

Toetsvragen Schrijven

Ben Wilbrink

Handreiking bij het bedenken van vragen over de leerstof, geschikt om in (eind-)toetsen te gebruiken.

oorspronkelijk ISBN 90-274-6674-0 SISO 450.6 UDC 371.26 UGI 566

1983 Het Spectrum Aula 809 Reeks Onderwijskundige informatie voor het Hoger Onderwijs, onder auspiciën van de Contactgroep Research Wetenschappelijk Onderwijs, redactie G. F. Bernaert, J. H. Daniëls, G. W. H. Heijnen, J. F. J. M. van Hout, M. J. A. Mirande, J. Nedermeijer, A. Pilot, D. W. Vaags en A. I. Vroeijenstein.

Het copyright berust bij de auteur. De tekst is vrij beschikbaar, respecteert u wel mijn auteursrechten.

Inhoud

Bij deze digitale versie

Voorwoord 7

1 Inleiding 10

1.1 Toetsvragen schrijven: kunst of kunde? 11

1.2 Uitgangspunten 13

1.3 Inhoudsoverzicht 17

2 Vraagsoorten, doorzichtigheid, rompvragen en abstractieniveau 20

2.1 Aanvul- of invulvragen 21

2.2 Keuzevragen 24

2.3 Opstelvragen 34

2.4 Doorzichtigheid 37

2.5 Rompvragen: meervoudig bruikbare vraagskeletten 44

2.6 Vermijd abstracte vraagstellingen 50

2.7 Literatuur 54

3 Leerstofinventarisatie 57

3.1 (Indirect) waarneembare termen 59

3.2 Abstracte termen en constructs 61

3.3 Theoretische termen 64

3.4 Onderlinge verknooptheid van termen 66

3.5 Varianten van 'definities' 70

3.6 Literatuur 78

4 Toetsvragen schrijven bij afzonderlijke termen 81

4.1 Vertalen 82

4.2 Definiëren 86

4.3 Voorbeelden geven 91

4.4 Voorbeelden herkennen en benoemen 98

4.5 Herkennen en benoemen bij formeel gedefinieerde termen 108

4.6 Beschrijvende uitspraken 113

5 Toetsvragen schrijven bij relaties tussen termen 115

5.1 Vertalen en afbeelden 17

5.2 Onderscheiden 122

5.3 Classificaties 128

5.4 Stappenschema's, algoritmen, routines 132

5.5 Wetmatige relaties 137

5.6 Literatuur 142

6 Toetsvragen schrijven bij tekst 143

6.1 Participatiecontrole 144

6.2 Thema's en hoofdpunten 146

6.3 Analyse 152

6.4 Inferentie 158

6.5 Literatuur 163

7 Problemen stellen 164

7.1 Over problemen gesproken 165

7.2 Inventarisatie 177

7.3 Vuistregels 182

7.4 Literatuur 191

8 Kwaliteit van toetsvragen 193

8.1 Regels bij het examineren 194

8.2 Punten om op te controleren 200

8.3 Onafhankelijke beoordeling van kwaliteit 208

8.4 Controlelijsten in beknopte vorm 217

8.5 Literatuur 218

Gerefereerde literatuur 220

Register 226

Bij deze digitale versie

Na twee decennia is er wel een sterke behoefte om de stijlfouten te verbeteren en de tekst beter leesbaar te maken, maar ik heb in de tussenliggende jaren geen redenen gevonden om aan de inhoud te tornen. Het blijft een leemte dat aan het vragen van verklaringen geen aandacht is besteed, maar het is mij evenmin duidelijk dat dit een onderwerp zou zijn dat om een bijzondere behandeling vraagt. Het vragen naar verklaringen is zelden contextvrij, en de betreffende context zal vaak kunnen bepalen wat adequate vraagformuleringen zijn. Als casus van de toepassing van de ontwerptechnieken uit dit boek is op mijn website een korte handleiding voor het ontwerpen van toetsvragen over een inleiding taalwetenschap beschikbaar.

De wijze waarop in dit boek het schrijven van toetsvragen is behandeld is vandaag nog bijna even uniek als dat in 1983 het geval was. Mij is sinds 1983 alleen de DOZ-publicatie (1991) over toetsen bekend waarin deze aanpak van ontwerpen van toetsvragen is overgenomen. *Toetsvragen schrijven* heeft geen school gemaakt, en vult nog steeds dezelfde lacune in de onderwijskundige literatuur. Dat blijkt mij bij tijd en wijle ook uit commentaar van mensen die dit boek op hun pad vinden. De koudwatervrees die sommige collega's voor de benadering in dit boek hebben, volgt mogelijk uit ontzag voor de omvangrijke toetsliteratuur die is gebaseerd op methoden uit de psychologische testleer. Was toetsen in het onderwijs maar zo eenvoudig als het testen voor selectie of diagnostiek. De cognitieve psychologie is een betere voedingsbodem voor een ontwerptheorie voor toetsvragen. Voor inspiratie of verbreding zijn de sindsdien door Reigeluth (1983, 1987) bezorgde boeken aan te bevelen, naast de in *Toetsvragen schrijven* zelf gebruikte literatuur, zoals Hempel, over de aard en structuur van disciplinaire kennis en theorieën. Door de beperking tot het ontwerpen van toetsvragen, blijft de veel bredere problematiek van beoordelen van personen buiten zicht, zie daarvoor evenwel Hofstee (1999).

Over het gebruik van toetsen in het onderwijs zijn ondertussen meerdere publicaties van mijn hand verschenen. Het is geen onderzoek dat specifiek is gericht op het onderbouwen van deze constructieve benadering van toetsvragen schrijven, maar biedt daar wel ondersteuning voor, al was het maar omdat ook in dat werk bijvoorbeeld het beginsel van doorzichtigheid een belangrijke rol speelt, zoals Adriaan D. de Groot dat in 1970 voorstelde. De ontwerptheorie voor toetsvragen vormt samen met mijn uitwerking van het tentamenmodel van Van Naerssen (zie mijn website voor de laatste ontwikkelingen) een begin van een didakometrie, waaraan evenwel de systemische processen nog ontbreken (zie mijn congrespapers

uit 1992 voor een toepassing van de sociale systeemtheorie van James Coleman op een dataset van propedeusegegevens (website)).

A. D. de Groot heeft het boek met plezier gelezen. Hij heeft me toestemming gegeven om zijn antwoord op mijn vraag 'Waarom toch die dogmatiek van de vierkeuzevraag?' te publiceren: destijds speelde bij de oprichting van het Cito het gevoel dat nieuwe instituten liever niet te belasten met een veelvoud van mogelijke vormen voor keuzevragen. Het blijft mensenwerk, de pretentie van wetenschappelijkheid is niet altijd even serieus te nemen. Twee decennia later blijkt het voor een auteur met een genuanceerde kijk op 'objectieve toetsen' toch even lastig om tegen de staande dogmatiek in zijn werk gepubliceerd en aanvaard te krijgen. Het eerste is gelukt, met de steun van Tom Dousma en Ad Horsten, het tweede staat nog open.

In deze elektronische versie is de oorspronkelijke paginering van Aula 809 aangehouden, zijn enkele fouten verbeterd, zoals de penguin die weer pinguin mag heten, maar is de helaas gebrekkige stijl onveranderd gelaten.

Ben Wilbrink
Leiden, september 2005
benwilbrink.nl

DOZ. *Toetsen en beoordelen*. Culemborg: PHAEDON, 1991. ISBN 90-72456-35-1. Een helder boek, geschreven door een werkgroep uit het HBO, voor het HBO.

Hofstee, W. K. B. *Principes van beoordeling: Methodiek en ethiek van selectie, examinering en evaluatie*. Lisse: Swets & Zeitlinger, 1999.

Reigeluth, Charles M. (ed.). *Instructional-design theories and models. Overview of their current status*. Hillsdale, New Jersey: Lawrence Erlbaum, 1983.

Reigeluth, Charles M. (ed.). *Instructional theories in action. Lessons illustrating selected theories and models*. Hillsdale, New Jersey: Lawrence Erlbaum, 1987.

Wilbrink, B. (1984). Toetsvragen schrijven. *Onderzoek van Onderwijs*, 1984, 13, 20-21.

Voorwoord

De onderwijskunde is een jonge discipline, en nog niet in staat op alle dringende praktische vragen antwoord te geven. Een zo'n leemte is het ontbreken van werkbare technieken die de docent nodig heeft om op efficiënte wijze kwalitatief goede toetsvragen te schrijven. Het belang van goede vuistregels voor het schrijven van toetsvragen is evident: hoe en waarop studenten getoetst en beoordeeld worden, bepaalt in niet geringe mate de kwaliteit en de opbrengst van het onderwijs. Zolang toetsvragen op intuïtieve en impressionistische wijze bedacht worden, omdat docenten nu eenmaal niet over betere mogelijkheden beschikken, zal het lastig zijn om de kwaliteit van het onderwijs te verbeteren, op welke manier dan ook.

Dit boek is een eerste poging een samenhangend geheel van vuistregels te bieden, vooral gericht op de grote diversiteit aan 'leerstof' die de student in het Hoger Onderwijs aangeboden wordt. Deze nadruk op de inhoudelijke kant van toetsvragen onderscheidt dit boek ook van de eerste Amerikaanse poging om de beschikbare technieken voor het schrijven van toetsvragen in een overzichtelijke vorm voor onderwijspractici te presenteren: *Technology of test item writing* van Roid en Haladyna (1982), daar wordt het ideaal juist gelegd in zoveel mogelijk leerstofonafhankelijke technieken. Beide boeken zijn dan ook als enigszins complementair te zien.

De geboden vuistregels zijn bruikbaar, maar zeker niet volmaakt. Een ernstige tekortkoming is bijvoorbeeld de voortdurende beknoptheid van behandeling. Dat lijkt paradoxaal: enerzijds wordt hier in een bijna volstrekte leemte voorzien, anderzijds is het hier gebodene niet meer dan een bescheiden begin van de opvulling van die leemte, en lijkt een meer uitgewerkte invulling binnen handbereik te liggen (maar binnen één boek ontbreekt daartoe ten enen male de ruimte). Zo ontbreekt een systematische behandeling van vuistregels voor vragen die gericht zijn op het geven van verklaringen, en dat terwijl het kunnen verklaren van verschijnselen zo'n prominente rol in het wetenschappelijk onderwijs speelt. Slechts enkele bladzijden behandelen het kunnen leggen van dwarsver-

banden (analyse en inferentie), waar eerder een behandeling in afzonderlijke hoofdstukken op zijn plaats zou zijn. Toetsen op het oplossen van problemen krijgt een karige behandeling in één enkel hoofdstuk. Ook brengt de algemeen gerichte behandeling met zich mee dat de lezer geen kant-en-klare vuistregels voor het eigen vakgebied zal aantreffen, maar dat deze telkens afgeleid moeten worden van de algemene regels en de voorbeelden die gegeven worden.

Een enkel woord nog over aard van de geboden vuistregels. Ik heb gekozen voor vuistregels die de vragen doen aansluiten bij de aard en de structuur van de leerstof waarover gevraagd wordt. Dat is in tegenstelling tot het in de onderwijskunde meer gangbare denken in psychologische termen, waaruit vuistregels zouden kunnen volgen die leiden tot toetsvragen die bepaalde denkprocessen van de student eisen. Het psychologische uitgangspunt overheerst in de onderwijskunde, maar wordt nu toch wel hier en daar genuanceerd (Anderson, Spiro en Montague, 1977), of verlaten (Furst, 1981). Het verrassende is nu dat verschillende eenvoudige wijsheden uit de wetenschapsfilosofische hoek heel bruikbare aanknopingspunten bieden. Het gaat dan om de aard en de functie van definities (veel toetsvragen blijken over definities te gaan), het onderscheid tussen observeerbare, abstracte en theoretische begrippen, en dergelijke. Niet alleen zijn daar uitstekende vuistregels voor het schrijven van toetsvragen op te baseren, maar de hele aanpak zal de docent meer vertrouwd voorkomen dan de traditioneel onderwijskundige, op denkprocessen gerichte benadering. Van Westrhenen (1977) is een van de weinigen die eenzelfde aanpak al eerder volgde. Deze meer filosofische oriëntatie hoeft de lezer niet af te schrikken: diepzinnigheden uit de loopgravenoorlog tussen empiristen, realisten en antirealisten (zie bijv. Suppe, 1977) raken de hier geboden vuistregels en hun toepassing niet. Ik hoop dat ik heb kunnen laten zien dat het bedenken van vragen over de leerstof fascinerende kanten heeft, dat het kan leiden tot verdieping van het inzicht in de aard van en de samenhangen in de informatie die de studenten ter verwerking voorgelegd wordt.

Rest mij nog de redactie te danken voor de moed, het geduld en de welwillendheid bij het bespreken van de vele

volgende versies van het manuscript. In het bijzonder wil ik voor hun commentaren bedanken: Jan Daniëls, Wim Vaags, Henk Wijffels, Marcel Mirande, Albert Pilot en Ton Vroeijenstijn.

BEN WILBRINK
Amsterdam 1982

'... an experience, a very humble experience, is capable of generating and carrying any amount of theory (or intellectual content), but a theory apart from an experience cannot be definitely grasped even as a theory.'

JOHN DEWEY, in: *Democracy and education*.

1 Inleiding

In dit boek wil ik enkele handreikingen doen aan docenten die voor de opgave staan om over de eigen leerstof vragen te bedenken en te stellen.

De nadruk ligt daarbij op vragen die in beoordelingssituaties gebruikt moeten worden: aan deze vragen zijn immers andere - en vaak strengere - eisen te stellen dan aan de vragen die de docent in meer informele didactische situaties gebruikt. Desondanks zijn de aanwijzingen in met name de hoofdstukken 4 tot en met 7 ook bruikbaar voor het bedenken van deze louter didactisch te gebruiken vragen.

Het bedenken van telkens weer nieuwe toetsvragen heeft docenten en zelfs professionele schrijvers van toetsvragen altijd veel moeite gekost. De reden is dat er geen technologie voor het schrijven van toetsvragen beschikbaar is, zodat de toetsvragenschrijver vaak het idee heeft dat iedere te bedenken vraag hem of haar voor een nieuw probleem stelt. Bij een 'technologie' voor het schrijven van toetsvragen zijn verschillende mogelijkheden denkbaar. Een bescheiden technologie is al lange tijd beschikbaar, wanneer goede overzichten van mogelijke vraagsoorten daaronder verstaan worden. Zo geven Dousma en Horsten (1980) een overzicht van meerkeuzevragen, aanvulvragen, opstelvragen en dergelijke, met hun voor- en hun nadelen, waaruit de lezer die ene vraagsoort kan kiezen die geschikt lijkt voor de te toetsen leerstof.

Zo'n overzicht is informatief, maar lost niet het probleem op hoe in de eerste plaats aan *ideeën* voor nieuwe vragen te komen, hoe je vragen kunt bedenken of liefst construeren.

Een ander type technologie is gericht op het kunnen formuleren van een grote diversiteit van toetsvragen, gebaseerd op een enkele grondvraag, in dit boek *rompvraag* genoemd. Het gaat ook hier om een uiterlijke kant van het schrijven van toetsvragen, ongeveer te beschrijven als een technologie voor het genereren van vraagvarianten volgens een tevoren zorgvuldig gespecificeerde algemene omschrijving van deze vraagvarianten. Een uitstekend overzicht van juist deze technologie wordt gegeven door Roid en Haladyna (1982).

De zwakte in deze technologische benadering zit ook weer in

het startpunt: het wordt overgelaten aan de inspiratie van de toetsvragenschrijver om het *idee* van een bepaalde rompvraag te genereren. Er wordt uitgelegd wat rompvragen zijn en hoe je ze gebruikt, maar niet hoe je ze bedenkt.

Een aanvullende technologie is dan ook nodig, een technologie die het mogelijk maakt om langs meer systematische weg vragen te *bedenken*, op ideeën voor toetsvragen te komen. Dat betekent dat naar vuistregels gezocht wordt die tamelijk rechtstreeks de leerstof omvormen tot vragen over zichzelf. Het zijn deze vuistregels die in dit boek gepresenteerd worden, waar passend aangevuld met het bruikbare uit beide eerder genoemde technologieën.

1.1 Toetsvragen schrijven: kunst of kunde?

Het is een wijdverbreide opvatting dat het schrijven van toetsvragen alleen goed mogelijk is voor wie dat van nature in de vingers heeft en over veel ervaring beschikt.

‘Iedere toets begint met een idee van de toetsvragenschrijver. Een van de moeilijkste problemen waarmee hij geconfronteerd wordt, is het produceren en selecteren van ideeën voor zijn toetsvragen. (...) Er is geen automatisme dat ideeën voor zijn toetsvragen genereert. Ze moeten worden uitgevonden of ontdekt, en in dat proces zijn toevallige gedachten en inspiraties erg belangrijk.’ Aldus Wesman in zijn hoofdstuk ‘Writing the test item’ in het standaardwerk van Thorndike (1971). In zijn inleiding stelt Wesman nogal nadrukkelijk dat het schrijven van toetsvragen in wezen creatief is - het is een kunst. ‘Op dezelfde wijze als er geen verzameling formules bestaat voor het maken van een goed verhaal of schilderij, kan er geen verzameling regels bestaan die de vervaardiging van goede toetsvragen garanderen.’

