken weer terug. Dat kan ten koste van de inhoudelijke diepgang gaan. Wanneer die vreemde ogen echter de ogen zijn van andere docenten die ook hun vak serieus nemen, kan externe validering juist een kans zijn om gezamenlijke vakkennis te ontwikkelen en deze te verdiepen.

Een lerarenopleiding is bezig de beoordeling van afstudeeronderzoeken te verbeteren. Naast de groep docenten die onderzoek begeleiden, wordt een groep van externe werkvelddeskundigen gevormd. De groep bestaat uit directeuren, coördinatoren, interne begeleiders of leraren die het studentonderzoek vanuit het werkveld begeleiden. Omdat het gaat om het beoordelen van onderzoek, wordt vastgesteld dat zowel de interne als de externe beoordelaars over een master moeten beschikken. Ieder onderzoek wordt door een interne en een externe beoordelaar beoordeeld aan de hand van een beoordelingskader. Jaarlijks organiseert de opleiding een bijeenkomst met alle werkveldbeoordelaars om te praten over de ervaringen bij het beoordelen. Vaak wordt het beoordelingskader na deze bijeenkomsten bijgesteld. Zonder uitzondering geven de werkvelddeskundigen aan dat dit werk leerzaam is. Uiteindelijk gaat de discussie altijd over de vraag hoe je studenten op een zinnige manier leert nadenken over onderwijs. Het is fijn om gezamenlijk met deze vraag bezig te zijn.

Vakmanschap in ontwikkeling: het voorbeeld van het academisch ziekenhuis

Het beoordelen van hogere niveaus vraagt om conceptuele kennis én gevoel. Dit gevoel wordt door ervaring gescherpt. Dit is een reden waarom het in de medische wetenschap heel gewoon is dat artsen, alvorens zij zich verder kunnen gaan specialiseren, een voorgeschreven aantal casussen van patienten met een bepaalde ziekte gezien moeten hebben. Vervolgens is het belangrijk om taal te vinden voor die ervaringskennis (Eisner, 1976). Hoe meer complexe, concrete situaties iemand gezien heeft, hoe beter iemand in staat is om onder woorden te brengen aan welke criteria goed werk beantwoordt. Zo is het ontstaan van een goed beoordelingskader een voortdurend proces van ontwikkeling. Een docententeam kan bijvoorbeeld blijven groeien in het beoordelen van opstellen. Nieuwe opstellen zijn steeds weer interessante casussen waarmee het vakmanschap verder ontwikkeld kan worden terwijl er waarschijnlijk steeds meer consensus ontstaat.

Complexer is het beoordelen van het werk van studenten als dat werk buiten de opleiding plaatsvindt. Dat is het geval in het beroepsonderwijs. Lang niet alle docenten in het beroepsonderwijs beoefenen het beroep waarvoor ze opleiden. Ze zijn ingehuurd vanwege hun academische kennis. Een docent onderwijskunde die lesgeeft aan een lerarenopleiding is niet automatisch een goede juf of meester op de basisschool. Een docent sociologie die lesgeeft aan een opleiding voor hulpverlening is niet automatisch een goede hulpverlener.

Over het algemeen gesproken is de kans groter om in de praktijk mensen aan te treffen die het vak beheersen terwijl ze zich minder bezighouden met de theorie. Aan de opleiding zijn juist vaak mensen verbonden die theoretisch goed onderlegd zijn maar zich minder bezighouden met de praktijk en zeker de ontwikkelingen van de laatste tijd niet van binnenuit kennen.

Deze situatie vormt een spiegel van het oude objectiviteitsdenken waarin theorie- en praktijkontwikkeling gescheiden terreinen zijn. Dat kan spanningsvelden opleveren bij het beoordelen van het werk van studenten en bij de ontwikkeling en de aanscherping van beoordelingskaders. De kans is aanwezig dat uiteindelijk de manier waarop de studenten over hun handelen in de praktijk rapporteren belangrijker wordt dan hun handelen zelf. Dit kan 'opgelost' worden door practici een belangrijke stem in het kapittel te geven. Voor het ontwikkelen van genuanceerde beoordelingskaders is het nodig om de juiste taal te vinden. Dan is het jammer als juist de mensen die goed zijn in conceptualiseren (dus theoretici) weinig praktijksituaties meemaken. Zo ontbreekt de basis om preciezere formuleringen te vinden voor wat het betekent om goed werk te leveren. Een probleem ontstaat ook als op de opleiding uitgegaan wordt vanuit theorieën waar in de praktijk niet mee gewerkt wordt.

In de medische wereld zijn theorie en praktijk beter geïntegreerd. Opleiders in academische ziekenhuizen houden zich bezig met theorie én praktijk. Studenten worden beoordeeld in een meester-gezel relatie. Dit meester-gezel model (niet voor niets een model van voor de verlichting) is nog steeds actueel. Het voorbeeld van het academisch ziekenhuis kan het beroepsonderwijs inspireren tot vernieuwende samenwerkingsverbanden met de beroepspraktijk.