Het zou niet alleen voor de toetsvragen schrijvende docent jammer zijn wanneer Wesman gelijk heeft. Veel spijtger is het voor de studenten. De onvermijdelijke tegenhanger van toetsvragen die uit de kunst zijn, is immers dat het beantwoorden van dergelijke vragen ook een kunst is.

Wanneer de toetsvragenschrijver geen werkbare technieken tot zijn beschikking heeft, dan kan al snel iedere nieuw te be-

denken vraag een hersenbreker worden. De vraag is dan: wat doe je als toetsvragenschrijver wanneer je nadenkt over het formuleren van een vraag over een bepaald onderwerp? Zijn er vuistregels te geven waardoor dat nadenken richting gegeven wordt, waardoor het minder afhankelijk wordt van toevallige eigenschappen van het onderwerp waarover een vraag geschreven moet worden? Die vuistregels worden in dit boek ontwikkeld. En eigenlijk ligt het 'programma' voor die vuistregels uitermate voor de hand. In het eerder genoemde standaardwerk van Thorndike over 'educational measurement' is het ook geformuleerd, zij het niet door Wesman: 'De toetsvragenschrijver die het ernstig meent met de inhoudelijke representativiteit (content validity) van de vragen, begint niet met het neerpennen van vragen zoals ze hem te binnen schieten, en ook niet door nuchter een vraag per bladzijde tekst te bedenken maar door ieder van zijn onderwijsdoelen te vertalen in taakomschrijvingen. Gewoonlijk levert dat het beste resultaat op wanneer globale doelstellingen nogal fijn onderverdeeld worden. Binnen ieder van die onderverdelingen is het niet voldoende om de onderwerpen louter op te sommen, maar moet voor ieder onderwerp aangegeven worden op welke vorm van beheersing van dat onderwerp het onderwijs gericht is.

Het moet duidelijk zijn of de student te maken krijgt met technische termen met in dagelijkse termen beschreven situaties, met afbeeldingen van situaties of met concrete dingen.

Ook specificaties van de verlangde antwoorden zijn van belang. Maar al te vaak wordt de meerkeuzevraagvorm als vanzelfsprekend beschouwd. Als de toetsvragenschrijver onbevangen nadenkt over zijn doelen, zal hij er vaak toe besluiten dat de taak vraagt om antwoorden die door de student geconstrueerd worden - om geschreven antwoorden op aanvulvragen, of om mondelinge antwoorden om belemmeringen zo klein mogelijk te houden.' Aldus Cronbach in Thorndike (1971, p. 458).

Dit 'programma' is in dit boek uitgevoerd, ook al heeft het daarmee ongetwijfeld niet zijn optimale vorm gekregen. Het inventariseren van de leerstof wordt in hoofdstuk 3 behandeld, en latere hoofdstukken vullen dat verder in. De verschillende vormen van stofbeheersing en hoe die in toetsvragen 'vertaald' worden, zijn het onderwerp van hoofdstukken 4 tot

en met 7. De kwestie van de te kiezen vraagvorm, meerkeuze of andere, wordt in hoofdstuk 2 aangevat langs de kritische lijn zoals Cronbach die heeft aangegeven.

1.2 Uitgangspunten

De vuistregels voor het schrijven van toetsvragen zoals ik ze in dit boek presenteer, berusten op een aantal gekozen uitgangspunten die het karakter van de vuistregels sterk bepaald hebben. Omdat deze uitgangspunten ook al aangeven wat de lezer wèl en wat níet mag verwachten, verdienen ze een korte bespreking.

Geen verschillen tussen studenten willen vastleggen. Er bestaat tegenwoordig een grote mate van overeenstemming waar het gaat om de wenselijkheid de beoordeling te richten op de mate waarin de student de stof beheerst, in tegenstelling tot beoordeling die verschillen tussen studenten wil vastleggen. Het gaat er allereerst om dat de toetsvragen representatief zijn voor wat er in het onderwijs is behandeld, dat zij overeenstemmen met de doelen van het onderwijs. Wanneer een hoge mate van stofbeheersing een algemeen doel van het onderwijs is, en wanneer dat doel ook bereikt wordt, dan zal de toets kunnen bestaan uit vragen die correct beantwoord worden door bijna iedereen die het onderwijs gevolgd heeft. Dan zullen verschillen tussen studenten onderling relatief klein zijn, en meer te maken hebben met pech of geluk dan met verschillen in de mate van stofbeheersing. Dit uitgangspunt is van belang voor de vuistregels voor het schrijven van toetsvragen zoals in dit boek gegeven, omdat toetsvragen die geschikt zijn om verschillen tussen studenten onderling mee vast te leggen, in strijd kunnen komen met de eis dat toetsvragen representatief moeten zijn voor de onderwijsdoelen. Bijvoorbeeld: met 'gemakkelijke' vragen kunnen geen verschillen tussen studenten gemeten worden, omdat iedereen deze vragen ongeveer even goed maakt, terwijl ze wel degelijk betrekking kunnen hebben op de 'kern van de stof' (de vragen zijn juist gemakkelijk omdat er aan die 'kern van de stof' zoveel aandacht is besteed).

De toets heeft feed-forward-effecten. Vanwege het belang dat

de student heeft bij een gunstig toetsresultaat, zal hij zijn voorbereiding zo goed mogelijk proberen af te stemmen op wat hij verwacht dat er gevraagd zal worden. Wat gevraagd kan worden, is van belang; wat waarschijnlijk *niet* gevraagd wordt, dat is geen aandacht en tijd waard. De toets straalt daarmee een sterk effect *naar voren* uit, ik noem dat feed-forward (In analogie met feed-back, die achteraf tot effecten leidt). Deze stand van zaken is niet zonder meer goed of slecht te noemen.

Slecht is het wanneer de toets met een grote mate van geheimzinnigheid is omgeven, en studenten zich daardoor in allerlei malle bochten gaan wringen om er toch achter te komen wat er van hen verwacht wordt. Het laakbare gedrag is niet dat van de student, maar integendeel het achterhouden van informatie door de docent. Laakbaar is het omdat studenten recht hebben op goede informatie over de vragen die ze mogen verwachten. Laakbaar is het ook omdat er een negatief gebruik van de toets wordt gemaakt, in plaats van een positief gebruik. Hoe je het ook wendt of keert, de toets zal altijd een feed-forward-effect hebben. Welnu, maak daar dan ook goed gebruik van, tennutte van het onderwijs, en in het belang van de studenten.

Een goed gebruik houdt in: geef zoveel mogelijk informatie over de te verwachten toetsvragen (maak de toetsing zo doorzichtig mogelijk), en zorg ervoor dat onderwijs en toets op elkaar afgestemd zijn (toets wat onderwezen is, onderwijs wat getoetst wordt). Het is duidelijk dat de aard van de toetsvragen in dit feed-forward-spel van belang is, evenals de representativiteit van de vragen gezamenlijk voor de doelen die in het onderwijs nagestreefd worden.

Onderwijsdoelen. De vragen die in de toets opgenomen worden, moeten in overeenstemming zijn met de doelen van het onderwijs: de toets moet representatief zijn voor wat er in het onderwijs behandeld is. Eigenlijk is dat al een vuistregel voor het schrijven van toetsvragen, zij het ook een wel heel erg algemene. Het uitgangspunt is niet letterlijk bedoeld: het gaat er niet om dat in de toets nog weer eens alle opgaven voorkomen die in het onderwijs behandeld zijn, maar het is wel de bedoeling dat de opgaven in de toets representatief zijn voor het uiteindelijk in het onderwijs nagestreefde en ook bereikte niveau van stofbeheersing. Een opmerkelijke implicatie hiervan is dat ook onderwerpen die perfect beheerst worden, in de

toets gevraagd zullen worden.

Het woord 'onderwijsdoelen' is in deze paragraaf al herhaaldelijk gevallen, zoals het ook in het vervolg nog vaak gebruikt zal worden. Ik veronderstel dat er inderdaad *doelen* zijn, dat het onderwijs geen chaotisch tijdverdrijf is. Dat hoeft echter niet in te houden dat de doelen ook expliciet als doelen gesteld zijn, en zeker niet in de grote mate van detail die het zou toelaten toetsvragen direct uit de onderwijsdoelen af te leiden. Ik ga ervan uit dat de docent, ook zonder onderwijsdoelen op papier beschikbaar te hebben, heel goed de grens kan trekken tussen vormen van stofbeheersing die er wel en die er niet onder zouden vallen, zouden de doelen expliciet geformuleerd zijn. Het uitgangspunt heeft de schijn van grote subjectiviteit, en een subjectief element zit er zeker in. Maar laten we niet over het hoofd zien dat vele dingen expliciet vastgelegd zijn: de leerstof (boeken, artikelen, syllabi), en de toetsvragen (een uitgebreide verzameling die representatief is voor de leerstof, of tenminste toetsvragen zoals die de laatste keren in de toets gebruikt zijn). Dit materiaal geeft aan waar het in het onderwijs om gaat, zoals ook expliciete doelen dat (in het kort) weergeven. Wanneer er gedetailleerde onderwijsdoelen geformuleerd zijn, is dat prachtig; ze kunnen dienen om alle betrokkenen scherper bewust te maken waar het in het onderwijs om gaat; ze kunnen het schrijven van toetsvragen vergemakkelijken; misschien verduidelijken ze ook wat onder een representatieve toets verstaan moet worden. Maar onmisbaar zijn ze niet: ongeschreven onderwijsdoelen zijn als het ware virtueel aanwezig achter de leerstof en het toetsvragenbestand. Wanneer er in dit boek gesproken wordt over onderwijsdoelen, mag dat letterlijk opgevat worden, maar er is geen bezwaar tegen om daarbij te denken aan virtuele doelen. Het uitgangspunt is gebaseerd op het onbehaaglijke gevoel dat de formulering van gedetailleerde onderwijsdoelen op zich een goede zaak is, maar dat de grote hoeveelheden tijd en energie die ermee gemoeid zijn, beter in het maken van onderwijs gestoken kunnen worden, bijvoorbeeld in het schrijven van goede toetsvragen.

Wie het maken van een leerstofinventarisatie opvat als een vorm van doelen stellen, zou het bovenstaande verkeerd kunnen begrijpen. Voor het schrijven van toetsvragen is het vrijwel altijd gewenst in enigerlei vorm een leerstofinventarisatie

te maken, en aan dat inventariseren zal in dit boek ruime aandacht gegeven worden.

De vuistregels moeten algemeen bruikbaar zijn. De disciplines die we binnen het wetenschappelijk onderwijs aantreffen, verschillen sterk van elkaar in karakter, ook wat de leerstof en de vragen daarover betreft. In het zicht van die verscheidenheid is de verleiding groot om vuistregels voor het schrijven van toetsvragen te beperken tot één cluster van verwante vakken. Het voordeel zou zijn dat de vuistregels een wat grotere 'kracht' zouden hebben, en dat er voorbeelden bij gegeven kunnen worden die rechtstreeks toegeschreven zijn op de betreffende vakken. Het nadeel is dat op deze wijze er een speciale onderwijskunde ontwikkeld zou moeten worden voor ieder afzonderlijk cluster van vakken (geneeskundige, natuurkundige, technische, sociaal-wetenschappelijke enz.).

Dit perspectief is zo onaantrekkelijk dat er niet anders dan voor de algemene benadering gekozen kan worden, ook al zal die moeizamer zijn, wat minder krachtige vuistregels opleveren, en docenten minder direct aanspreken omdat er meer aan het eigen initiatief wordt overgelaten. De voorbeelden van toetsvragen zijn gekozen op een zo groot mogelijke doorzichtigheid voor lezers uit verschillende disciplines. Dat brengt onvermijdelijk met zich mee dat er soms de reuk van trivialiteit aan kleeft, maar bedenk dat eenvoud de voorbeelden voor u toegankelijker maakt (opgaven uit de thermodynamica zullen een psycholoog niets zeggen).

Basis creëren voor toetsanalyse en cesuurbepaling. De kwaliteit van de vragen die samen de toets uitmaken, is bepalend voor de mogelijkheden die er zijn bij het analyseren van toetsresultaten, en voor de af te leggen verantwoording waar het de gekozen zak/slaag-grens betreft. Ook vanuit deze optiek is het gewenst dat het schrijven van toetsvragen geen kunst, maar kunde is. Toetsanalyse en cesuurbepaling zijn onderwerpen die in dit boek niet terloops ook nog behandeld kunnen worden. Ik moet volstaan met te verwijzen naar op dit moment beschikbare publikaties (Wilbrink 1978, 1979, 1980). Voor een overzicht van de meer klassieke methoden van toetsanalyse en cesuurbepaling kan de lezer terecht bij Dousma en Horsten (1980).

1.3 Inhoudsoverzicht

Dit boek heeft een nogal zware informatieve lading, waarbij niet alles voor iedereen even relevant is. Het is de bedoeling dat de lezer aan de hand van de gegeven vuistregels zelf toetsvragen gaat schrijven. Daarom bevatten de verschillende hoofdstukken in de eerste plaats concrete aanwijzingen, lijsten met mogelijkheden waaruit keuzen gemaakt moeten worden, geboden en verboden. Met daaromheen gerangschikt een meer uiteenzettend verhaal. De verschillende hoofdstukken vormen een eenheid in deze zin dat een globaal kennismaken ervan een noodzakelijke achtergrond vormt om van de vuistregels in dat ene hoofdstuk waarin u speciaal geïnteresseerd bent, verstandig gebruik te kunnen maken. Daarom is het een goede strategie om eerst het boek globaal door te nemen, en daarbij aantekening te maken van de passages of hoofdstukken die voor u van direct praktisch belang zijn, om die vervolgens grondiger te gaan bekijken.

Hoofdstuk 2 bespreekt de diverse vraagvormen waaruit te kiezen valt, met de voor- en nadelen ervan, bijvoorbeeld die van meerkeuzevragen. De nadruk wordt gelegd op de aard van de te toetsen leerstof: die bepaalt welk vraagtype geschikt is en dan ook gebruikt moet worden. De meeste lezers zullen al enige ervaring in het schrijven van toetsvragen hebben opgedaan, en beschikken over een bescheiden verzameling van vragen. Zij zijn mogelijk bijzonder geïnteresseerd in technieken die de uitbouw van een kleine vragenverzameling tot een wat grotere verzameling vergemakkelijken (zie 2.5). Er wordt aandacht besteed aan het belang van een juist niveau van abstractie in de toetsvragen, en dat komt meestal neer op een *lager*, meer concreet niveau dan gebruikelijk: liever met nieuwe voorbeelden werken dan met algemene verbale omschrijvingen en definities. Dat maakt het ontwerpen van nieuwe toetsvragen en vooral van meerdere varianten ook gemakkelijker.

Hoofdstuk 3 lijkt een omtrekkende beweging te zijn: er worden geen vuistregels voor het schrijven van toetsvragen gegeven, maar voor het maken van een leerstofinventarisatie. Zo'n inventarisatie is een noodzakelijke tussenstap: je moet tenslotte weten waarover wel en waarover geen vragen gesteld zullen

worden. De te investeren tijd komt er dubbel en dwars weer uit, omdat een goede inventarisatie het schrijven van toetsvragen gesmeerd zal doen verlopen. Toch zijn dit maar bijkomstige redenen. Belangrijker is dat de docent zich er rekenschap van geeft wat de *aard* van de leerstof is, wat er redelijkerwijs wel en wat er niet over gevraagd kan worden. Pas dan is het mogelijk om uit datgene wat er wel gevraagd kan worden - en dat is altijd veel meer dan van studenten verwacht mag worden - die beperkte keuze te maken die goed overeenstemt met de onderwijsdoelen.

De hoofdstukken 4 en 5 geven aanwijzingen voor het schrijven van toetsvragen bij de *termen* en hun *relaties* die uit de leerstofinventarisatie van hoofdstuk 3 zijn gerold.

Hoofdstuk 6 gaat dieper in op het speciale geval van het 'literatuur'-tentamen, waarin minder de nadruk ligt op een diepgaande beheersing van een beperkt aantal nieuwe termen en relaties dan op een *overzicht* van enkele bredere onderwerpen. Juist het schrijven van toetsvragen over tekst wordt vaak als erg lastig ervaren.

Hoofdstuk 7 bespreekt eveneens een speciaal onderwerp: het oplossen van problemen. Voor een deel gaat het daarbij om vaardigheden die in voorgaande hoofdstukken al aan de orde zijn geweest, zoals het behoorlijk formuleren of herformuleren van de probleemstelling (zie 5.1), het analyseren van het probleem (zie 6.3), en het gebruik maken van kennis en inzichten waarover je zelf beschikt (inferentie, zie 6.4). Het accent ligt echter op het eigen karakter van het, aanpakken en oplossen van problemen, dat planmatig gebeurt en waarbij verschillende stadia te onderscheiden zijn; bij het ontwerpen van probleemstellingen kan dat inzicht in de structuur van het aanpakken en oplossen van problemen goed uitgebuit worden. Evenals in hoofdstuk 3 ligt ook in dit hoofdstuk de nadruk op het expliciteren van de aard van de leerstof, in dit geval de probleemaanpak die de student zich eigen moet maken.