Stappen bij het multidimensionale model

Om het verschil aan te geven tussen het objectivistische denken binnen het platte wereldbeeld en de noodzaak van betrokkenheid binnen het multidimensionale wereldbeeld, geven we eerst de vragen weer die docenten zich stellen als zij werken volgens het objectivistische model. Dat sluit overigens nauw aan bij het instrumentele paradigma.

- Hoe zit de wereld in elkaar?
- Wat moeten de studenten weten?
- Hoe meet ik of mijn studenten de stof kennen?

Werken docenten vanuit het multidimensionale model dan stellen ze zich de volgende vragen:

- Hoe past mijn vak in de verschillende lagen van het model (zie hoofdstuk 7, afbeelding 7.3):
- Waar heeft mijn vak technische aspecten? Met welke instrumenten kan ik meten of mijn studenten deze technische aspecten beheersen?
- Dankzij welke hogere vragen komen die technische aspecten tot hun recht? Daarbij kan gedacht worden aan:
 - *Kwalitatieve vragen*: toepassingsvragen, waarbij het gaat om goed of beter vakmanschap en verschillende interpretaties
 - *Esthetisch/expressieve aspecten*: vragen van mooi of mooier, kracht waarmee ideeën of gevoelens tot uitdrukking komen
 - *Burgerschapsvragen*: manier van omgaan met verschillende interpretaties
 - *Ethische vragen*: verantwoording van verschillende interpretaties in het licht van hogere waarden
- Vanuit welke beroepswaarden en waarden over persoonsvorming ga ik de studenten beoordelen? Kan ik die expliciet maken?
- Hoe komen wij op onze opleiding tot gedeelde waarden en hoe helpen docenten elkaar om een gezamenlijk referentiekader te ontwikkelen en te verfijnen?

Gespreksvraag: Zijn de bovenstaande vragen bij het multidimensionale model toepasbaar op uw vak?

Gebruik van ict en het multidimensionale model

Dit boek is geschreven in het kader van het Surf-project 'Voortgangstoetsing in de doorstroom mbo-hbo.' Docenten Nederlands, rekenwiskunde en Engels hebben een leerlijn ontwikkeld van mbo naar hbo met daarbij digitale toetsen. Digitale toetsen hebben hier geen summatieve, maar een formatieve functie. Ze zijn bedoeld om studenten van feedback te voorzien zodat studenten zelfstandig kunnen oefenen. Hoe past het in het multidimensionale model? Enerzijds is het duidelijk dat het hier voor een groot deel gaat om technische vaardigheden waar toetsing met digitale instrumenten kan helpen. Anderzijds spelen andere vragen. In een cartoon onder de titel 'De ideale toekomst' zit een kind achter de computer. Moeder roept: "Hee, je moet al lang naar bed!" "Ja, ja, nog één sommetje", zegt het kind. Docenten weten dat deze ideale toekomst nog niet is aangebroken. Lang niet alle studenten zijn gemotiveerd om



zelfstandig aan het werk te gaan met moeilijke sommen, ook niet als dat via de computer gaat. Toch heeft het project wel een ideaal van zelfstandigheid dat aansluit bij het vormingsparadigma. Docenten staan dan voor een dilemma. Enerzijds willen ze studenten zoveel mogelijk zelf laten doen, anderzijds moeten ze sturen zodat studenten inderdaad aan het werk gaan. Het digitale programma biedt prachtige mogelijkheden voor leergierige studenten die de stof redelijk snappen. De studenten die de stof moeilijk vinden, hebben echter vaak niet genoeg aan de digitale feedback. Ze hebben een docent nodig die zich in hun manier van denken verplaatst. Zij hebben behoefte aan levens-echte formatieve feedback.

In de praktijk van het onderwijs ervaren docenten vaak dat het veel vraagt van studenten om het technische niveau onder de knie te krijgen waardoor de aandacht voor hogere niveaus in de verdrukking kan komen. Dat technische kennis betekenis krijgt vanuit hogere niveaus is belangrijk. Goede docenten vragen zich dus altijd af hoe kennis betekenis kan krijgen voor studenten en daarmee toegankelijk kan worden. Dit neemt niet weg dat er ook een reële vraag kan ontstaan hoeveel tijd daarvoor is. Vanuit projecten als 'Flipping the Classroom' (Berrett, 2012) zijn goede ervaringen opgedaan met een werkwijze waarbij ict een belangrijke rol speelde bij de instructie en oefening van technische aspecten van het vak. Ict was hierbij ingebed in de sociale structuur van het gezamenlijke onderwijsproces. Daardoor werd docenten een bepaalde hoeveelheid werk uit handen genomen zodat er meer tijd ontstond voor levensechte formatieve feedback. Soms kreeg dit de vorm waarbij docenten hun studenten voorafgaand aan colleges een digitale toets lieten maken. Zo verkregen de docenten informatie over de problemen die studenten hadden, waar zij hun college op konden afstemmen. Andere docenten toetsten vooraf of studenten hun stof daadwerkelijk gelezen hadden. De betere voorbereiding van studenten kwam de diepte van de discussie ten goede waardoor vaardigheden van een hogere orde meer tot hun recht kwamen. Het blijft de moeite waard om te zoeken naar mogelijkheden waarop intelligent gebruik van ict docenten werk kan besparen zodat er meer ruimte komt voor formatieve feedback op vaardigheden van een hogere orde. Ict wordt dan niet ingezet als middel tegen alle kwalen. Ook voor ict geldt dan dat het wordt ingezet op een niveau dat gecontroleerd wordt door een hoger niveau.