Hoofdstuk 8 vormt de logische sluitsteen op 'toetsvragen schrijven': nagaan of de ontworpen toetsvragen voldoen aan redelijke kwaliteitscriteria. De nadruk ligt daarbij op de controle voorafgaand aan de toetsafname, omdat het achteraf verwijderden van ondeugdelijk gebleken toetsvragen vervelende gevolgen heeft voor de studenten. De controle zelf wordt

uitgevoerd: de procedure die daarbij gevolgd wordt, te waarborgen dat verschillen in opvatting tussen verschillende examinatoren ook zullen blijken, schriftelijk vastgelegd worden, en niet in onderlinge gedachtenwisseling weggemasseerd worden. Dit is de praktische consequentie van de eis dat toetsvragen (en wat als juiste beantwoording geldt) de intersubjectieve overeenstemming van examinatoren moeten hebben.

Samenvattend: Hoofdstuk 2 is een algemeen hoofdstuk. De hoofdstukken 3, 4 en 5 geven aanwijzingen voor het schrijven van toetsvragen bij termen en relaties, de bouwstenen van alle leerstof. De hoofdstukken 6 en 7 bouwen daarop voort, waarbij hoofdstuk 6 speciaal van belang is voor toetsen over tekst en hoofdstuk 7 voor toetsen waarin problemen gesteld worden. Hoofdstuk 8 behandelt de altijd onmisbare controle op de kwaliteit van de ontworpen toetsvragen.

2 Vraagsoorten, doorzichtigheid, rompvragen en abstractieniveau

In dit hoofdstuk passeren enige grondregels voor het schrijven van toetsvragen de revue. Om te beginnen worden de vormen de vraagvormen besproken: opstelvragen, aanvulvragen en keuzevragen. Daaruit moet een weloverwogen keuze gedaan worden; welke vraagvorm voor welk doel geschikt is, hangt minder van algemene voor- en nadelen af dan van de aard van de leerstof en wat erover gevraagd wordt. Dat is vooral van belang bij overwegingen om keuzevragen te gebruiken. Enige flexibiliteit is bij het kiezen van een vraagvorm meestal wel mogelijk: er is geen bezwaar tegen om in dezelfde toets verschillende vraagvormen naast elkaar te gebruiken, mits dat maar overzichtelijk gebeurt.

De toets hoort nog steeds tot het onderwijs: het is van belang dat de manier waarop studenten zich kunnen voorbereiden op het afleggen van de toets, bijdraagt aan onderwijsdoelen. Het sleutelwoord is hier *doorzichtigheid*. Om die doorzichtigheid te vergroten, moet er vooral op het karakter van de toetsvragen gelet worden.

Een bijzonder aandachtspunt in dit boek vormt het abstractieniveau waarop de toetsvraag geformuleerd wordt. Wie te algemeen of te abstract geformuleerde vragen schrijft maakt het niet alleen zichzelf moeilijk steeds maar weer nieuwe vragen van dit slag te bedenken, maar berokkent ook het onderwijs schade. Praktisch wordt dit vertaald naar het gebruiken van rompvragen: vraagskeletten waarin telkens nieuwe (concrete) voorbeelden uit een gespecificeerde lijst ingevuld worden om 'nieuwe' vragen te maken. Meer principieel is het van belang om ook bij het schrijven van toetsvragen ervan uit te gaan dat stofbeheersing op een abstract niveau alleen maar betekenis kan hebben wanneer dat onderbouwd is met kennis van de concrete zaken die geabstraheerd zijn.

2.1 Aanvul- of invulvragen

VOORBEELD

Amerika werd ontdekt door

Het voorbeeld is een prototype van de *aanvulvraag*, wat niet wil zeggen dat het ook het meest fraaie voorbeeld is dat er te bedenken valt. Dezelfde vraag kan anders ingekleed worden tot een *invulvraag*:

VOORBEELD

Amerika werd door ontdekt.

Een *kort-antwoordvraag* wordt het door de vragende vorm te gebruiken :

VOORBEELD

Door wie werd Amerika ontdekt?

De eenvoud van deze vraagvormen, die ik in het vervolg samenvat onder de term 'aanvulvragen, is heel aantrekkelijk. Deze vragen zijn de directe uitdrukking van wat de docent graag van de student wil horen. Ook wie meerkeuzevragen gaat schrijven, doet er goed aan te beginnen met aanvulvragen te schrijven, om die daarna tot meerkeuzevragen uit te bouwen door geschikte alternatieven te construeren.

Het karakter van de aanvulvraag is dan wel de eenvoud, maar daar volgt nog niet uit dat eenvoudig gestelde vragen ook goede vragen zijn. Het gegeven voorbeeld illustreert dit op een niet eens al te overdreven wijze. Bedenk eens welke goede, maar mogelijk niet bedoelde antwoorden er op deze vraag zouden kunnen komen:

- een Italiaans kapitein - zeevaarders
- de Spanjaarden - de Chinezen
- de Vikingen - de Feniciërs
- Bjarni of Leif

Zoals de aanvulvraag is gesteld, sluit hij geen van de 'onbedoelde' antwoorden uit. Het spreekt bijna vanzelf dat onbedoelde antwoorden die niet evident fout zijn, volledig goed gerekend moeten worden. Dat kan weleens problemen geven, bijvoorbeeld wanneer de beoordelaar zich niet realiseert dat Columbus een Italiaan was, of wanneer de antwoorden twijfelachtig worden, zoals in het voorbeeld het antwoord 'zeevaarders.' Scherper formuleren van de vraag kan dergelijke verwikkelingen voorkomen:

VOORBEELD

De man die Amerika in 1492 ontdekte, heette

Nu we toch nog even met onverwachte antwoorden bezig zijn: er zit een zeker cultureel vooroordeel in de Columbusvraag, en het is niet uitgesloten dat sommige studenten erop antwoorden met een uiteenzetting dat Columbus een land 'ontdekte' dat al bewoond was.

Aanvulvragen zijn bij de meeste leerstof of onderwijsdoelen te gebruiken. Het volgende lijstje voorbeelden geeft daar een indruk van, ook al heb ik daar geen vragen in opgenomen die een meer uitvoerige vraagstelling hebben.

VOORBEELDEN

- Een bal valt van een 28 meter hoog balkon; hoe lang duurt zijn val?
.....
- Wat is de vergelijking van de raaklijn in het punt $x = 2$ aan de kromme met de vergelijking $y = 2x^2 + 5x - 1$?
- Wat is het meest voorkomende mineraal in de aardkorst?
- Geef de formule voor de tweede bewegingswet van Newton
- De vier meest voorkomende stikstofbasen van DNA zijn,, en

Aanvulvragen zijn er in soorten. De Columbusvraag is een voorbeeld van het quiz-achtige type: een korte vraag, een antwoord dat bestaat uit een enkel woord of een naam, en een vrijwel onmiddellijk te beantwoorden vraag (voor de student die het antwoord weet, tenminste). Rekenopgaven vragen meer tijd ter beantwoording: het antwoord moet berekend, geconstrueerd of afgeleid worden. Wie daarbij niet wil toetsen of de student wel kan rekenen, maar of de student de juiste formule of wet weet te kiezen, kan daarmee rekening houden door 'gemakkelijke' getallen te kiezen of in plaats van een rekenopgave beter rechtstreeks te vragen naar de formule die voor de berekening van de opgave gebruikt zou kunnen worden (en de berekening zelf achterwege laten). Vervolgens zijn er vragen die heel uitgebreid gesteld worden, omdat de probleemsituatie een uitgebreide beschrijving nodig heeft, en die

toch heel kort te beantwoorden zijn. Dergelijke vragen kosten veel leestijd, waar soms aan tegemoet gekomen kan worden door meerdere aanvulvragen te stellen bij dezelfde beschreven situatie. Tenslotte zijn er aanvulvragen die een wat langer antwoord vereisen: een definitie, een opsomming van zaken, een schets, een bewijsvoering, niet alleen het antwoord van een berekening, maar ook de berekening zelf, een samenvatting. Kortom, er is een geleidelijke overgang van aanvulvragen naar opstelvragen, en het is niet van belang daar een precieze grens in aan te geven.

Het nakijken of scoren van aanvulvragen hoeft geen tijdrovende bezigheid te zijn, zeker niet bij vragen die met een enkel getal, een naam of een woord beantwoord moeten worden. Wanneer voor het toetsformulier een handige lay-out wordt gekozen, waarbij de antwoorden ongeveer in één kolom aan de rechterkant komen, kan een scoringsjabloon gemaakt worden dat gewoon bestaat uit een correct beantwoord toetsformulier. Ik vermeld dit hier, omdat bij de vraagformulering daar ook al rekening mee gehouden kan worden: liever geen invulvragen, de vraag zodanig formuleren dat een zo kort mogelijk antwoord gegeven kan worden. Bij gebruik van doordrukantwoordformulieren kan het scoren eventueel ook aan de studenten worden overgelaten. In dat geval kan de docent volstaan met steekproefsgewijs daar controle op uit te oefenen, en in ieder geval alle formulieren met een score vlak bij de grens zakken-slagen te controleren. Bij aanvulvragen moet er altijd op gelet worden of de vraag voldoende duidelijk maakt welk soort antwoord er verlangd wordt, zoals de Columbusvraag al demonstreerde. Het is vaak maar al te gemakkelijk om een opgave of vraag anders te interpreteren dan de vragensteller bedoeld had. Marten Toonder is een meester in het uitwerken van deze thematiek in zijn Bommelstrips. Wie wel eens geprobeerd heeft om een computerprogramma te schrijven weet hoe lastig het kan zijn volstrekt eenduidige opdrachten te formuleren. Voor het communiceren met computers kan ik verwijzen naar het werk van Belnap en Steel (1976), *The logic of questions and answers*. Nu is zo'n computer maar een domme machine en zo'n ding is heel wat moeilijker te hanteren dan een intelligente student. In het onderwijs dat aan de toets voorafgaat, heeft de student kennis gemaakt met het soort vragen dat hij in de toets mag

verwachten, en welk soort antwoorden daarop verwacht wordt. Anders gezegd: in de formulering van de aanvulvraag moet slechts dan heel precies aangegeven worden welk soort antwoord er van de student wordt verwacht, wanneer dat niet impliciet duidelijk is uit het gegeven onderwijs.

Er valt verder over aanvulvragen niet veel te vertellen, en dat kan als een voordeel van deze vraagvorm beschouwd worden: het gebruik ervan kent weinig problematische kanten. Er wordt wel eens gezegd dat het onderscheid tussen aanvulvragen en meerkeuzevragen is dat bij aanvulvragen de student het antwoord moet produceren, terwijl hij het bij meerkeuzevragen moet herkennen. Dat is een wat te simpel beeld, denk bijvoorbeeld aan wiskundeopgaven in meerkeuzevorm, waar bij het antwoord gewoon berekend zal moeten worden. (Zie voor dat veronderstelde verschil tussen 'herkennen' en 'herinneren' ook Langerak, 1979).

2.2 Keuzevragen

Bij keuzevragen 'kiest' de student zijn antwoord uit een of meer lijstjes aangeboden alternatieven. De meest eenvoudige keuzevraag is de ja/nee-vraag, op de voet gevolgd door de tweekeuzevraag. De vierkeuzevraag wordt veel gebruikt, soms ook de vijfkeuzevraag. Er kan gevraagd worden het *beste* uit de gegeven alternatieven aan te kruisen; in de exacte vakken zal doorgaans gevraagd worden het *juiste* alternatief aan te kruisen. Het is een wijdverbreide gewoonte om vragen te schrijven met één juist alternatief, maar er is veel voor te zeggen om voor daartoe geschikte leerstof de mogelijkheid te gebruiken dat *meerdere* alternatieven juist zijn, of geen enkele (de toetsinstructie moet dat duidelijk aangeven). Een typisch voorbeeld van een meerkeuzevraag:

VOORBEELD

Welk dier legt eieren in het water?

1. slang
2. kikker
3. schildpad
4. walvis
5. zwaan

Het voorbeeld laat ook een typische tekortkoming van talrijke meerkeuzevragen zien: wanneer het de bedoeling is na te gaan of de leerling weet dat kikkers eieren in het water leggen, is het niet duidelijk wat die andere dieren voor rol spelen in deze keuzevraag. Het ei van Columbus zou hier toch zijn om een gewone aanvulvraag te stellen:

VOORBEELD

Waar legt een kikker eieren?

Keuzevragen zijn op te vatten als een speciaal soort aanvulvragen:

VOORBEELD

Legt een kikker eieren in het water of op het land?

Dit illustreert dat keuzevragen minder algemeen bruikbaar zijn dan aanvulvragen, hoewel er ook onderwerpen zijn waar nu juist de keuzevraag goed bij gebruikt kan worden. Voor de schrijver van keuzevragen zit het probleem in het bedenken van de aan te bieden alternatieven, en dan vooral de afleiders. Neem bijvoorbeeld deze aanvulvraag:

VOORBEELD

Los x op uit $x/3 = x - 1$ $x = \dots\dots\dots$

Dit is hetzelfde voorbeeld dat Timmer (in de 1969 editie van De Groot en Van Naerssen, hoofdstuk 10) gebruikt om er de problemen van het schrijven van meerkeuzevragen voor wiskunde aan te demonstreren. Het juiste antwoord is $3/2$. Wil je een vierkeuzevraag maken, dan heb je drie alternatieven nodig die *fout* zijn. Die alternatieven worden met een ongelukkig suggestieve term wel *afleiders* genoemd, en ik zal dat in het volgende ook doen. Het is duidelijk dat de afleiders in ieder geval *getallen* moeten zijn, maar welke? Neem je een paar willekeurige getallen, dan zijn die al gauw herkenbaar als afleider. Getallen die lijken op het juiste antwoord, zoals -1 , $1/4$, en $3/4$, zijn bruikbaar: de student die de opgave niet kan berekenen, zal dan moeten raden. Maar je kunt ook getallen nemen die het resultaat zijn van bepaalde fouten bij het berekenen: uit de foute stap $x = 3x - 1$ volgt

bijvoorbeeld het antwoord $x = 1/2$. Een keuzevraag met dergelijke afleiders toetst mede of de student geleerd heeft de daaraan corresponderende fouten te vermijden; deze afleiders kunnen alleen gebruikt worden wanneer in het onderwijs het vermijden van deze fouten is behandeld.

Welke getallen ook als afleiders gekozen worden, het blijft bij deze keuzevragen altijd mogelijk om het juiste antwoord op averechtse wijze te vinden door vanuit de genoemde alternatieven *terug te redeneren of te rekenen*. Dat probleem is op gekunstelde wijze te ondervangen door als alternatieven niet bepaalde getallen te nemen, maar intervallen waarbinnen het juiste antwoord wel of niet ligt. Ik raad dergelijke constructies af, omdat de wijze van vragen stellen daardoor nog een stap verder af komt te staan van de werkelijkheid waarover het onderwijs verondersteld wordt te gaan. Anders gezegd: een toets is geen quiz. Deze truc met intervallen als alternatieven komt niet alleen voor bij opgeven waar getallen als antwoord bij horen. De volgende afrader laat dat zien, en illustreert ook het vervreemdende effect van deze wijze van vragen stellen.

AFRADER

De naam van Socrates' beroemdste discipel begint met de letter:

1. A t/m E
2. F t/m J
3. K t/m O
4. P t/m T
5. U t/m Z

(Wesman, in: Thorndike, 1971).

De afleiders spelen bij keuzevragen dus een heel belangrijke rol: goede meerkeuzevragen zijn alleen dan mogelijk wanneer de afleiders op tenminste enigszins 'natuurlijke' wijze bij de vraagstelling passen. Wanneer je mag veronderstellen dat de student bij het beantwoorden van een aanvulvraag kiest uit een klein aantal mogelijkheden, dan kun je diezelfde mogelijkheden natuurlijk ook gebruiken als alternatieven voor dezelfde vraag in meerkeuzevorm. Ik geef een aantal simpele voorbeelden van aanvulvragen die zich tot een dergelijke natuurlijke omzetting tot meerkeuzevragen lenen:

VOORBEELDEN

- Welk waddeneiland was aan het eind van WO II het toneel van een dramatische strijd van Georgiërs tegen Duitsers?
- Welk kabinet nam het besluit duikboten aan Taiwan te leveren?.....
- Wat is de hoofdstad van Gelderland?
- Noem een reden waarom walvissen als zoogdieren geïncubateerd worden
- Strontium-90 is een radioactieve stof die bij proeven met waterstofbommen vrijkomt. Het beendergestel neemt het op omdat het er welke andere stof vervangt?