Het toetsplan en moeilijk meetbare vaardigheden

Het maken van een toetsplan hadden we aanvankelijk onder het instrumentele paradigma geschaard en beschouwd als onderdeel van een visie

op onderwijs waarbij de uitkomsten van tevoren vast staan. De kritiek van Dagmar in hoofdstuk 1 was terecht: niet alles van waarde is meetbaar. We hebben nu echter een werkelijkheidsbeeld gepresenteerd waarin zowel technische als de creatieve en menselijke aspecten een plaats hebben. Aan technische aspecten kan vanuit het instrumentele paradigma recht gedaan worden. Het vormingsparadigma doet recht aan de creatieve en menselijke aspecten.

Dat betekent dat ook het toetsplan ruimte biedt voor vaardigheden van hogere orde: voor eigen inbreng van de student, creativiteit en originaliteit, levensrecht en betekenisvol leren. Het toetsplan – maar dat hoort bij deze interpretatie eigenlijk evaluatieplan genoemd te worden – beschrijft de manier waarop vastgesteld wordt of studenten aan de eisen voldoen en gebruikt daarbij een scala aan middelen. Het verheldert ook de waarden van waaruit studenten beoordeeld worden en beschrijft hoe docenten die waarden delen en elkaar scherp houden. De criteria die gehanteerd worden, zullen rationeel gesproken nooit honderd procent waterdicht zijn. Daar staat tegenover dat het vakmen- sen zijn die met de criteria omgaan. Een evaluatieplan kan helpen om richting te geven aan het vormingsproces. Docenten kunnen elkaar versterken door samen dezelfde waarden centraal te stellen en zo voorkomen dat individuele leraren op een eiland met andersoortige leerprocessen worstelen.

Docent Walter wil dat zijn pabostudenten nieuwsgierig en onderzoekend te werk gaan met mathematische systemen. Hoe krijgt deze moeilijk meetbare doelstelling een plaats in het toetsplan dat Walter voor zijn domeinteam rekenwiskunde heeft geschreven?

Er is gekozen voor een grote variëteit van toetsvormen. De instrumentele kant van zijn vak houdt in dat de studenten functioneel rekenvaardig zijn. Deze instrumentele vaardigheden komen aan de orde via een digitale toets met gesloten vragen. Als studenten zelf kunnen rekenen, wil dat echter nog niet zeggen dat ze onderzoekend te werk gaan bij rekenen. Walter heeft te maken met studenten die meer van kinderen houden dan van rekenen. Daar speelt hij op in om nieuwsgierigheid en onderzoekend leren te bevorderen. Hij confronteert studenten met het denkwerk van kinderen. Op papier of via een filmpje laat hij zien hoe een leraar eerst de manier waarop kinderen denken onderzoekt en pas dan besluit wat het kind nodig heeft. Op het tentamen krijgen ze een casus van een leerling voor wie ze een aanpak moeten bedenken door diens denkwijze te achterhalen. Zo neemt Walter een deel van onderzoekend leren op in zijn toetsplan.

Verder maken studenten gedurende hun opleiding een lessenserie rekenwiskunde met een verantwoording daarbij. Ook dat bestrijkt het terrein van onderzoekend leren. Bij deze materialen wordt door de docenten formatieve feedback gegeven. De summatieve evaluatie houdt in dat studenten aangeven wat ze met de formatieve feedback hebben gedaan.

De rest van dit moeilijk meetbare ideaal wordt getoetst in de stage en daar beoordeeld. Walter heeft daarover regelmatig overleg met de interne coaches of mentoren op de stagescholen.

Studenten laten zich sterk sturen door wat geëvalueerd wordt. Omdat de toetsen goed aansluiten bij de brede vaardigheden die studenten uiteindelijk moeten beheersen, kan dit 'studeren voor de toets' geen kwaad. Er is immers op gelet dat ook de hogere orde vaardigheden (volgens de taxonomie van Bloom) voldoende aan bod komen. En er is gekozen voor evaluatievormen die recht doen aan de aard van wat er geleerd moet worden.

Wat zijn uw vormingsidealen en komen deze in uw evaluatieplan naar voren?

Slot

We hebben in de loop van dit boek gezien dat het beroep van docent complexer is dan vele mensen denken (zie ook Labaree, 2000). Goede docenten zijn normatieve professionals die in unieke situaties steeds weer eigen oplossingen moeten vinden. Hun persoonlijke betrokkenheid en hun vakmanschap is ook nodig bij het beoordelen van het werk van studenten. Waar het objectiviteitsdenken de docent als normatieve professional naar de achtergrond dringt, vraagt het beoordelen van complexe vaardigheden om vakmanschap en kennis van zaken. Dit veronderstelt een gemeenschap waarin docenten gezamenlijk met hun vak bezig zijn en hun inschattingen en waarden met betrekking tot de beoordeling op elkaar af kunnen stemmen. Tijd en aandacht voor samenwerking is dus noodzakelijk. Vormingsidealen kunnen dan opnieuw een plaats krijgen.

Hoofdstuk 9

Doel en middelen

Het onderwijs zal gericht zijn op de volle ontwikkeling van de menselijke persoonlijkheid en op de versterking van de eerbied voor de rechten van de mens en de fundamentele vrijheden. Het zal het begrip, de verdraagzaamheid en de vriendschap onder alle naties, rassen of godsdienstige groepen bevorderen (Universele verklaring van de rechten van de mens, 1948).

In dit boek hebben we ons beziggehouden met vorming, toetsing en evaluatie. Toetsen zijn middelen om het leerproces te bevorderen. De cijfers die uit toetsen voortkomen, vormen belangrijke maar gebrekkige indicaties van wat studenten geleerd hebben.

Toetsing kan niet losgemaakt worden van vorming. Slechts een gedeelte van een betekenisvol curriculum valt door middel van toetsen te evalueren. Een breed scala aan evaluatiemiddelen is nodig om recht te doen aan datgene wat studenten nodig hebben om te kunnen participeren in de maatschappij en de beroepspraktijk. Toetsen afnemen is één van die middelen.

Als resultaten van onderwijs gevat worden in getallen, kan bij managers of beleidsmakers de indruk ontstaan dat men grip heeft op de kwaliteit van onderwijs. De relatie tussen doel en middelen wordt dan nogal eens omgedraaid. Toetsen zijn dan geen middelen meer om het leerproces te bevorderen. Het leerproces wordt een middel om de toetsresultaten te bevorderen. Vaak wordt dit gerechtvaardigd vanuit economische motieven. Omdat we in een kennis-economie leven, hebben we goed onderwijs nodig. Wat goed onderwijs is, blijkt uit ranglijsten waarmee landen met elkaar concurreren.

Zo dreigt toetsing losgemaakt te worden van vorming en telt alleen nog maar wat gemeten wordt. Essentiële vaardigheden als kritisch en creatief denken



– uitgerekend die vaardigheden die de kenniseconomie zo nodig heeft – worden tweederangs. Het curriculum verenigt zich tot kennis die weinig meer met de praktijk van het leven te maken heeft. De volle ontwikkeling van de menselijke persoonlijkheid, waarover de Universele verklaring van de rechten van de mens spreekt, raakt uit beeld. Voor een aanzienlijk deel van de bevolking wordt onderwijs een producent van negatieve zelfbeelden.

In haar boek *Niet voor de winst* (2011) heeft Martha Nussbaum laten zien dat een brede ontwikkeling van alle talenten ten goede komt aan de sociale samenhang in de maatschappij. Kunst, literatuur, talen, muziek, filosofie, geschiedenis en cultuur zijn geen overbodige luxe. Ze zijn essentieel voor de vorming van elke burger. Een brede ontwikkeling is nodig om de democratie levend te houden. De poging om economische doelen centraal te stellen bijt dus – paradoxaal genoeg – in zijn eigen staart. Juist als we het onderwijs ondergeschikt maken aan de economie, kunnen we verwachten dat problemen ontstaan op het gebied van innovatief denken, sociale samenhang en participatie. Met alle gevolgen voor de economie van dien. Het belang van goed onderwijs is intrinsiek. Een economie kan pas functioneren als de eigen aard van onderwijs wordt erkend. Toetsen moeten middelen blijven om het leerproces te bevorderen.

Het onderwijsbeleid slingert tussen het instrumentele paradigma en het vormingsparadigma. Beide paradigma's met elkaar verenigen lukt maar zeer ten dele, omdat we gevangen zitten in een objectiviteitsdenken dat onderwijs wil vatten in systemen die van buitenaf bestuurbaar zijn. In dit boek hebben we een handreiking gedaan voor een alternatief waarbij het objectiviteitsdenken plaatsmaakt voor normatieve professionalisering en voor vakmanschap van de docent.

Veel professionals maken deel uit van systemen waarin ze hun studenten zien lijden onder de manier waarop toetsen gebruikt worden. Andere professionals, die met 'zachtere' middelen werken, merken dat ze te weinig grip hebben op het leerproces waardoor echte kwaliteit uitblijft. Standaardantwoorden zijn niet te geven. Normatieve professionalisering veronderstelt docenten die in hun eigen vakmanschap geloven en zelf het heft van onderwijsontwikkeling in handen nemen. Ze baseren zich zowel op onderzoek als op praktijkervaring, op verstand én gevoel. Ze nemen hun professionele ruimte in en baseren zich desnoods op het feit dat de vrijheid van onderwijs in de grondwet is verankerd.

Niet in de eerste plaats de deskundige of de beleidsmaker van buitenaf maar de docent zelf kan zien wat studenten nodig hebben. In dit boek hebben we geprobeerd om achtergrondinformatie en hulpmiddelen aan te reiken waarmee docenten in samenspraak met elkaar een visie op evaluatie kunnen vormen.

Bijlage

Uitslag vragenlijst onderwijsvisies

De onderzoekers Kaldeway, De Heer en Bulterman-Bos (2014) maken onderscheid tussen opbrengstgericht werken en talentgericht werken. Opbrengstgericht werken sluit aan bij het instrumentele paradigma, talentgericht werken bij het vormingsparadigma.