Voor ieder van de genoemde aanvulvragen geldt dat de student die tenminste een minimale kennis van zaken heeft, in ieder geval weet uit welke mogelijkheden er gekozen moet worden. Deze student krijgt geen informatie cadeau wanneer er meerkeuzevragen van gemaakt worden door deze 'natuurlijke' mogelijkheden als alternatieven te nemen. Het voordeel van meerkeuzevragen is in deze gevallen minimaal, en bestaat alleen uit de mogelijkheid om deze vragen door een computer te laten verwerken. Daar staat als nadeel tegenover dat de student, die *niets* van het onderwerp weet, nu een raadkans krijgt, die veel groter is dan bij de aanvullende vraagvorm.

Dit soort 'natuurlijke' meerkeuzevragen kan vragen omvatten die het als aanvulvraag niet zouden doen omdat het omschrijven van het verlangde antwoord te lastig of te omslachtig is: de genoemde alternatieven geven tevens aan welk soort antwoord verlangd wordt. Door de aangeboden alternatieven wordt de zoekruimte van de student ingeperkt, en dat kan soms wenselijk zijn. Dergelijke meerkeuzevragen zijn misschien minder geschikt voor eindtoetsen, maar kunnen in tussentijdse toetsen en als oefenopgaven goede diensten bewijzen.

Wanneer de student moet kunnen *onderscheiden* tussen bepaalde dingen, gebeurtenissen enz. komt de meerkeuzevraagvorm ook op natuurlijke wijze in aanmerking. Bijvoorbeeld wanneer de leerling bepaalde dieren moet kunnen classificeren als wel of niet tot de zoogdieren behorend, ligt het voor de hand om een lijstje van die dieren als alternatieven te ne-

men. Dan is het ook natuurlijker om het aantal 'juiste' alternatieven vrij te laten: minimaal geen, maximaal alle alternatieven zijn zoogdieren. Wanneer dat tot verwarring aanleiding zou kunnen geven omdat in een toets de meeste vragen van het type 'het beste alternatief' of 'het juiste alternatief' zijn, kan de meerkeuzevraag gepresenteerd worden als een *meervoudige ja/nee-vraag*.

VOORBEELD

Is dit een zoogdier?

1. struisvogel ja/nee
2. walvis ja/nee
3. veldmuis ja/nee
4. vleermuis ja/nee
5. pinguïn ja/nee

Meerkeuzevragen zijn dus geschikt voor opgaven waar de student op een bepaald onderscheidingsvermogen getoetst wordt. Het zal geen verbazing wekken dat het gevaar van het lichtvaardig gebruiken van meerkeuzevragen in het omgekeerde geval ligt: vragen waarbij het helemaal niet om onderscheidingsvermogen gaat, kunnen door een ondoordachte set afleiders dat karakter krijgen, en daarmee in feite een 'vaardigheid' toetsen die in het geheel niet tot de onderwijsdoelen behoort. Bedenk dat een student die zo'n vraag in de aanvulvorm zonder meer goed beantwoord zou hebben, dezelfde vraag in meerkeuzevorm fout kan gaan maken omdat hij door ongelukkige afleiders op het verkeerde spoor wordt gebracht. Dat kan gebeuren omdat het zelf bedenken van een juist antwoord nog niet uitsluit dat een van de andere alternatieven een nog beter antwoord zou zijn; en de spelregels van de meerkeuzetoets verlangen dan het aanstrepen van het nog betere antwoord. De student die een meerkeuzevraag beantwoordt zal doorgaans alle alternatieven op hun mogelijke juistheid nagaan. Wanneer er afleiders gekozen worden die niet in het onderwijs aan de orde zijn geweest, komen studenten in de problemen en verliest de toets aan geldigheid.

Dan moeten nu nog de matchingvraag en de ja/nee-vraag besproken worden. De *matchingvraag* is mogelijk bruikbaar in al die gevallen waarin dingen enz. *in paren* komen en waarbij

van de student het nodige onderscheidingsvermogen met betrekking tot juiste en onjuiste combinaties verwacht mag worden.

VOORBEELD

Schrijf het nummer van de titel voor de naam van de auteur ervan.

1. Het vijfde zegel Boon
2. De Kapellekensbaan Hermans
3. De tranen der acacia's Mulisch
4. Archibald Strohalm Vestdijk
5. Het Geuzenboek
6. Onder professoren

Bij *ja/nee-vragen* is het zaak scherp op te letten of bij de gegeven uitspraak wel zo'n absoluut oordeel als 'ja' of 'nee', dan wel 'juist' of 'onjuist' verlangd mag worden. Er zijn maar weinig korte uitspraken die ongeclausuleerd 'juist' of 'onjuist' zijn, en zou je zo'n uitspraak kiezen als ja/nee-vraag, dan komen al die studenten met de vraag in moeilijkheden, die condities en omstandigheden kunnen bedenken, die de uitspraak in zijn juistheid of onjuistheid kunnen aantasten. Die studenten moeten namelijk ook antwoord zien te vinden op de vraag of de vragensteller diezelfde condities of omstandigheden in gedachten had, of misschien niet.

VOORBEELD (goed)

Is een walvis een zoogdier? Juist/onjuist

AFRADER

De aarde beschrijft een cirkelvormige baan om de zon. Ja/nee

Een ander probleem bij ja/nee-vragen kan erin liggen dat het niet altijd even gemakkelijk is om 'goede' uitspraken te bedenken die onjuist zijn en waarbij het zo is dat uit de onderwijsdoelen duidelijk is dat de student deze uitspraken als onjuist moet kunnen herkennen. En het gebruik van ja/nee-vragen impliceert dat ongeveer de helft ervan als correct antwoord 'nee' moeten hebben (anders zouden onvoorbereide studenten een hoge score kunnen behalen door alle vragen met 'ja' te beantwoorden).

Tenslotte: weersta de verleiding om zinnen uit het studieboek te knippen, en daar ja/nee achter te plakken. Dat heeft niets met behoorlijk toetsen te maken.

Niet onvermeld mag blijven dat studenten 'gevoelsmatige' bezwaren tegen keuzevragen kunnen hebben. Dat kan een teken zijn dat keuzevragen misschien minder goed passen bij deze leerstof. Hoe dat ook zij, het *feit* dat studenten gevoelsmatige bezwaren hebben, is op zich reden genoeg om te zoeken naar toetsvragen die geen negatieve gevoelens oproepen.

Wanneer er weinig tijd beschikbaar is voor het nakijken, overweeg dan of studenten het werk zelf na kunnen kijken, met steekproefsgewijze controle daarop (van sommige studenten het hele werk nakijken, van allen enkele vragen nakijken).

Beperk meerkeuzevragen tot drie (en soms tot twee) alternatieven. Het schrijven van meerkeuzevragen gaat niet altijd even gemakkelijk. Vooral het bedenken van de tweede en derde afleider levert nogal eens moeite op; ofwel omdat de vraag eigenlijk maar één zinnige afleider toestaat, of omdat het lastig is alle afleiders voor de onwetende student even aantrekkelijk te maken. Bovendien doorzien studenten die moeizaam bedachte afleiders al gauw. De aanbeveling is simpel: gebruik die afleiders niet, maak dan liever drie- of tweekeuzevragen. Wanneer uit de vragenanalyse blijkt dat de meeste vragen wel één afleider hebben die door vrijwel niemand is aangekruist, dan functioneert die derde afleider niet en kan hij beter weggelaten worden. Hetzelfde kan gelden voor de tweede afleider wanneer ook die maar door een enkele student gekozen wordt. Weglaten van die afleiders levert geen verlies van toetskwaliteit op. Afleiders weglaten die toch niet functioneren, verandert immers de eigenschappen van de toets niet nadelig. Heeft dit gevolgen voor de zak/slaag-grens? Nee, omdat die derde (of tweede) afleider toch niet functioneert. Het is weliswaar bij tweekeuzevragen gemakkelijker om door raden aan een score te komen die enigszins boven de 50% ligt, maar dat heeft geen speciale betekenis. Dat zo'n derde afleider door bijna niemand aangestreept wordt, is een aanwijzing dat geheel onwetende studenten niet aan het tentamen deelnemen. Voor de scores van de studenten verandert de reductie in het

aantal alternatieven niets.

Nu zijn er ook vierkeuzevragen waar alle afleiders wel in redelijke mate functioneren. Is het aan te raden om ook dan op driekeuzevragen over te gaan? Dat zou het bedenken van vragen immers erg vereenvoudigen. Het antwoord is bevestigend, en de reden daarvoor ligt in de relatieve efficiëntie van driekeuzevragen. De raadkans is wel iets groter dan bij vierkeuzevragen, maar dat wordt gecorrigeerd door meer vragen (ongeveer 20%) in de toets op te nemen. Omdat studenten telkens één alternatief minder te bestuderen hebben, kunnen in dezelfde tijd meer keuzevragen beantwoord worden. Bijkomend voordeel van een groter aantal vragen is de betere dekking van de stof die zo verkregen wordt. Kan ik met mijn twee- of driekeuzevragen gebruik maken van hetzelfde computerprogramma dat toetsen met vierkeuzevragen analyseert? Ja, dat kan. Dezelfde antwoordformulieren kunnen ook gebruikt worden, al is het eleganter om studenten een antwoordformulier voor te leggen dat in aantal alternatieven per vraag overeenstemt met het toetsformulier. Er is geen dwingende reden om altijd tot vier alternatieven per vraag te gaan. In het bijzonder moet hier afgeraden worden om daartoe vragen van het volgende type te gebruiken.

AFRADER

- A. Meerkeuzevragen moeten bij voorkeur vier alternatieven hebben.
 - B. Een computerprogramma voor vierkeuzevragen kan geen driekeuzevragen verwerken.
1. A is juist en B is juist
 2. A is juist en B is onjuist
 3. A is onjuist en B is juist
 4. A is onjuist en B is onjuist

Maak van zo'n vraag gewoon twee afzonderlijke vragen, in plaats van één pseudo-vierkeuzevraag. Uitspraak B is onjuist, het computerprogramma kan deze tweekeuzevragen zonder meer verwerken. Bedenk ook dat deze afrader ertoe kan leiden dat studenten zich vergissen bij het aankruisen van het juiste alternatief: dat gegoochel met vier gekunstelde alternatieven kan net iets meer koelbloedigheid vragen dan in de tentamensituatie opgebracht kan worden.

Raadkansen. Een belangrijk nadeel van (meer)keuzevragen is de invloed van raadkansen op de toetsresultaten. Nu bestaan er formules waarmee je toetsresultaten zou kunnen 'corrigeren voor raden'. Dat is noodzakelijkerwijs altijd een correctie in statistische zin, en zo'n correctie is alleen zinvol voor wie geïnteresseerd is in de toetsresultaten voor een hele *groep* studenten. Bijvoorbeeld bij de analyse van toetsresultaten voor de evaluatie van het onderwijs ligt het in de rede om dat deel van het resultaat dat op raden berust, af te trekken. De score voor een individuele student kun je op deze wijze ook wel corrigeren, maar dit is veel minder zinvol, en lokt het misverstand uit als zou je daarmee corrigeren voor het effect van het raden dat *deze* student gedaan heeft. Het is nu eenmaal niet mogelijk om in deze zin individuele scores voor raden te corrigeren: aan de toetsresultaten is niet af te lezen op welke vragen de student geraden heeft, zelfs niet op hoeveel de student geraden heeft, en al helemaal niet hoeveel vragen de student 'goed' geraden zou hebben. Ik besteed hier zoveel aandacht aan omdat het raden een negatieve invloed heeft op de kwaliteit van de toetsing. Door het raden wordt er immers in de toetsresultaten een extra toevalselement ingebracht, een toevalselement dat bij aanvulvragen *niet* bestaat (tenzij ook dan de student uit een klein aantal alternatieven kiest).

Objectiviteit. Keuzevragen worden wel aanbevolen omdat ze *objectief* zouden zijn. Daarmee wordt bedoeld dat het bij deze vragen ondubbelzinnig vaststaat welke antwoorden goed en welke fout gerekend worden. Op triviale wijze is dat inderdaad zo: voor het nakijken van keuzevragen wordt doorgaans een *scoringsleutel* opgesteld die aangeeft welk alternatief bij iedere vraag als het juiste antwoord gescoord moet worden. Is de sleutel eenmaal vastgesteld, dan kan de computer het verdere werk opknappen. Bij het nakijken kan het dan niet meer gebeuren dat de ene beoordelaar een alternatief als juist aanmerkt terwijl een andere beoordelaar hetzelfde alternatief als onjuist scoort. Meningsverschillen tussen de ene en de andere beoordelaar spelen zodoende *bij het nakijken* geen rol meer. Op een ander moment spelen die verschillen echter wel degelijk een rol: *bij het vaststellen van de scoringsleutel*. Dan kan wel degelijk blijken dat de ene docent een afwijkend oordeel heeft over wat het 'beste' alternatief is, of over de juistheid

van wat een andere docent als 'afleider' ziet. Blijkt dat, dan wordt zo'n keuzevraag natuurlijk verbeterd of weggegooid. In de praktijk gaat het vaak anders: ofwel wordt in zo'n geval gewoon *afgesproken* hoe deze vraag gescoord zal worden, ofwel de keuzevragen zijn niet eens door meer dan één docent kritisch bekeken, zodat helemaal niet bekend is of er wel overeenstemming bestaat over wat het beste alternatief is en over de onjuistheid van als onjuist bedoelde afleiders. In beide gevallen zijn het de studenten die er het slachtoffer van worden, en wel op dezelfde wijze als zij het slachtoffer kunnen worden van het enigszins subjectieve oordeel van de beoordelaar van aanvulvragen.

Keuzevragen zijn in principe even gevoelig voor het subjectieve oordeel van de docent als aanvulvragen dat zijn, en daar staat de computer helemaal buiten.

De enige remedie is: door goede kwaliteitscontrole vooraf op de concept-toetsvragen, inclusief wat daarbij gedacht wordt dat 'juiste' antwoorden op de vragen zijn, zo veel mogelijk subjectieve momenten proberen uit te sluiten (zie hfdst. 8).

Degenen die keuzevragen propageren vanwege hun objectiviteit, omschrijven dat in termen van de uiterlijke kenmerken van de vragen, niet van de inhoudelijke kwaliteit. Met andere woorden: de objectiviteit van degene die de alternatieven formuleert en de scoringsleutel vaststelt, is niet gegarandeerd.

In de vergelijking met aanvulvragen zijn keuzevragen dus ongeveer even objectief. En inderdaad worden in de Amerikaanse literatuur zowel aanvulvragen als keuzevragen tot de objectieve vragen gerekend. De tegenstelling is dan ook niet zozeer die tussen aanvulvragen en keuzevragen, als wel die tussen *objectieve vragen* en opstelvragen (zie 2.3). Met name het pleidooi van De Groot (in: De Groot en Van Naerssen) voor het gebruik van keuzevragen kan (mijns inziens: moet) gelezen worden als een pleidooi voor het gebruik van objectieve vragen: aanvulvragen of keuzevragen. In zijn *Methodologie* heeft De Groot (1961, blz. 239 e.v.) een overzicht gegeven van de talrijke onbedoelde effecten die op kunnen treden bij het nakijken van opstelvragen. Wanneer opstelvragen goed vervangen kunnen worden door objectieve vragen, dan is het voordeel van de grotere objectiviteit meestal van doorslaggevend belang. Overigens, wanneer opstelvragen niet goed te vervangen zijn door objectieve vragen, dan kan

door enkele eenvoudige maatregelen (zie hoofdstuk 8) het nakijken ervan verregaand 'geobjectiveerd' worden.

Samenvatting. Door hun speciale vorm zijn keuzevragen heel geschikt voor het toetsen van speciale vormen van stofbeheersing; in dit boek zullen daar talrijke voorbeelden van gegeven worden. Wanneer keuzevragen niet op 'natuurlijke' wijze bij de leerstof passen, brengt het gebruik van deze vraagvorm enkele nadelen met zich mee, die afgewogen moeten worden tegen het enige voordeel dat deze vraagvorm in die situaties heeft: automatische scorbaarheid. Is dit een eerlijke presentatie van de voor- en nadelen van keuzevragen? Mijns inziens wel. Twee dingen zijn daarbij te bedenken, aan de objectieve scorbaarheid, en vooral de automatische scorbaarheid en bijpassende computerverwerking die in de jaren '50 mogelijk werd, heeft de keuzevraag zijn populariteit en ruime toepassing te danken; en die scorbaarheid was vooral van belang voor de *gestandaardiseerde* toets, die vrijwel gelijktijdig met de gestandaardiseerde psychologische test op grote schaal toepassing vond. De tweede opmerking is dat 'objectiviteit' geen specifiek voordeel van keuzevragen is, zoals hierboven besproken.

In 8.2 worden aanwijzingen gegeven die van belang zijn bij het formuleren van keuzevragen, resp. het controleren van die formulering.

2.3 Opstelvragen

Opstelvragen zijn vragen die een korte uiteenzetting, een verklaring, een opsomming, een schets, een samenvatting en dergelijke vragen. Kortom, opstelvragen worden typisch in een of meerdere zinnen beantwoord, terwijl aanvulvragen een antwoord van een of enkele woorden eisen. De overgang tussen beide vraagvormen is uiteraard niet scherper dan dat af te bakenen. De opstelvraag biedt de student enige vrijheid in de wijze van formuleren, en inderdaad kan die wijze van formuleren zelf in *sommige gevallen* ook ter beoordeling staan. Er is doorgaans niet alleen vrijheid in de wijze van formuleren, maar ook in de keuze van wat in het antwoord vermeld wordt.