Geef bij elke vraag je score aan: bij het meest linkse bolletje hoort een score 1, vervolgens 2, 3 en 4; bij meest rechtse bolletjes score 5.		
Vul in de 2e kolom in:		
Score vraag 1		
6 minus score vraag 2		
Score vraag 3		
6 minus score vraag 4		
Score vraag 5		
6 minus score vraag 6		
6 minus score vraag 7		
Tel je scores op		
Deel het totaal door 7 (één cijfer achter de komma)		Score 1 = maximaal opbrengstgericht Score 5 = maximaal talentgericht Score 3 = precies gemiddelde score
Reken eventueel om in een percentage: Trek van je score 1 punt af en vermenigvuldig het resultaat met 25.		Het percentage geeft je mate van talentgerichtheid aan. 100 minus het percentage geeft je mate van opbrengstgerichtheid aan.

Literatuur

- Aalst, H.F., & Kok, J.J.M. (2004). Driehoeken van het Nieuwe Leren. *Jeugd in School en Wereld*, 89, 4, 11-15.
- AERA, APA & NCME (American Educational Research Association, American Psychological Association & National Council on Measurement in Education) (1999). *Standards for Educational and Psychological Testing*. Washington DC: American Educational Research Association, zie standaard 13.7.
- Bergh, H. Van den (2006). *Zeker weten door zuiver meten?* Oratie uitgesproken bij de aanvaarding van het ambt van bijzonder hoogleraar Didactiek en toetsing van het taalvaardigheidsonderwijs, vanwege de Vereniging van Leraren in Levende Talen. Universiteit Utrecht, faculteit geesteswetenschappen en letteren.
- Berrett, D. (2012). How 'flipping' the classroom can improve the traditional lecture. *The Chronicle of Higher Education*, Feb. 19, 2012. <http://cft.vanderbilt.edu/teaching-guides/teaching-activities/flipping-the-classroom/>.
- Biesta, G. (2007). Why "what works" won't work: evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57 (1), 1-22.
- Biesta, G. (2012). *Goed onderwijs en de cultuur van het meten. Ethiek, politiek en democratie*. Den Haag: Boom/Lemma.
- Boer, E. de, Burg, A. van der, Hoek, A., Kerkhoffs, J., Kuhlemeier, H., Os, M. van, Raanhuis, M., & Willemse, P. (2005). *Competentiegericht leren en beoordelen in vmbo en mbo Ontwikkelingen, knelpunten en oplossingsrichtingen tussen theorie en praktijk*. Utrecht: WVOI.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: the classification of educational goals; Handbook I: Cognitive Domain*. New York: Longmans, Green.
- Bruner, J. (1996). *The culture of education*. Cambridge, MA: Harvard University Press.
- Bulterman-Bos, J.A. (2007). "Bij de groep horen": hoe de imitatietendens het leraren moeilijk kan maken om adaptief onderwijs vorm te geven. *Pedagogische Studiën*, 84(3), 159-175

- Bulterman-Bos, J.A. (2008a). Will a clinical approach make educational research more relevant for practice? *Educational Researcher*, 37(7) 412-420.
- Bulterman-Bos, J.A. (2008b). Clinical study: A pursuit of responsibility as the basis of Education Research. *Educational Researcher*, 37(7) 439-445.
- Bulterman-Bos, J.A., Terwel, J., Verloop, N. & Wardekker, W.L. (2002). Observation in teaching: toward a practice of objectivity. *Teachers College Record* 104(6), 1069-1100.
- Bulterman-Bos, J., Verloop, N., Terwel J. & Wardekker, W. (2003). Reconciling the Pedagogical goal and the Measurement Goal of Evaluation: The Perspectives of Teachers in the Context of National Standards. *Teachers College Record*, 105/3, 344-374.
- Bulterman-Bos (2010). *Geef de onderwijsontwikkeling terug aan leraren*. Lectorale rede Christelijke Hogeschool Ede.
- Bulterman-Bos, J.A. (2012). Wetenschap en vakmanschap en de toekomst van de onderwijs/leerwetenschappen. *Pedagogische Studien*, 89, 174-184.
- Casteleijns, J, Segers M. & Struyven, K. (2011). *Evalueren om te leren. Toetsen en beoordelen op school*. Baarn: Coutinho.
- Commissie externe validering examenkwaliteit hoger beroepsonderwijs (2012). *Vreemde ogen dwingen. Eindrapport*. Den Haag: HBO-raad, vereniging van hogescholen.
- Corte, E. de, Geerlings, C.T., Lagerweij, N.A.J., Peters, J.J., & VandenBerghe, R. (1981). *Beknopte didaxologie*. Groningen: Tjeenk Willink.
- Creemers, B., & Slegers, P. (2003). *De school als organisatie*. In: N. Verloop & J. Lowyck (red.), *Onderwijskunde*. Groningen: Wolters-Noordhoff.
- Crombrugge, H. van (2011). *Vorming tussen gisteren en morgen*. In: Onderwijsraad, *Essays over vorming in het onderwijs*, Den Haag: Onderwijsraad, pp. 17-22.
- Dijsselbloem, J. (2008). *Tijd voor onderwijs*. Commissie Parlementair Onderzoek Onderwijsvernieuwingen (commissie Dijsselbloem) KST113842. 's-Gravenhage: Sdu Uitgevers
- Dochy, F., & Janssens, S. (2003). Evaluatie en assessment. In: N. Verloop & J. Lowyck (Eds.) *Onderwijskunde, een kennisbasis voor professionals*. Groningen/Houten: Wolters-Noordhoff.
- Eisner, E.W. (1976). Educational connoisseurship and criticism. Their form and functions in educational evaluation. *Journal of Aesthetic Education*, 1976, 135–150
- Eisner, E. W. (2004). Educational objectives, help or hindrance? In: D.J. Flinders & S. J. Thornton (Eds.). *The curriculum studies reader, second edition*. New York and London: Routledge Falmer.
- Gelder, L. van, Oudkerk Pool, T., Peters, J., & Sixma, J. (1973). *Didactische analyse*. Groningen: Wolters-Noordhoff.
- Gelwick, R. (1977). *The way of discovery. An introduction to the thought of Michael Polanyi*. New York: Oxford University Press.
- Good, Th.L. (2010). Forty years of research on teaching 1968-2008: What do we know that we didn't know then? In: R. J. Marzano (Ed.), *On excellence in teaching*. Bloomington IN: Solution Tree.

- Greene, M. (1969). Hobbes and the modern mind. In: M. Greene (Ed.). *The Anatomy of Knowledge*. Amherst: University of Massachusetts Press.
- Gude, R. (2012). Tijdloze Bildung. Over de Agora als model voor Bildung. In: Stralen, G. van & Gude, R. *En denken! Bildung voor leraren*. Leusden: ISVW Uitgevers.
- Hattie J. & Timperley, H. (2007). *The Power of Feedback, Review of Educational Research*. American Educational Research Association.
- Hattie, J.A.A., (2009). *Visible learning*. New York: Routledge.
- Hentig, H. von (1996). *Bildung. Ein Essay*. München/ Wien: Carl Hanser.
- Hoogland, J. (2005). Verlichting. Tussen persoonsideaal en beheersingsstreven. In: Muynck, A. de & Kalkman, B. (red.), *Perspectief op leren. Verkenningen naar onderwijs en leren vanuit de christelijke traditie*. Gouda: Projectgroep Leren in perspectief.
- Houten, D. van (2008). Professionalisering: een verkenning. In: Jacobs, G., Meij, R., Tenwolde, H. & Zomer, Y., *Goed werk. Verkenningen van normatieve professionalisering*. Amsterdam: SWP.
- IJzendoorn, M.H. (1997). Empirisch analytische pedagogiek. In: Miedema, S. (red.). *Pedagogiek in meervoud. Wegen in het denken over opvoeding en onderwijs*. Houten/ Diegem: Bohn Stafleu Van Loghum, 73-116.
- Jansen, H. (2009). *De leraar als koelie. Een gesprek over het leven, het werken en de identiteit van de leraar*. Amersfoort: Uitgeverij Agiel.
- Jong, T. de (2006). Nieuw leren en oude kennis: over bestaande evidentie voor de effectiviteit van “nieuwe” en “gecombineerde” vormen van leren. *Pedagogische Studien*, (83), 74-99.
- Kaldeway, J.A., Heer, C. de & Bulterman-Bos, J. (2014). *Opbrengstgericht of talentgericht werken? Een onderzoek naar onderwijsvisies*. Tijdschrift voor lerarenopleiders [in press].
- Kaldeway, J.A., Bulterman-Bos, J. (2013). De cognitieve switch in het licht van talentdiversiteit. *Tijdschrift voor lerarenopleiders*, 34 (1), 55-65.
- Kelchtermans, G. (1999). Kwetsbaarheid en professionele identiteit van leerkrachten basisonderwijs. Een exploratie van morele en politieke wortels. *Pedagogisch Tijdschrift* 24(4), 471-492.
- Kessels, J.P.A.M. & Korthagen, F.A.J. (1996). The relationship between theory and practice: Back to the classics. *Educational Researcher* 25(3), 17-22.
- Koninklijke Nederlandse Akademie van Wetenschappen (KNAW)(2009). *Rekenonderwijs op de basisschool. Analyse en sleutels tot verbetering*. Amsterdam: KNAW.
- Kohn, A. (2011). The case against grades. *Educational Leadership*, 69/3, 28-33.
- Koretz, D (2008). *Measuring Up. What educational testing really tells us*. Cambridge MA: Harvard University Press.
- Krathwohl, D.R., (2002). A revision of Bloom's Taxonomy: an overview. *Theory into practice*, Volume 41, Number 4.
- Kunneman, H. (2005). *Voorbij het dikke ik. Bouwstenen voor een kritisch humanisme*. Amsterdam: SWP.