Het is duidelijk dat deze vrijheden het risico met zich mee-

brengen dat er in de beoordeling van de antwoorden meer ruimte is om subjectieve verschillen tussen beoordelaars tot uiting te laten komen dan dat bij de objectieve vraagvormen - aanvul- en keuzevragen - het geval is. Naast de opstelvragen hebben ook het *opstel* en het *werkstuk* dezelfde kenmerken van vrijheid in de beantwoording en kwetsbaarheid voor subjectieve beoordeling. In dit boek hebben uitspraken over opstelvragen dan ook evenzeer betrekking op deze beide, verder niet afzonderlijk besproken, vraagvormen.

Opstelvragen zijn geschikt om te toetsen op zaken als: helderheid van formuleren, compositie of structuur in antwoord of werkstuk, nauwkeurigheid, oorspronkelijkheid, diepgang van analyse, en dergelijke. Maar het is niet *vanzelfsprekend* dat antwoorden op deze zaken beoordeeld worden. Daarvoor moeten tevoren met de studenten afspraken gemaakt worden (*doorzichtigheid*, zie 2.4), terwijl ook de beoordelaars daartoe goed geïnstrueerd moeten worden. Het blijkt immers telkens weer dat de ene docent sterk kan verschillen met de ander wat betreft de zaken waarop hij of zij speciaal let bij het beoordelen; daarom moeten daar van tevoren duidelijke afspraken over gemaakt worden. Daar komt bij dat op genoemde zaken alleen dan beoordeeld mag worden, wanneer zij tot de onderwijsdoelen behoren (zie ook 8.1).

Het behoeft geen uitleg dat juist dit soort opstelvragen zich niet lenen tot vervanging door objectieve vraagvormen. Dat betekent, wanneer de zo te toetsen vaardigheden belangrijke onderwijsdoelen zijn, dat het tijdrovende nakijken weliswaar een nadeel van deze opstelvragen is, maar een onvermijdelijk nadeel: er is geen goed alternatief voorhanden, en daarom mag deze prijs betaald worden.

De opstelvraag is ook de vraagvorm bij uitstek voor de categorie vragen die de student voor een op te lossen *probleem* stellen: zie hoofdstuk 7 dat daaraan gewijd is.

Dat hoofdstuk is een goede illustratie van de wijze waarop het mogelijk is om *structuur* aan te brengen in een bepaalde klasse van opstelvragen, een structuur die het mogelijk maakt om heel precies aan te geven wat er getoetst wordt, en waarop antwoorden beoordeeld kunnen worden. Diezelfde structurering maakt het ook mogelijk om omvangrijke opstelvra-

gen terug te brengen tot kleinere opstelvragen, in sommige gevallen ook tot objectieve vragen.

Er wordt vaak gevraagd of opstelvragen hetzelfde ‘meten’ als objectieve vragen. Onderzoek daarnaar levert geen eenduidige resultaten op (Mellenbergh, 1971). De vraag is ook niet erg zinvol, want ofwel datgene wat via de opstelvraag getoetst wordt, laat zich niet in objectieve vragen toetsen (compositie, structuur enz.), ofwel de opstelvraag kan wel door objectieve vragen vervangen worden, en dan hebben die laatste zeker de voorkeur. Opstelvragen hebben het nadeel dat uitgebreide voorzorgen genomen moeten worden om een eerlijke beoordeling te waarborgen: tevoren moeten *modelantwoorden* opgesteld worden die als leidraad bij het nakijken moeten dienen, terwijl meestal ook een bijbehorend *beoordelingsvoorschrift* opgesteld moet worden, waarin precies aangegeven wordt hoe de puntentoekening dient te gebeuren. Het nakijken en beoordelen is tijdrovend, en daar komt dan nog bij dat ieder antwoord door tenminste twee beoordelaars onafhankelijk van elkaar beoordeeld moet worden om een minimale controle op onbedoelde verschillen tussen beoordelaars te hebben. Zie ook 8.1 en 8.3. Een ander nadeel van de opstelvraag is dat er in één toets veel minder vragen over de stof gesteld kunnen worden dan bij het gebruik van objectieve vragen. Dit is een nadeel omdat de toetsresultaten dan sterker verstoord worden door het feit dat de student over een enkel gevraagd onderwerp toevallig weinig of juist erg veel weet. Het is niet zo dat een toets met opstelvragen minder representatief voor de te toetsen leerstof is, want daarvoor is uitsluitend van belang dat ieder onderwerp in die leerstof een

Tabel 2.1. Beoordeling van tandheelkundige werkstukken door drie instructeurs

werkstuk nr.:	1	2	3	4	5	6	7	8	9	10	totaal v/d instructeur
instructeur 1	8	11	14	7	10	11	7	14	9	10	101
instructeur 2	8	14	9	9	11	14	12	9	9	12	107
instructeur 3	6	9	6	13	10	14	13	8	11	9	99
hoogste oordeel	8	14	14	13	11	14	3	14	11	12	
laagste oordeel	6	9	6	7	10	11	7	8	9	9	

even grote *kans* heeft in de toets opgenomen te worden als andere onderwerpen.

Houd bij de afweging van een en ander er rekening mee dat de genoemde nadelen van de opstelvraag sterker meewegen naarmate de vragen meer in de richting van het opstel gaan, terwijl ze minder worden wanneer de vragen meer het karakter van de aanvulvraag (kort-antwoordvraag) krijgen.

Tot slot geeft tabel 2.1 een illustratie van de mate waarin beoordelaars van elkaar kunnen verschillen, zelfs wanneer de beoordeling gebonden is aan een lijst van 15 criteria waarop beoordeeld moet worden (bron: Tromp, 1979, de gegevens van Tromp zijn uitgebreider dan de tabel kan laten zien).

2.4 Doorzichtigheid

De toets heeft een heel eigen plaats in het onderwijsgebeuren, maar wel er *binnen*, niet aan de zijlijn. Studenten besteden erg veel van hun tijd aan de voorbereiding op toetsen; hoe belangrijk is het dan niet dat ze goed weten welke soort vragen gesteld zullen worden, dat ze zich goed op dergelijke vragen kunnen voorbereiden, en dat ze zich tijdig een helder beeld kunnen vormen hoe het er met de eigen beheersing van de stof voorstaat? De Groot (1970) introduceerde voor dit complex van wenselijkheden de term *doorzichtigheid* (transparency).

Op verschillende manieren heeft de toetsvragenschrijver met deze doorzichtigheid te maken. Natuurlijk moeten afzonderlijke vragen doorzichtig zijn: niet alleen voorbereidbaar, maar ook begrijpelijk in hun relatie tot de onderwijsdoelen of de leerstof. Daarnaast vraagt doorzichtige toetsing dat studenten toegang hebben tot een tamelijk groot bestand van (oude) toetsvragen, en dat betekent dat het schrijven van nieuwe toetsvragen efficiënter zal moeten gebeuren dan voorheen. Tenslotte moet een doorzichtige toets representatief zijn voor de behandelde stof, en dat betekent dat er speciaal op gelet moet worden dat vragen over de *kern* van de stof, vragen die veel studenten daarom goed zullen weten te beantwoorden, niet ten onrechte uit de toets weggelaten worden omdat ze 'te gemakkelijk' zouden zijn.

Doorzichtigheid. Het is wenselijk dat toetsen doorzichtig zijn. Voor afzonderlijke toetsvragen zijn er twee betekenissen van 'doorzichtig' te onderscheiden. Is het duidelijk (1) wat de vraag toetst, en (2) hoe de student zich op de beantwoording heeft kunnen voorbereiden? Toetsvragen moeten in beide betekenissen doorzichtig zijn *voor studenten*; daarnaast is het gewenst dat ze ook voor collega-docenten doorzichtig zijn.

Is het duidelijk wat de vraag toetst? Het moet de student duidelijk zijn wat het onderwerp van de vraag is, en naar welke vorm van beheersing ervan gevraagd wordt. Daarvoor is het niet voldoende dat de vraag ondubbelzinnig gesteld is: het is mogelijk dat de student wel kan weten wat als een goed antwoord op de vraag zal gelden, terwijl het onduidelijk blijft wat het kunnen geven van dat goede antwoord precies met de bestudeerde stof te maken heeft. Ideaal zijn dermate doorzichtig geformuleerde vragen dat ook studenten die een fout antwoord geven, *weten* dat het gegeven antwoord fout is, en weten welk stukje kennis of inzicht hun ontbreekt. Ongelukkige afleiders bij meerkeuzevragen bedreigen deze doorzichtigheid.

Kan de student zich voorbereiden op vragen als deze? Wanneer het niet duidelijk is wat een vraag toetst, is het de student niet mogelijk zich gericht op zo'n vraag voor te bereiden. Andersom kunnen vragen doorzichtig zijn in de eerste betekenis van die term, en desondanks vrijwel onvoorbereidbaar zijn. Voor iedere toetsvraag moet gelden dat studenten weten dat zo'n *soort* vraag verwacht mag worden, en dat zij zich daar efficiënt op voor kunnen bereiden. Deze tweede betekenis kan ook omgekeerd uitgelegd worden: wanneer een slimme student die weinig van de stof af weet in het voordeel is boven de minder briljante student die de stof tot in de puntjes beheerst, dan is de vraag zeer ondoorzichtig en 'meet' hij voornamelijk het verschil in intelligentie.

Is het nu zo erg wanneer de student zich niet gericht kan voorbereiden op de soorten toetsvragen die in een afsluitende toets gebruikt worden? Ik meen van wel. Dergelijke vragen geven de student het idee dat er geen rechtstreeks verband is tussen de manier waarop hij de stof bestudeerd heeft, en de toetsvragen erover. De student kan zich niet anders op de toets voorbereiden dan zich domweg zoveel mogelijk van de stof in het hoofd te prenten. In zo'n ondoorzichtige situatie gaan studen-

ten zoeken naar mogelijkheden om op een efficiëntere manier het begeerde 'voldoende' toetsresultaat binnen te halen. Bijvoorbeeld is dat een minimale voorbereiding, met de gedachte: slaag ik de eerste keer niet, dan wel in de herkansing. Dit 'feed-forward'-mechanisme, de invloed van ondoorzichtig toetsen op het gedrag van de student, is ook op positieve wijze te zien. Want hetzelfde mechanisme kan uitgebuit worden door de toets doorzichtiger te maken. Dat spoort de student aan tot een studiestrategie die mikt op het in één keer slagen, omdat hij zich nu efficiënt inhoudelijk kan voorbereiden. Dit is een van de manieren om te bevorderen dat de onderwijsdoelen ook gerealiseerd worden.

Een doorzichtige toets is een toets waar de student zich efficiënt op kan voorbereiden. Bij die voorbereiding is het belangrijk dat de student kan oefenen op vragen van dezelfde soort als in de toets gebruikt zullen worden. En het gaat niet alleen om oefening: bovendien krijgt de student zo een scherper inzicht in het niveau van zijn eigen beheersing van de stof, en hoe zich dat verhoudt tot de eisen die de toets stelt. Ook dat is een aspect van de doorzichtigheid, de voorbereidbaarheid, van de toets: weten wanneer je de stof voldoende beheerst om met een gerust hart de toets af te kunnen leggen.

Bij dit alles is het telkens weer van belang dat de student kan beschikken over een redelijk uitgebreide verzameling (oude) toetsvragen. Helaas blokkeren sommige docenten die beschikbaarheid door gebruikte toetsvragen geheim te houden.

Vrijgeven van vragen na afloop van de toets. Bij een psychologietentamen in het najaar van 1978 nam een student de tentamenvragen mee, tegen het daartoe ingestelde verbod in. De examencommissie schortte daarop de tentamenuitslag voor deze student op. Het College van Bestuur (van de Universiteit van Amsterdam) sprak uit dat het opschorten van de tentamenuitslag volstrekt onaanvaardbaar is, en beraadde zich vervolgens over de meer algemene problematiek van het geheimhouden van tentamenvragen. De bezwaren tegen geheimhouding werden afgewogen tegen de bezwaren van vrijgeven. In deze afweging van principiële tegen praktische bezwaren kwam het College tot het besluit dat studenten de vragen na afloop van het tentamen mee mogen nemen. De Commissie

voor de Bestuurshervorming (commissie-Slagter) sprak zich in dezelfde zin uit (9 december 1980, advies aan de C-3 examencommissie van de subfaculteit geneeskunde van de Universiteit van Amsterdam).

Welke principiële argumenten liggen er ten grondslag aan de nadrukkelijke adviezen tentamenvragen vrij te geven? De Commissie voor de Bestuurshervorming: 'Met deze openbaarheid (van examens, Academisch Statuut art. 190 derde lid), die voor schriftelijke examens beperkt kan zijn tot openbaarheid van de stukken (achteraf), wordt bereikt dat vragen, antwoorden en beoordeling voor de belanghebbende controleerbaar zijn. Dit is niet alleen van betekenis voor de student die wil nagaan of hij juist is beoordeeld, bijvoorbeeld met het oog op een in te stellen procedure ex artikel 40 WUB (nu gewijzigd ingevolge de Wet tweefasestructuur wo). Het is naar de mening van de Commissie ook van onderwijskundig belang dat studenten kennis kunnen nemen van dit materiaal om een inzicht te krijgen in de omvang en zwaarte van de eisen die bij het examen worden gesteld, ten einde zich daarop zo goed mogelijk voor te bereiden. Een tweede principiële argument betreft de mogelijke ongewenste gevolgen van een geheimhoudingsbeleid. Bij belangrijke tentamens zullen studenten proberen geheime oude vragen te bemachtigen. Dat kan zelfs georganiseerd worden aangepakt door in onderlinge taakverdeling na afloop tentamen alle vragen te reconstrueren, uit te tikken, en openbaar te maken. Daar is niets op tegen, omdat immers de vragen voor iedereen gelijkelijk toegankelijk gemaakt worden. De docent doet er beter aan de vragen zelf vrij te geven dat voorkomt ook fouten bij het reconstrueren van de vragen.

Heel anders is het wanneer stiekeme vragencircuits ontstaan, waarbij studenten proberen bekende geraakte vragen voor medestudenten verborgen te houden om de eigen relatieve slaagkans voor het tentamen niet ongunstig te beïnvloeden. Dit soort competitieve sfeertjes komen inderdaad voor. Dan ontstaat er *ongelijkheid* tussen studenten qua informatie waarover ze kunnen beschikken, en dat maakt het tentamen zowel onrechtvaardig als onrechtmatig. Dat is ook het gevolg van handeltjes in toetsvragen, waar ze voor een rijksdaalder per stuk van de hand gaan om de eigen vakantiekas te spekken. Ga ervan uit dat studenten zullen proberen geheime vragen te kra-

ken wanneer u het belangrijk genoeg vindt om ze geheim te houden. En trek er de consequentie uit: dat het niet anders kan of de vragen moeten vrijgegeven worden, om de toets doorzichtig te houden.

Een en ander kan echter wel tot praktische problemen leiden: wanneer toetsvragen na afloop vrijgegeven worden, dan kan het niet anders of voor iedere nieuwe toets moeten telkens ook nieuwe vragen bedacht worden, althans veel meer dan voorheen gebruikelijk was. De reden waarom vragen tot nu toe wel geheim gehouden werden, was immers dat er herhaald gebruik gemaakt kon worden van die vragen.

Nu is dit boek juist geschreven om de docent uit deze verlegenheid te helpen, om haar of hem vuistregels te geven die het bedenken van vragen over de leerstof vergemakkelijken en systematiseren. Daarnaast geven deze en de beide volgende paragrafen enkele aanwijzingen waardoor het aantal *varianten* op een gegeven vraag vergroot kan worden, waardoor het ook mogelijk is om van al bekend gemaakte vragen nieuwe varianten te maken die als 'nieuwe' vragen in volgende toetsen bruikbaar zijn.

Bij het bedenken van toetsvragen bestaat de neiging om 'te gemakkelijke' vragen al bij voorbaat uit te sluiten, en anderzijds om vragen te ontwerpen die boven de opgegeven stof uitgaan als 'bonusvragen' voor 'betere' studenten.

Beide neigingen werken inefficiënt uit: het weglaten van gemakkelijke vragen omdat juist die vragen minder bedenktijd kosten, en het opnemen van bonusvragen omdat die vragen lastig te bedenken zijn. Daar komt bij dat weglaten van gemakkelijke vragen en opnemen van bonusvragen schade berokkent aan de representativiteit van de toets, en daarmee ook aan zijn doorzichtigheid.

Gemakkelijke vragen. Er zijn twee redenen die de toetsvragenschrijver ertoe brengen om gemakkelijke vragen al bij voorbaat uit te sluiten: de toetsliteratuur bevat vaak de aanbeveling om geen al te gemakkelijke vragen te gebruiken, en het lijkt ook overbodig om vragen te gebruiken die vrijwel iedereen goed zal maken. Toch is het zinvol om gemakkelijke vragen in de toets op te nemen, en er zal aangetoond worden dat het technische advies om gemakkelijke vragen te vermijden,

in zijn algemeenheid misleidend is.