- Labaree, D.F. (2000). On the nature of teaching and teacher education: difficult practices that look easy. *Educational Researcher* 51(3), 228-233.
- Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review* 55 (2), 178-194.
- Litt, Th. (1927/1965) *Führen oder Wachsenlassen. Eine Erörterung des Paedagogischen Grundproblems*. Stuttgart: Verlag Ernst Klett.
- Marzano, R.J. (2003) *What works in schools*, Alexandria VA: ASCD.
- Masschelein, J. & Simons, M. (2012). *Apologie voor de school. Een publieke zaak*. Leuven/ Den Haag: Acco.
- Ministerie OCW (2012). *Betere rekenprestaties scholieren basisonderwijs*. Nieuwsbericht 18-1-2012
- Muynck, A. de & Vos, P.H. (2014). 'Inspire the desire to aspire'. In: Keersmaekers, P., van Kerkhoven, M. & Vanspeybroeck, K., *Vormend onderwijs. Onderwijs vormen*. Antwerpen: Halewijn. [In press].
- Muynck, A. de & Roeleveld, M.E. (2005). *Competentiegericht leren, opleiden en begeleiden*. Amsterdam: Buijten & Schipperheijn, Dixit Reeks, deel 2.
- Muynck, A. de, Both, D.D. & Visser-Vogel, E. (2013). *Opbrengstgericht leren, meer dan presenteren. Een integrale aanpak van OGW en HGW*. Bussum: Coutinho.
- Mylle, J. (2011). *Beknopte didactiek en instructie*. Garant: Antwerpen/Apeldoorn.
- NRC, 5-11-2009. *Rapport: daling niveau rekenonderwijs ligt aan leraar*.
- Nussbaum, M. (2011). *Niet voor de winst. Waarom de democratie de geesteswetenschappen nodig heeft*. Amsterdam: Ambo.
- Onderwijsraad (2013). *Een smalle kijk op onderwijskwaliteit. Stand van educatief Nederland 2013*. Den Haag: Onderwijsraad.
- Palmer, P.J. (1998). *The Courage to Teach. Exploring the Inner Landscape of a Teacher's Life*. San Francisco CA: Jossey-Bass.
- Palmer, P. (2005). *Leraar met hart en ziel*. Groningen/Houten: Wolters-Noordhoff.
- Pikkareinen E. (2012). Signs of reality. The idea of General Bildung by J.A. Comenius. In: Siljander, P., Kivelä & Sutinen, A. (Eds). *Theories of Bildung and Growth. Connections and controversies between continental educational thinking and American Pragmatism*. Rotterdam/Boston/Taipei, Sense publishers. P. 19-30.
- Philipse, H. (2005). *Verlichtingsfundamentalisme? Open brief over verlichting en fundamentalisme aan Ayaan Hirsi Ali. Mede bestemd voor Piet Hein Donner*. Amsterdam: Bert Bakker.
- Polanyi, M. (1958). *Personal Knowledge: Towards a post-critical philosophy*. Chicago: University of Chicago Press.
- Procee, H. (1991). *Over de grenzen van culturen. Voorbij relativisme en universalisme*. Amsterdam/Meppel: Boom.
- Roefs, E. (2010). *Inspirerende docenten. Inzichten en verhalen uit het hoger beroepsonderwijs*. Antwerpen/Apeldoorn: Garant.

- Roeleveld, M.E. (2006). Over vorming en expliciete en impliciete factoren in het vormingsproces. In: Muynck, A. de & Vos, P.H. (red.). *Leren voor het leven. Vorming en christelijk onderwijs*. Amsterdam: Buijten & Schipperheijn. 33-57.
- Ryan, R.M., & Deci, E.L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54-67.
- Schumacher, E.F. (1973). *Small Is Beautiful: Economics As If People Mattered*. New York: HarperCollins.
- Shapin, S. (2005). *De wetenschappelijke revolutie*. Amsterdam: Balans
- Stamperius, H. (2011). *God en de Verlichting*. Vught: Skandalon
- Steenbergen, H. (2009). *Vrije en reguliere scholen vergeleken. Een onderzoek naar de effectiviteit van Vrije scholen en reguliere scholen voor voortgezet onderwijs*. Academisch proefschrift: GION, Gronings Instituut voor Onderzoek van Onderwijs.
- Struyven, K., Dochy, F. & Janssens, S. (2003). Students' perceptions about new modes of assessment: a review. In: Segers, M., Dochy, F., Cascallar, E. (Eds.). *Optimising new modes of assessment: In search for qualities and standards*. Boston: Kluwer Academic.
- Toulmin, S. (2001). *Return to reason*. Cambridge MA: Harvard University Press.
- Taylor, C. (2007). *A secular age*. London: Belknap Press and Harvard University Press.
- Tielemans, J. (2001). *Psychologie en didactiek in het onderwijs*. Leuven/Apeldoorn: Garant.
- Thijs, A. & Van den Akker, J. (red.) (2009). *Leerplan in ontwikkeling*. Enschede: SLO.
- Trouw, 5-11-2009. *Rekenniveau daalt door docent*.
- Vanderstraeten, Raf. (2001). 'Vorming: een sociaal-historisch ideaalbeeld van het opvoedingssysteem'. In: Vanderstraeten, Raf (red.). *Algemene vorming: constructie van een pedagogisch ideaal*. Leuven/Apeldoorn: Garant. 45-66.
- Werf, G. van der (2005). *Leren in het studiehuis. Consumeren, construeren of engageren*. Inaugurale rede. Groningen: GION.
- William, D. (2011). What is assessment for learning? *Studies in Educational Evaluation* 37, 3-14.