Wanneer gemakkelijke vragen overeenstemmen met de onderwijsdoelen is het zeker legitiem om ze te gebruiken. Juist onderwerpen die tot de kern van de stof behoren, worden uitvoerig behandeld en studenten zullen die onderwerpen goed beheersen. Met andere woorden: het is niet onwaarschijnlijk dat vragen gemakkelijker zijn naarmate ze meer de kern van de stof betreffen of de belangrijker onderwijsdoelen bestrijken. Het omgekeerde verband is beter bekend: naarmate vragen meer over bijzaken, details en voetnoten gaan, worden ze moeilijker.

Weglaten van gemakkelijke vragen kan ertoe leiden dat belangrijke onderwijsdoelen niet in de toets terug te vinden zijn. Dat is niet alleen jammer voor de representativiteit van de toets, maar heeft ook tot gevolg dat studenten de aandacht op bijzaken richten in plaats van op hoofdzaken. Studenten verwaarlozen de kern van de stof ten gunste van de details en de voetnoten, waardoor een goede beheersing van de kern van de stof niet meer vanzelfsprekend is. Neem daarom ook die vragen in de toets op die iedere behoorlijk voorbereide student goed zal beantwoorden en die onmisbaar zijn in iedere toets die de behandelde stof moet weerspiegelen. Bijkomend voordeel van zo'n meer representatieve toets is dat ze ook een beter beeld geeft van wat er met het onderwijs bereikt is.

Omdat vragen die tot de kern van de stof behoren, makkelijker te bedenken zijn, kan het voorbereiden van die toets efficiënter worden door die vragen niet meer uit te sluiten. Niet alle vragen die de kern van de stof raken, zullen voor de student gemakkelijk zijn, maar voor zover ze dat wel zijn, vraagt het beantwoorden ook minder tijd. Ook dat is winst, omdat in dezelfde tijd meer vragen gesteld kunnen worden, waar door zowel een betere dekking van de stof wordt verkregen als een kwalitatief betere toets. Immers, een groter aantal vragen vermindert de invloed van toevalligheden en maakt het toetsresultaat voor de student beter voorspelbaar (de toets wordt daarmee doorzichtiger).

Gemakkelijke vragen die tot de kern van de stof behoren, zullen goed onderscheiden tussen studenten die het onderwijs behoorlijk gevolgd hebben, en studenten die dat niet gedaan hebben en toch aan de toets deelnemen. Men mag verwachten dat de kern van de stof niet zo vanzelfsprekend is dat iedere

buitenstaander vragen daarover goed zou kunnen beantwoorden. Voor de duidelijkheid: wanneer bijna alle studenten deze vragen goed maken, mag dat niet uitgelegd worden als zouden deze vragen niet onderscheiden tussen beide categorieën studenten. De meest waarschijnlijke uitleg is immers dat bijna alle studenten de kern van de stof beheersen, en in die zin het onderwijs goed gevolgd hebben. Wie in een bepaald geval twijfelt aan de juistheid van deze uitleg, kan de proef op de som nemen, en de gemakkelijke vragen over de kern van de stof voorleggen aan studenten die het onderwijs nog niet gevolgd hebben.

Bonusvragen. De aanbeveling om gemakkelijke vragen niet te schuwen, impliceert niet dat moeilijke vragen extra argwanend bekeken moeten worden. Ook voor moeilijke vragen geldt dat ze een plaats in de toets verdienen wanneer ze volgen uit de onderwijsdoelen, en zeker wanneer ze de kern van de stof raken. Niet alle vragen over de kern van de stof behoeven gemakkelijk te zijn, ook niet wanneer er in het onderwijs erg veel aandacht aan is besteed. Sommige onderwerpen zijn nu eenmaal moeilijk. Goede voorbereiding en uitgebreide oefening kunnen niet altijd garanderen dat dan ook alle vragen goed gemaakt worden. Er is geen enkel bezwaar tegen het gebruik van zulke moeilijke vragen.

Heel anders ligt dat met vragen die hun moeilijkheid daaraan danken dat ze buiten de behandelde stof gaan, en die veeleer een beroep doen op intelligentie dan op verworven kennis en inzicht. Zulke vragen worden wel bedacht en gebruikt onder het motto dat ook de geniale student aan zijn trekken moet komen. Dat zijn twijfelachtige opvattingen. Daarmee wordt de overige studenten geen recht gedaan. Vragen die buiten de opgegeven stof gaan, horen in een toets niet thuis, daar is geen discussie over nodig. Die vragen zijn ook niet voor te bereiden hoe goed je de stof ook bestudeert, voor het kunnen beantwoorden van bonusvragen is het veeleer van belang dat je slimmer dan anderen geboren bent.

Negatieve gevolgen van ondoorzichtigheid. Leerervaringen zijn ruwweg in twee categorieën in te delen: ervaringen die meer inspelen op het verwerven van vaardigheden ('cogni-

tieve'), tegenover ervaringen die meer het gevoelsmatige benadrukken, de houding tegenover de stof (het affectieve). Het afleggen van een toets is een geweldig intense ervaring voor de student, en het is dan ook te verwachten dat bij gelegenheid van een afsluitende toets er sterke positieve of negatieve gevoelens ten opzichte van het getoetste vak kunnen optreden. Wie erop uit is om studenten afkeer van het eigen vak bij te brengen, kan ervoor zorgen dat de toetsvragen zo ondoorzichtig mogelijk zijn (in beide eerder onderscheiden betekenissen). Wie van de toetsgelegenheid gebruik wil maken om de motivatie van de student een extra duwtje in de positieve richting te geven, die zorgt er om te beginnen voor dat de toetsvragen doorzichtig zijn, terwijl het daarnaast de moeite waard is om de vragen zo te kiezen dat de student aan de toets het gevoel kan overhouden dat hij of zij een behoorlijk inzicht in het getoetste vak heeft verkregen, en dat inzicht nog verder heeft kunnen verdiepen tijdens de toetsing.

2.5 Rompvragen: meervoudig bruikbare vraagskeletten

Iedereen is vertrouwd met vraagstellingen die qua vorm aan elkaar identiek zijn, maar waarbij de inhoud gevarieerd is.

VOORBEELD

$$23 + 56 = ?$$

$$23 + 20 = ?$$

$$23 + 11 = ?$$

Het skelet van deze vragen is te schrijven als ' $23 + \dots = ?$ ', op de opengelaten plaats wordt een getal tussen 10 en 100 gekozen om de opgave te completeren. Vraagskelet *samen met* de specificatie van wat op de opengelaten plaats(en) ingevuld kan worden, noemen we een *rompvraag* (de Engelse term hiervoor is *item form*).

VOORBEELD ROMPVRAAG

$$\dots + \dots = ?$$

specificatie: getallen van 10 tot 90, maar zodanig dat de som niet groter is dan 100.

Het gebruik van rompvragen kan een krachtig hulpmiddel zijn bij het aanmaken van grotere hoeveelheden toetsvragen. In veel gevallen zal de rompvraag ook bruikbaar zijn in deze zin dat in achtereenvolgende toetsen telkens hetzelfde vraagskelet gebruikt kan worden; dat is niet telkens dezelfde vraag, omdat herhaald gebruik van hetzelfde vraagskelet gebeurt met telkens een andere keuze uit de specificatie van het 'vlees', de woorden, getallen en dergelijke, die op de open plaatsen ingevuld kunnen worden.

In het algemeen bestaat de specificatie van wat in het vraagskelet ingevuld kan worden, uit een *lijst* of een *verzameling*. Een lijst is een verzameling die bestaat uit een opsomming. Andere manieren om een verzameling te bepalen zijn: de omschrijving, zoals in het gegeven voorbeeld, per definitie (zie ook 3.5), per constructievoorschrift, enz...

Een vraag wordt gemaakt door het vraagskelet in te vullen met een getal, voorbeeld enz. uit de specificatie. De keuze voor een bepaald getal, voorbeeld e.d. kan door toeval bepaald worden, waarbij rekening te houden is met getallen, voorbeelden e.d. die al in het onderwijs of in recente toetsen gebruikt zijn.

Het vraagskelet hoeft niet rigide te zijn, maar kan ook gevarieerd worden.

VARIANT

23

...

___+

?

VARIANT

23 + ? = ...

Dergelijke varianten kunnen de vraag inhoudelijk vrijwel onveranderd laten (de eerste gegeven variant), of ook de inhoud zelf veranderen (de tweede variant). Wat mogelijk is met de vorm van sommen, kan natuurlijk ook met de redactie van verbale opgaven. Let er wel op of bij herformuleren van een vraagskelet dezelfde specificatie volledig geldig blijft. Het is aan te bevelen altijd enkele alternatieve vraagskeletten

te maken bij iedere rompvraag. Wanneer altijd hetzelfde vraagskelet gehanteerd zou worden, ontstaat de mogelijkheid dat studenten het vraagskelet leren zien als een hint voor het beantwoorden van de opgave. Dit is minder onschuldig dan het lijkt, het produceren van bepaalde kennis wordt onbedoeld gekoppeld aan een bepaalde vorm van vraagstelling, en dat maakt het de student straks onmogelijk om in andere situaties deze relevante kennis te produceren (zie ook 7.1 over produkties).

Dat rompvragen handig zijn bij rekenopgaven, oké, maar hoe gebruik je ze voor andere leerstof? Wel, op analoge wijze. Voor andere leerstof dan wiskundige zien opgaven er nogal eens als volgt uit.

VOORBEELD

Beschrijf hoe je twee getallen bij elkaar optelt, wanneer beide niet kleiner dan 10 en niet groter dan 100 zijn.

Hier maak je niet een rompvraag van door de getallen te veranderen:

AFRADER

Beschrijf hoe je ... getallen bij elkaar optelt, wanneer beide niet kleiner dan ... en niet groter dan ... zijn.

Het probleem met dit voorbeeld is dat het op heel algemeen, *abstract* niveau geformuleerd is. Dergelijke abstracte vragen laten geen rompvraag toe, eenvoudig omdat de rompvraag gebruik maakt van precies die zaken die in de abstracte vraag bij elkaar genomen worden door de abstractie. Bij rekenopgaven is de remedie vanzelfsprekend: formuleer de vraag met concrete getallen en vraag die bij elkaar op te tellen. En daar is heel eenvoudig een rompvraag van te maken (het eerste voorbeeld dat ik van een rompvraag gegeven heb). Bij opgaven over andere leerstof geldt nu hetzelfde: werk met *geconcretiseerde* vragen, vragen waarin met voorbeelden en dergelijke gewerkt wordt (zie ook 2.6). Dan zijn rompvragen te maken door een verzameling van voorbeelden aan te leggen of anderszins te omschrijven

VOORBEELD

Three Mile Island, Harrisburg, is één van de kernreactoren waar een ernstig ongeluk mee is gebeurd. juist / onjuist

Specificatielijst (alleen USA):

- Fermi, Detroit, snelle kweek, 1966 (gedeeltelijke kernsmelt)
- Idaho Falls, 1961 (gedeeltelijke kernsmelt)
- Browns Ferry, 1975
- Three Mile Island, Harrisburg, 1979
- onjuiste voorbeelden: overige Amerikaanse kernreactoren

Merk op dat een specificatie van voorbeelden soms uitgebreid moet of kan worden met een lijst van niet-voorbeelden (zie ook hoofdstuk 4). De vraag in het voorbeeld zou een 'onbenullige toetsvraag' kunnen zijn, bedoeld om te toetsen of de student een artikel of boek over ongelukken met kernreactoren tenminste gelezen heeft (zie over deze wijze van toetsen verder 6.1).

VOORBEELD

Bijgaande foto is een foto van een

- a) hazepootje
- b) muizeoor
- c) vogelpootje
- d) cypreswolfsklauw.

De specificatie bij een vraagskelet kan vaak op heel verschillende manieren ingevuld (of aangelegd) worden. Het plantje 'muizeoor' uit bovenstaand voorbeeld kan op verschillende manieren getekend zijn, gefotografeerd zijn, foto's die al dan niet een belangrijk stuk van de omgeving laten zien waarin het muizeoor groeit. De verzameling kan zelfs gedroogde exemplaren bevatten, of ongerepte exemplaren op locatie (dan maakt de vraag onderdeel uit van het veldwerk). Het opstellen van een specificatie hoeft voor een deskundige op het desbetreffende vakgebied niet lastig te zijn. Wanneer de verzameling uit voorbeelden bestaat, is te bedenken dat voorbeelden op hun beurt ook weer in soorten komen. Doe een passende keuze uit de volgende mogelijkheden:

- voorbeelden die al in tekst, syllabus, werkbespreking behandeld zijn;

- voorbeelden waar de student zich op geïnteresseerd heeft
- nieuwe voorbeelden (en daar gaat het eigenlijk om), zoals:
 - * meest voorkomende, modale, prototypische voorbeelden;
 - * gewone, gemiddelde, normale voorbeelden;
 - * buitenissige, extravagante, uitzonderlijke voorbeelden;
 - * grensgevallen (waar ook deskundigen met elkaar van mening verschillen);
 - * uitzonderingen (die afwijken van een definitie, en desondanks toch ...);
 - * voorbeelden onder restrictie (uit een bepaald tijdperk, gebied enz. te kiezen);
 - * overgeneralisatie-gevoelige;
 - * misvattingen en fouten (o.a. uitkomsten verkregen bij een rekenfout);
 - * willekeurige niet-voorbeelden (zij het ook onder bepaalde restricties te kiezen);
 - * tegengestelden, antonymen en dergelijke (gaat gevaarlijk in de richting van intelligentietest-achtige vraagstellingen).

Wat voor voorbeelden geldt, gaat ook op voor *toepassingen* van wetten, wetmatigheden, technieken en dergelijken. De hoofdstukken 4 en 5 gaan op voorbeelden en toepassingen verder in.

Het laatste voorbeeld is een rompvraag die wat het vraagskelet betreft ook voor andere onderwerpen toepasbaar is (bijvoorbeeld voor hazepootje, vogelpootje, en andere plantesoorten).

Vaak zullen vraagskeletten bij bepaalde onderwerpen ook bruikbaar zijn bij andere onderwerpen. Het bedenken van een vraagskelet is niet altijd even eenvoudig (daar gaat eigenlijk de rest van dit boek over), zodat het economisch gebruik maken van al bedachte en in de praktijk beproefde vraagskeletten voor verwante onderwerpen zeker altijd de moeite waard is. Sommige vraagskeletten zijn zo algemeen bruikbaar dat de rompvraag er nauwelijks in te 'zien' is:

VOORBEELD

Benoem ... [foto, voorwerp e.d.]

VOORBEELD

Evalueer ... [wiskundige uitdrukking]

VOORBEELD

Maak ... [en dan volgt een specificatie]

De suggestie bij dit alles is dat een rompvraag het mogelijk maakt om onder verwisseling van voorbeelden dezelfde vraag herhaalde malen te gebruiken in opeenvolgende toetsen. Is dat verwisselen van voorbeelden wel een voldoende waarborg tegen het uit het hoofd leren van vragen, in dit geval van vraagskeletten? Dat hangt ervan af. Wanneer zo'n vraagskelet in hoge mate specifiek is voor het onderwerp waar de vraag over gaat, dan bestaat dat gevaar inderdaad. Maar voor hoeveel vraagskeletten over uw leerstof zou dat kunnen gelden? Waarschijnlijk is dit probleem altijd wel te omzeilen door geen barokke, opvallende vraagskeletten te bouwen. Een verwante vraag is deze: hoe verschillend moeten twee vragen zijn om van de tweede vraag te kunnen zeggen dat het 'nieuwe' vraag is, vergeleken met de eerste? Bloom c.s. (1956) formuleerden het antwoord zo: dat een goede (toepassings)toetsvraag een probleem moet behelzen dat als zodanig aan de student bekend kan zijn, maar dat stelt op een manier waar hij waarschijnlijk nog niet eerder aan gedacht heeft ('a problem known to the student but a new slant that he is unlikely to have thought of previously'). Maar dat is de zaak op de kop zetten, het gaat er in het onderwijs immers om dat de student leert bepaalde soorten vraagstelling te beantwoorden, bepaalde klassen van problemen aan te pakken en op te lossen. Welnu, een 'nieuw' probleem of een 'nieuwe' vraag zal doorgaans moeten passen in een *bekend* slag, type of soort probleem of vraag. Het nieuwe kan bestaan uit een andere formulering of een andere vorm, maar dat is doorgaans van tamelijk ondergeschikt belang. Wat een vraag 'nieuw' maakt, is veeleer een voorbeeld dat de studenten nog niet bekend is, een situatie en dergelijke. Met andere woorden: rompvragen leveren verzamelingen vragen op, waarbij de vragen ten opzichte van elkaar 'nieuw' zijn doordat nieuwe voorbeelden uit de specificatie erin gebruikt worden. Tenzij de specificatie voorbeelden bevat die als twee druppels water op elkaar lijken, levert ieder voorbeeld een nieuwe vraag op. De nieuwigheid kan eventueel benadrukt worden door een andere formulering van het vraagskelet te gebruiken.