Personalia

Dr. Jacquélien Bulterman-Bos

Dr. Jacquélien Bulterman-Bos is lector Docent en Talent bij de Christelijke Hogeschool Ede en sinds 2014 verbonden aan de University of California (Berkely, VS). Ze deed ervaring op in het basisonderwijs, werkte in het havo-mbo, op de lerarenopleiding vo van het hbo en was docent op de VU en de UVA. In het bijzonder is ze geïnteresseerd in de manier waarop de onderwijswetenschap zich kan verbinden met de onderwijspraktijk. Daarover publiceerde ze in diverse internationale wetenschappelijke tijdschriften. Haar lectorale rede luidde: *Geef de onderwijsontwikkeling terug aan leraren*. Ze was betrokken bij het opzetten van de master Academisch Meesterschap van de UVA en de VU waar ervaren leraren onderwijsexperts worden die zowel een bijdrage leveren aan de onderwijswetenschap als aan de ontwikkeling van de onderwijspraktijk.

Dr. Bram de Muynck

Dr. Bram de Muynck is lector Christelijk leraarschap bij Driestar Hogeschool Gouda en leidt in die hoedanigheid het onderzoeksprogramma van het lectoraat. In de periode 2003 - 2011 was hij daar ook lector Onderwijs & Identiteit. Tevens is hij docent aan de masteropleiding Leren en Innoveren (Driestar Hogeschool, Christelijke Hogeschool Ede en de Gereformeerde Hogeschool) en inhoudelijk leider van de internationale minor Christian Education, Crossing the Borders. Ook is hij projectleider van één van de projecten van het Center of Expertise Persoonlijk Meesterschap. Dit expertisecentrum, een samenwerkingsverband van zeven hogescholen, heeft als doel kennis te ontwikkelen en te verspreiden die bijdraagt aan de vorming van 'persoonlijk meesterschap'.



De DIXIT-reeks wordt sinds 2005 uitgegeven bij Buijten & Schipperheijn *Motief* te Amsterdam. De reeks heeft als doel de rol van levensovertuigingen in beroepspraktijken te thematiseren.

Dixit is een vorm van het Latijnse werkwoord *dicere*, 'zeggen'. DIXIT is een reeks waarin onderzoeksresultaten van de hogescholen uit Gouda, Ede en Zwolle gepresenteerd worden. De redactie wil met deze reeks onderwijskundige boeken publiceren, voor diverse doelgroepen. De primaire doelgroep wordt gevormd door HBO-studenten, op Bachelor- en Master-niveau. Ook HBO-opgeleide professionals kunnen in deze reeks expertise vinden die recent op de hogescholen ontwikkeld is.

De DIXIT-redactie bestaat in 2014 uit dr. Bram de Muynck (lector *Christelijk Leraarschap*, Driestar Educatief te Gouda), dr. ir. Jan van der Stoep (lector *Religie in Media en Publieke Ruimte*, Christelijke Hogeschool Ede) en dr. Jan Hoogland (lector *Samenlevingsvraagstukken*, Gereformeerde Hogeschool te Zwolle). Drs. Wouter van Grootheest, docent aan de Christelijke Hogeschool Ede, is redactiesecretaris.

DIXIT

IS ALLES VAN WAARDE MEETBAAR?

Hoe ontwikkel je als onderwijsinstelling een evenwichtig toetsbeleid? In het (hoger) beroepsonderwijs, het voortgezet onderwijs en het basisonderwijs neemt het gebruik van toetsen toe. Tegelijkertijd verschijnen er kritische rapporten over toetsing en ontstaan collectieven om toetsen af te schaffen. Voor de onderwijspraktijk is het niet gemakkelijk om in te spelen op de wisselende accenten van de politiek. Nu eens staan de eigen identiteitsvorming en de zelfwerkzaamheid van jongeren centraal, dan weer de externe eisen waar zij aan moeten voldoen. Hoe ga je daarmee om als docent, als afdelingsleider of als examencommissie? Is alles van waarde meetbaar?

Dit boek helpt docenten, beleidsmakers en andere onderwijsbetrokkenen bij het ontwikkelen van een toetspraktijk die verder kijkt dan de waan van de dag. Een toetspraktijk die het belang van kennis erkent, maar ook het belang van persoonlijke vorming en creatief denken.

Het boek is bijzonder geschikt om te gebruiken voor gezamenlijke scholings- en beziningsmomenten. Het bevat herkenbare praktijkvoorbeelden, handige achtergrondinformatie en onderzoeksgegevens. Met discussievragen en een zelftoets worden onderwijsprofessionals uitgedaagd om de eigen visie te verhelderen en te herijken.



Dr. Jacqueline Bulterman-Bos is lector Docent en Talent bij de Christelijke Hogeschool Ede en sinds 2014 verbonden aan de University of California, Berkeley (VS).

Dr. Bram de Muynck is lector Christelijk Leraarschap bij Driestar Hogeschool Gouda.



Buijten & Schipperheijn *Motief* – Amsterdam

ISBN 978-90-5881-787-7



9 789058 817877