2.6 Vermijd abstracte vraagstellingen

‘Tussen moordenaars en hun slachtoffers staat de technologie en de symbolische taal, en deze laatste maken het mogelijk om moord, verminking en vernietiging op een schier onbeperkte schaal uit te voeren. De moord op de mensheid is nu niet alleen mogelijk geworden, maar is al voorbereid en georganiseerd; de uitvoering wacht nog slechts op een teken, Deze mogelijkheid werd geschapen door de toegepaste wetenschap en door die gecompliceerde vormen van menselijke organisatie die we beschaving noemen: beide zijn de voortbrengselen van de symbolische taal.’ Aldus Anatol Rapoport in zijn boek over semantiek (1975, blz. 17). De taal symboliseert de werkelijkheid, daardoor maakt ze het in zekere zin mogelijk om efficiënter met diezelfde werkelijkheid om te gaan, De prijs die daarvoor betaald wordt, is het risico dat symboliek en realiteit van elkaar vervreemd raken.

Een middel dat kan leiden tot de zelfvernietiging van de mensheid moet wel een bijzonder krachtig middel zijn. In het onderwijs wordt iedere telg van het menselijk geslacht bijgebracht hoe je met dat middel om kunt gaan of moet gaan, hoe je met de taal omgaat, hoe je wetenschappelijke taal gebruikt.

De boodschap is duidelijk: de taal is een abstractie van de werkelijkheid, en om met zo’n abstractie op een behoorlijke wijze om te kunnen gaan, is een voortdurend terugkoppelen naar de achterliggende werkelijkheid noodzakelijk. De wetenschap heeft daar een geheel eigen ritueel voor ontwikkeld, waarbij de werkelijkheidswaarde van nieuwe theorieën, of van nieuwe gevolgtrekkingen uit gevestigde theorieën, aan ‘de empirie’ getoetst wordt. Het is een hoofdbestanddeel van wetenschappelijk onderwijs om studenten in dat ritueel in te wijden. Op een veel nederiger niveau moet de docent de koppeling tussen taal en werkelijkheid, tussen abstractie en concrete zaken, ook voortdurend in de gaten houden. In onderwijs en toetsing is het zaak er voortdurend op te letten of onderwezen en getoetste onderwerpen op het juiste niveau van abstractie behandeld worden. Een simpel voorbeeld: het onderwerp ‘olifanten’ behandel je niet door leerlingen te informeren dat het zoogdieren met flaporen zijn, en toets je niet door deze definitie terug te vragen.

In het bijzonder bij het toetsen is de verleiding groot om ge-

noegen te nemen met vragen die zich bewegen op het verbale abstractieniveau, terwijl de onderwijsdoelen toch evident voorschrijven dat de student met *concrete* dingen, technieken, gebeurtenissen enz. moet kunnen omgaan.

AFRADER

De gedragsverandering die plaatsvindt als resultaat van het herhaaldelijk presenteren van een stimulans terwijl het individu probeert daar effectief op te reageren, staat bekend als:

- 1) rijping
- 2) reactie
- 3) leren
- 4) conditionering
- 5) bekrachtiging

(Mouly en Walton, 1962, item 2.675).

Het gaat er nu om een redelijk antwoord te vinden op de vraag hoe concreet een vraag gesteld moet worden om concreet genoeg te zijn. Omdat een toets doorgaans een papier-en-potlood aangelegenheid is, zijn de mogelijkheden om de vragen aan concrete objecten enz. te formuleren, slechts beperkt. Op het scala van mogelijkheden tussen concreet en abstract moet de keuze zodanig bepaald worden dat de relevantie van de vraag geen geweld wordt aangedaan, dat de vraag geen gekunsteld karakter krijgt. Enige mogelijkheden van concreet tot abstract zijn:

- presentatie van het ding of verschijnsel zelf (een olifant);
- een natuurgetrouwe afbeelding of model;
- een schematische afbeelding;
- een symbolische afbeelding of weergave;
- een verbale beschrijving van een specifiek voorbeeld;
- een verbale beschrijving van het bedoelde begrip (de afrader hieronder);
- alleen de term waarmee het begrip aangeduid wordt (het woord 'olifant').

Het gegeven rijtje omspant abstractieverschillen die gaan van het aanwijzen van een olifant tot het alleen maar gebruiken van de term 'olifant'.

Op ieder aangegeven niveau zijn nog weer fijnere onderscheidingen in abstractie te maken: het maakt verschil of die olifant

alleen maar op afstand bekeken kan worden, of dat het dier aangeraakt, besnuffeld of beklommen mag worden. Een verbale beschrijving van een olifant kan zich beperken tot 'een zoogdier met flaporen', of zo beeldend zijn dat je het dier bijna denkt te kunnen zien en aanraken.

De verbale beschrijving is voldoende 'concreet' in al die gevallen waarin verondersteld mag worden dat de student zich bij die beschrijving een tamelijk levendige voorstelling kan vormen. Dat kan er in grensgevallen toe leiden dat de docent speculeert op het voorstellingsvermogen van de student: dat is sterk af te raden, omdat de ene student zich nu eenmaal gemakkelijker dergelijke voorstellingen vormt dan de andere. Dit voorstellingsvermogen hoeft zeker geen verband te houden met de mate van beheersing van de stof, noch met algemene intelligentie, Probeer liever de vraag te concretiseren, maak gebruik van beschikbare audiovisuele media of van de mogelijkheden om (zwartwit)foto's mee te reproduceren op het toetsformulier.

Toetsvragen op een te algemeen niveau verlokken de student tot het uit het hoofd leren van tekst, definities enz. Toetsvragen op een meer concreet niveau vragen een andere voorbereiding van de student: hij moet de stof actiever verwerken, zelf nieuwe voorbeelden bedenken, nieuwe voorbeelden leren herkennen en dergelijke. Passief door de student verwerkte informatie, bijvoorbeeld voor een literatuurtentamen, werkt net zo ongunstig uit als eenzijdig abstract verwerkte informatie. Een toetsvraag die passief verwerkte informatie terugvraagt, en tegelijk te abstract, te algemeen is, is de volgende:

AFRADER

Een belangrijk verschil tussen de klassieke en de keynesiaanse economie betreft de veronderstelde oorzaak van werkloosheid.

Bij welke school past de gedachte dat deze het gevolg is van onvoldoende vraag en daardoor gedaalde produktie?

Meer actieve behandeling van dit gegeven vraagt om het herkennen van bepaalde uitspraken die economen uit beide scholen gedaan hebben over werkloosheid. Deze uitspraken kan de docent gemakkelijk verzamelen, en hij kan er een deel van in het onderwijs en als oefenmateriaal gebruiken, en de

rest bij het schrijven van toetsvragen.

Een bijkomend voordeel van een meer concreet, minder algemeen niveau van vraagstelling is dat het dan heel wat gemakkelijker is om een groter aantal varianten van dezelfde vraag te schrijven. De afrader hierboven laat weinig variatie in vraagstelling toe. Zodra in plaats van de algemene formulering van een begrip een specifiek voorbeeld wordt ingevuld, zijn er legio variaties te maken door te putten uit de onbeperkte mogelijkheden om voorbeelden van begrippen te bedenken. Ieder nieuw voorbeeld levert een nieuwe toetsvraag op over hetzelfde onderwerp (rompvraag, zie 2.5).

Is het aan te raden altijd de mogelijkheid van een lager abstractieniveau te overwegen, dan moet daar wel als clause bij dat dat lagere abstractieniveau werkelijkheidswaarde moet hebben en niet gekunsteld is. In de exacte wetenschappen, die overigens model kunnen staan voor voorbeeldige vragen over de leerstof, komen gekunsteld concrete toetsvragen nog wel eens voor.

AFRADER

Een slee met een massa van 10 kg heeft een waargenomen versnelling van 5 m/s^2 . Hoe groot is de netto kracht die erop inwerkt?

Antwoord: $F = ma = 10 \text{ kg} \cdot 5 \text{ m/s}^2 = 50 \text{ N}$

Mogelijk is het de bedoeling van de vragenschrijver om de student een concreet probleem voor te leggen, met de vraag hoe dat probleem op te lossen is. In dit geval door er Newtons tweede bewegingswet op toe te passen. In plaats daarvan wordt een vraag voorgelegd waarin voor de goede verstaander direct af te lezen is dat F gevraagd is terwijl m en a gegeven zijn. De afrader vraagt slechts het manipuleren met getallen in de formule $F = ma$. De afrader past beter bij het onderwijsdoel 'oplossen van vergelijkingen met één onbekende', dan bij 'Newtons tweede bewegingswet weten toe te passen'.

VOORBEELD (beter)

Beschrijf twee concrete situaties waarin Newtons tweede bewegingswet van toepassing is.

Leerstof die alleen op abstract niveau behandeld en getoetst

wordt, brengt de student in de vervelende situatie dat hij nimmer de concrete toepassingen weet te maken, en juist daardoor ook de beheersing van de abstractie niet anders kan bereiken dan door uit het hoofd leren. Dan wordt noch beheersing op concreet niveau noch die op abstract niveau bereikt.

Abstracties zijn slechts zinvol in samenhang met de meer concrete zaken die geabstraheerd worden. Wie die relatie verbreekt, doet er beter aan in het geheel geen onderwijs te geven.

Over de wiskundige Peano (1858-1932) maakt Kneebone (1952) de volgende opmerking, waarmee deze paragraaf passend afgesloten wordt: 'He never tired of trying to persuade other teachers that they fail in their task unless they are able to make every fact which they have to communicate come to life in the student's mind, so that he sees it as something directly relevant to his own thinking.'

2.7 literatuur

BELNAP, N. D. Jr., & T. B. STEEL, Jr., *The logic of questions and answers*, London, Yale University Press, 1976.

Dit boek behandelt exact wat de titel ervan belooft: een logisch filosofische verdieping voor de kunst van het vragenstellen. Bevat een rijke bibliografie, opgedeeld in 'Logic and philosophy of language', 'Linguistics', 'Automatic question answering', & 'Psychology and pedagogy'.

THORNDIKE, R. L. (ed.), *Educational measurement*, Washington D.C., American Council on Education, 1971.

Dit is het handboek over toetsen en beoordelen. In het bijzonder: A. G. WESMAN: 'Writing the test item', R. L. THORNDIKE 'Reproducing the test' (over lay-out, druktechnieken e.d.), en W. E. COFFMAN: 'Essay examinations'.

DE GROOT, A. D., & R. F. VAN NAERSSSEN (red.), *Studietoetsen, construeren, afnemen, analyseren*, deel I, Den Haag, Mouton, 1973.

Deze tekst verwoordt, bij monde van De Groot, een uitgesproken standpunt voor het gebruiken van meerkeuzevragen.

VAN DEN BRINK, W. P., 'Het optimale aantal alternatieven per item', *Tijdschrift voor Onderwijsresearch*, 1979, 4, 151-158.

Een overzicht, en een advies: 'Vierkeuze-items leveren voortreffelijke driekeuze-items op indien men het slechtste alternatief aan de vuilnisman meegeeft.'

DOUSMA, T., & A. HORSTEN, *Tentamineren* (Aula 804), Utrecht, Het Spectrum, 1980.

Bestrijkt de belangrijkste onderwerpen die ook in De Groot en Van Naerssen te vinden zijn, maar is toegankelijker (en beknopter) geschreven.

POPHAM, W. J., *Modern educational measurement*, London, Prentice-Hall, 1981.

Een gunstig, maar toch representatief voorbeeld van de behandeling van het onderwerp toetsvragen schrijven in Amerikaanse stijl.

COHEN, M. J., *Studierechten in het wetenschappelijk onderwijs*, Zwolle, Tjeenk Willink, 1981.

Bespreekt studierechten en doceerplichten. Zie i.h.b. blz. 95 'transparantie of kenbaarheid', en blz. 97 'kenbaarheid als beginsel van behoorlijk bestuur?' voor de meer juridische aspecten van doorzichtigheid.

SPARKS, C. P., 'Open versus secure testing', *Personnel Psychology*, 1980, 33, 1-2.

Inleiding op een themanummer van *Personnel Psychology* op dit thema. In de Verenigde Staten is namelijk een heftige strijd losgebarsten rond het achteraf vrijgeven van gestandaardiseerde tests die gebruikt worden bij selectieve toelatingen tot instellingen van Hoger Onderwijs. In steeds meer staten wordt vrijgeven van items een wettelijke verplichting.

HOFSTEE, W. K. B., 'Bureaucratisering van de beoordeling? Een beschouwing naar aanleiding van het voor en tegen van multiple-choice-toetsen', *Universiteit en Hogeschool*, 1981, 27, 286-292.

ROID, G., & T. HALADYNA, *A technology for test-item writing*, London, Academic Press, 1982.

Dit boek representeert 'the state of art' in de Verenigde Staten wat betreft vuistregels voor het schrijven van toetsvragen.

CROMBAG, H. F. M., 'Over het inrichten van leersituaties', in: A. I. VROEIJENSTIJN (red.), *Kwaliteitsverbetering hoger onderwijs*, Voorburg, CBOWO/Stichting Nationaal Congres, 1981.

Over het belang van het 'concretiseren' van het onderwijs.

RAPOPORT, A., *Semantics*, New York, Thomas Y. Crowell, 1975.

Niet alleen van belang vanwege de hieruit geciteerde passage over te ver doorgevoerde abstractie van de werkelijkheid, maar ook als algemene achtergrond, naast het al genoemde boek van BELNAP & STEEL, voor de kunst van het stellen van betekenisvolle vragen over leerstof (over de wereld).

3 Leerstofinventarisatie

Het voorgaande hoofdstuk behandelde verschillende vraagvormen, algemene punten waar bij het schrijven van toetsvragen op te letten is, en het verhogen van de efficiëntie bij het schrijven van toetsvragen; het laatste vooral door het gebruik van rompvragen. Maar nog niets over het *bedenken* van vragen of rompvragen. Wil dat 'bedenken' ook efficiënt en zonder herhaaldelijk blokkeren verlopen, dan vraagt dat een degelijke voorbereiding voordat met het schrijven van de vragen begonnen wordt. Die voorbereiding bestaat uit het inventariseren van de leerstof, op zo'n manier dat daarna het schrijven zoveel mogelijk een routinematige opgave kan zijn. De leerstofinventarisatie structureert al enigszins wat er over deze stof gevraagd kan worden, en levert daarmee een leidraad bij het schrijven van de toetsvragen.

In de onderwijskunde is de meest begane weg tot het inventariseren van leerstof de omweg via het formuleren van onderwijsdoelen. In tegenstelling daarmee sluit de hier gekozen wijze van inventariseren heel nauw aan op de *inhoud* van de leerstof, de aard van de nieuwe *termen* die daarin geïntroduceerd worden, met hun onderlinge relaties. Drie soorten termen zijn in bijna iedere wetenschap te onderscheiden: waarneembare termen, abstracte termen en theoretische termen. Zowel in het onderwijs zelf als bij het toetsen spelen deze termen een eigen rol. Men zou nu kunnen zeggen dat het onderwijs bedoelt de student met deze wetenschappelijke terminologie op adequate wijze te leren omgaan, waaronder ook begrepen is het verkrijgen van een overzicht hoe deze termen onderling gerelateerd zijn. Bij voorkeur strekt de leerstofinventarisatie zich uit tot het schematisch expliciteren van de relevante samenhangen tussen de gespecificeerde termen.

In dit hoofdstuk ook aandacht voor het *definiëren* van termen. In het onderwijs spelen definities vaak een hoofdrol, die in de toetsing dan ook gecontinueerd pleegt te worden. Termen zijn veelal op verschillende manieren te definiëren, sommige ter men laten maar één bepaald type definitie toe, terwijl er ook termen zijn die zich in het geheel niet laten definiëren.

Het is van wezenlijk belang dat de docent zich een scherp inzicht verwerft in het karakter van definities, om zelf te kunnen bepalen wat wel en wat niet zinvol is bij het verwerken van bepaalde definities in toetsvragen.

Wetenschappelijke terminologie. Wat voor de student altijd hetzelfde blijft, of hij nu natuurkunde, geneeskunde, sociologie of rechten studeert, is dat hij zich de terminologie van dat vakgebied eigen moet maken. Niet alleen de gedefinieerde betekenissen van deze vaktermen, maar ook de wijze waarop je zinvol over de problemen in dat vakgebied kunt spreken met gebruikmaking van de daarvoor ontwikkelde terminologie. Neem een studieboek dat tot de opgegeven stof voor de toets hoort, en inventariseer alle nieuw geïntroduceerde termen en alle al bekende termen die van een nieuwe of een verdiepte betekenis voorzien worden. Bij studieboeken met een gedetailleerde inhoudsopgave zijn die termen daar bijna alle in terug te vinden. Dekt zo'n opsomming van termen de hele leerstof? Ja en nee. Nee omdat ook de onderlinge relaties tussen bepaalde termen een wezenlijke rol zullen spelen, zodat ook die relaties bij de inventarisatie betrokken moeten worden. Nee omdat een term niet meer is dan een label dat aan een bepaald onderwerp gehecht is, maar dit is juist gewenst omdat de term als label ook gebruikt kan worden als 'stand-in' voor dat onderwerp. Ja, omdat op deze wijze de hele stof in de opsomming van termen (en hun relaties) gerepresenteerd wordt. Bij het schrijven van toetsvragen wordt de zekerheid dat de hele stof in deze vragen betrokken wordt, verkregen uit het werken aan de hand van de opgestelde lijst van termen (en hun relaties). Daartoe is het handig in de lijst van termen enige ordening aan te brengen, zowel naar hun onderlinge relaties (door te schematiseren bijvoorbeeld), als naar onderscheiden soorten van termen. In de volgende paragrafen worden termen in drie soorten verdeeld: (indirect) waarneembare, abstracte en theoretische termen. Deze indeling van soorten termen is van direct belang bij het schrijven van toetsvragen omdat globaal gesproken bij waarneembare termen een ander slag toetsvragen past dan bij theoretische termen. Met andere woorden: wanneer aangegeven is dat een term behoort tot de categorie van waarneembare termen, dan is daarmee ook een beperking gegeven in het aantal te overwegen soorten van

toetsvragen over die term: het zoeken naar toetsvragen wordt daarmee richting gegeven.

3.1 (Indirect) waarneembare termen

In iedere wetenschap spelen waarneembare verschijnselen een belangrijke rol. Bij toegepaste wetenschappen is dat vanzelfsprekend. Maar ook waar de uitbouw van de theorie vooropstaat, spelen waarneembare verschijnselen een onmisbare rol bij het toetsen van hypothesen die uit die theorie afgeleid worden.

Waarneembare termen zijn er in velerlei soort, maar alle hebben ze gemeen dat ze verwijzen naar wat op ongecompliceerde wijze voor ons waarneembaar is.

Voor een overzicht heb ik het volgende lijstje van voorbeeld categorieën opgesteld.

-*Dokumenten, oorkonden, akten, overeenkomsten.* Het gaat hier om de stukken zelf zoals ze in archieven bewaard worden, die je in je handen kunt nemen, en niet om de betekenis die een bepaald document heeft, noch om de interpretatie, exegese, of vertaling ervan. Bijv.: de Unie van Utrecht.

-*Boeken, gedichten, rapporten, geschriften.* Wittgensteins *Filosofische Onderzoekingen*, Gorters *Mei*, *Rapport van de Commissie van Drie*. Het bronnenmateriaal van iedere wetenschapper, en voor geschiedkundigen en theologen studieobject.

-*Wetsartikelen, jurisprudentie.* Het gaat hier om de letterlijke tekst van de artikelen; de interpretatie daarvan is een ander verhaal. Rechterlijke uitspraken vallen op hun beurt ook weer onder de waarneembare termen.

-*Gebruiksvoorwerpen, kunstvoorwerpen, dingen.* Rembrandts 'Nachtwacht', een vuistbijl, een Rietveld-fauteuil, een kloostermop.

-*Patiënt, cliënt, opdrachtgever, 'de mens'.*

-*Personen.* Plato, Piet Hein, Picasso.

-*Lichaamsonderdelen, planteonderdelen.*

-*Dieren, planten.*

-*Gedragingen.* Babinskyreflex, glimlach, grimlach, aanwijzen.

-*Gebeurtenissen.* Storm, vulkaanuitbarsting, zonsverduis-

tering, demonstratie, overval, veldslag, ondertekening.

- *Instrumenten*. Scalpel, pen, sterrekijker, tachistoscoop, roetglaasje.

- *Meetinstrumenten*. Thermometer, toetssteen, Groninger Intelligentie Test, enquête, lakmoespapiertje.

- *Aanwijzing van meetinstrument*. Wijzerstand van drukmeter, hoogte van kwikkolom, score op neuroticismeschaal, antwoordpatroon op enquête, steekproefresultaat.

- *Kaarten, tekeningen, situatieschetsen*. Bouwtekening, stafkaart.

- *Modellen*. Model van DNA-structuur, model van supertanker.

- *Symbolen, tekens, acronyemen*. PCB, DNA, =, +, £, \$.

Het valt in de gegeven opsomming op dat heel veel van deze waarneembare termen al als bekend beschouwd kunnen worden, nog voordat de student aan de bestudering van het onderhavige vak is begonnen. In andere gevallen maakt de student in zijn studie op 'vanzelfsprekende' wijze kennis met de relevante waarneembare termen, die dan ook al even vanzelfsprekend niet het speciale onderwerp van de toetsing vormen. In alle overige gevallen zijn de waarneembare termen meer expliciet onderwerp van onderwijs, wordt de student er vertrouwd mee gemaakt, leert de student om te gaan met deze termen op de wijze die in het vak gebruikelijk is. En op dat laatste zullen ook de toetsvragen gericht zijn, over herkennen, benoemen, het geven van voorbeelden, aanwijzen e.d.

In de titel van deze paragraaf heb ik aangegeven dat er ook *indirect* waarneembare termen zijn. Daarbij gaat het om gebeurtenissen, dingen, eigenschappen, e.d., die niet waar te nemen zijn, waar instrumenten voor de waarneming ervan nodig zijn, of waarbij een bepaalde proefopstelling nodig is.

Ook vallen de termen eronder waarbij bepaalde waarnemingen *geïnterpreteerd* moeten worden. Een paar voorbeelden.

- *Droom*. Het verslag dat iemand doet over zijn droom, is concreet, valt onder de waarneembare termen: de droom waar het naar verwijst, is niet direct waar te nemen, slechts op indirecte wijze via het droomverslag.

- *Magnetisme*. Of een voorwerp magnetisch is of niet, blijkt

pas bij het uitvoeren van een *proef*. Magnetisme - en heel veel andere eigenschappen van dingen, personen enz. - is conditioneel gedefinieerd (waarover later meer): alleen aan zijn gevolgen 'waar te nemen'.

-*Witte dwerg*. Sommige sterren horen tot de categorie 'witte dwergen.' Die sterren zijn uitstekend zichtbaar, tenminste op foto's, maar louter aan hun uiterlijke verschijning is het niet uit te maken of je met een witte dwerg te doen hebt of niet. Daar is extra informatie voor nodig.

In het onderwijs leveren indirect waarneembare termen bijzondere problemen op, zoals uit de gegeven voorbeelden al mag blijken. Ook bij toetsvragen over deze termen moet er met dat bijzondere karakter rekening worden gehouden, om onjuistheden en dubbelzinnigheden in de vraagstelling te vermijden. Vandaar de extra moeite om binnen de categorie waarneembare termen het onderscheid te maken tussen datgene wat direct, met het 'ongewapend oog' waarneembaar is, en datgene wat al evenzeer een concreet karakter heeft, maar toch niet direct waarneembaar is.

Toetsvragen kunnen op dat indirecte karakter gericht zijn: 'Waaruit kan blijken dat ...', 'Hoe kom je erachter of ...' 'Welke proef is nodig om ...'. Het kan van belang zijn te toetsen of de student het indirect waarneembare karakter van de onderhavige term weet te onderkennen. Dergelijke vragen impliceren dus vaak het tegelijk toetsen van bepaalde wel waarneembare termen. Van indirect waarneembare termen zijn moeilijk nieuwe voorbeelden te vragen, tenzij de student over de daarvoor benodigde kennis al uit een andere bron beschikt. Het laten herkennen (identificeren, onderzoeken) van nieuwe voorbeelden is natuurlijk wel mogelijk.

3.2 Abstracte termen en constructs

Wanneer een politicoloog over 'de regering' spreekt of wanneer de econoom over 'de markt' spreekt, hebben zij het niet over concreet aanwijsbare zaken. Dergelijke *abstracte termen* zijn een soort verzamelnaam voor heel uiteenlopende activiteiten, uitingen, instanties. Voor een abstracte term is doorgaans geen uitputtende lijst te maken van zaken enz. die er

onder vallen. Neem bijvoorbeeld de abstracte term 'intelligent gedrag' uit de psychologie: concrete gedragingen die als intelligent aan te merken zijn, zijn niet in algemene termen uitputtend te beschrijven. Wie over intelligent gedrag spreekt, bedoelt daarmee alle verschillende uitingsvormen samen te vatten of te abstraheren.

Het zal niet verbazen dat abstracte termen vaak moeilijk te beschrijven zijn. Dat is ook in de wetenschap zo, en dat geldt in optima forma voor de wetenschappen die juist in de fase verkeren dat de betekenis van abstracte termen 'afgetast' wordt. In de psychologie wordt 'intelligent gedrag' telkens verschillend omschreven, ook of misschien juist binnen een en hetzelfde boek. Hetzelfde in de sociologie, bijvoorbeeld bij een term als 'maatschappelijke klasse'. Het gaat er dan niet om die ene betekenis te vinden die de juiste is, maar om in de opsomming van verschillende betekenissen het bedoelde begrip een goede dekking te geven. Een voorbeeld van Wittgenstein over de betekenis van 'de betekenis van een woord': "De betekenis van een woord is zijn gebruik in een taal; zijn gebruik in de praktijk; zijn rol in het taal spel; zijn plaats in de grammatica. De betekenis van een woord is datgene wat uitgelegd wordt in de uitleg van zijn betekenis. De betekenis van een woord is zijn doel, De betekenis van een zin is zijn gebruik; de methode waarop hij geverifieerd kan worden; wat uitgelegd wordt in de uitleg van zijn betekenis. De rechtvaardiging of de grond voor een bewering vormt zijn betekenis." (Voor de bronvermelding voor ieder van deze uitspraken, zie Baker en Hacker 1980, blz. 685).

Het zal niet eenvoudig zijn om dergelijke abstracte termen te vangen in korte toetsvragen. Het gaat vaak om structuren, instituties en dergelijke. De wijze van behandeling van deze termen in tekstboeken zal vaak een verbaal uiteenzettende zijn. De toetsvragen die daarbij passen worden behandeld in hoofdstuk 6.

Constructs. Een andersoortige abstractie is die waarbij de complexe werkelijkheid vereenvoudigd wordt tot een hanteerbaar ideaalmodel. In de natuurkunde hebben we zo de abstracte termen vacuüm, puntmassa, snelheid op tijdstip t , of nul graad Kelvin. Strikt genomen zijn dit onbestaanbaarheden, al kunnen velen zich er misschien een voorstelling van maken

als toestanden die in de limiet benaderbaar zijn. Deze abstracte termen zijn zo gekozen om daarmee in staat te zijn wiskundige modellen of theorieën van redelijke eenvoud, of althans van nog hanteerbare complexiteit, te kunnen bouwen. Ook in andere wetenschappen worden om deze reden abstracte termen gebruikt, denk aan de 'homo economicus' of de 'rationele mens' in de economie, een term als 'Jan Modaal' of 'de gemiddelde student'. In de regel zijn er van deze 'ideaaltypische' abstracte termen geen concrete voorbeelden te geven: die bestaan niet. Met uitzondering van Jan Modaal er zijn immers nogal wat mensen die een modaal inkomen hebben, dat is het meest voorkomende inkomen.

Een ander voorbeeld: in de testleer wordt wel gesproken over de 'ware score' en de 'meetfout'; zo kan de score die een student op de toets behaald heeft, opgesplitst worden in een (niet waarneembare, ook niet noodzakelijkerwijs bestaande, maar wel denkbare) ware score, en een (evenmin waarneembare enz.) meetfout, die samen optellen tot de waargenomen score. Deze abstracte termen zijn in het onderwijs nogal eens problematisch, omdat studenten het specifieke abstracte karakter ervan niet gemakkelijk begrijpen. Abstracte termen van dit type zal ik ook wel aanduiden als *constructs*, als 'gemaakte' abstracties, ter onderscheiding van het soort abstracte termen in het eerste deel van deze paragraaf besproken.

Toetsvragen over en rond constructs zijn op talrijke verschillende manieren vorm te geven. Juist door het gemaakte karakter van deze constructs zullen ze vaak een exact omschreven definitie hebben, waarmee in toegepaste situaties gewerkt moet kunnen worden. Constructs worden ontworpen op hun toepasbaarheid, en op het kunnen toepassen valt dan ook een natuurlijke nadruk. Vragen met een meer indirect karakter liggen meer in de sfeer van wat in hoofdstuk 6 over toetsvragen over tekst behandeld wordt: bijvoorbeeld het herkennen van impliciete aannames, aangeven van consequenties die volgen uit bepaalde gekozen axioma's, de beperkingen inherent aan het gebruik van constructs aan kunnen geven, en dergelijke. Voorzover een construct afgeleid is van andere constructs: deze afleiding kunnen produceren, zelf constructs kunnen vormen bij het aanpakken van complexe problemen, en gekozen constructs, axioma's of veronderstellingen kritiseren, behoren ook tot de mogelijkheden.

3.3 Theoretische termen

Een gerijpte wetenschap is uitgegroeid van het ordenen, beschrijven, en classificeren van empirische gegevens naar het bouwen van theorieën; van het doen van voorspellingen op grond van waargenomen empirische wetmatigheden tot het doen van voorspellingen op grond van afleidingen uit de theorie.

Een wetenschappelijke theorie bevat dan ook, naast waarneembare en abstracte termen, een heel bijzondere categorie termen die we *theoretische* termen zullen noemen. Waarom bijzonder? Wel, men kan zich afvragen waar theoretische termen eigenlijk aan refereren, omdat met de waarneembare en (sommige) abstracte termen al het 'waarneembare' uitgeput is. Theoretische termen verwijzen niet naar waarneembaarheden; en dat is minder een verrassende opmerking als wel een noodzakelijke. Neem een theoretische term uit de natuurkunde zoals 'aantrekkingskracht' of 'massa'. Of een beruchte theoretische term uit de psychologie als 'intelligentie'. Dat zijn termen waarbij niets concreets aanwijsbaar is, het zijn termen waarbij het niet mogelijk is een korte dekkende beschrijving van hun 'betekenis' te geven, het zijn termen waarvan de betekenis vooral gegeven is in de vorm van hun relaties tot de andere termen in de theorie. In het bijzonder zal het doorgaans niet mogelijk zijn om een bevredigende *definitie* van theoretische termen zoals intelligentie of zwaartekracht te geven. Het is daarom zinloos om over voorgestelde definities te bekvechten, of om studenten een definitie van eigen voorkeur te laten leren.

In het onderwijs is de behandeling van theoretische termen doorgaans verre van eenvoudig, juist omdat ze hun betekenis ontlenden aan de diverse relaties die ze binnen een bepaalde theorie hebben tot de andere termen in die theorie. Door het behandelen van de belangrijkste van die relaties wordt de betekenis langzamerhand overgebracht. De betekenis van een theoretische term groeit bij het verder kennis nemen van het vak, en zal nooit helemaal 'af' zijn.

Een wetenschappelijke theorie is vaak in wiskundige kleren gestoken, waarbij de theoretische termen en hun bijzondere onderlinge relaties in formules zijn vastgelegd: $E = mc^2$, $F = ma$ enz. Nieuwe theoretische termen worden, in de di-

dactische opzet van het studieboek, meteen in formule gepresenteerd, en de student leert op een aantal relevante manieren met zo'n formule, en daarmee met deze theoretische termen, om te gaan. Naast wiskundig onderbouwde theorieën zijn er uiteraard ook theorieën die meer verbaal van karakter zijn. zoals de persoonlijkheidstheorie van Freud of de evolutietheorie van Darwin, waarin de diverse theoretische termen even streng aan elkaar gerelateerd kunnen zijn als dat in wiskundige formalismen het geval is.

Theorieën handelen over gepostuleerde eigenschappen van dingen, personen, gebeurtenissen enz. In het dagelijks spraakgebruik hebben dergelijke eigenschappen een 'observeerbare' status. In de wetenschappelijke theorie is dat veelal *niet* het geval, spelen ze hun rol als theoretische termen.

Ik noemde al de persoonlijke eigenschap 'intelligentie', en hetzelfde geldt voor eigenschappen in de natuurkunde: 'lengte', 'massa', 'energie', 'snelheid', 'temperatuur'. In de natuurkundige theorie is de koppeling van deze termen aan waarneembare termen soms heel weinig problematisch, en bestaat de verleiding om de theoretische term 'lengte' simpelweg gelijk te stellen aan de waarneembare term 'lengte'. Zo'n simpele gelijkstelling is al niet meer mogelijk waar het gaat om de 'snelheid op het tijdstip t ', omdat deze snelheid niet een corresponderende waarneembare term heeft, maar zijn 'waarde' berekend moet worden uit meerdere waarnemingen. In de psychologie is de koppeling van de theoretische term 'intelligentie' aan waarneembare verschijnselen heel weinig direct; ik kom daar in 3.4 bij wijze van voorbeeld nog uitgebreider op terug.

Omdat de betekenis van theoretische termen vooral in hun relaties tot andere termen ligt, zullen toetsvragen alleen die relaties kunnen betreffen. Er bestaat niet zoiets als een toetsvraag die zich strikt beperkt tot één term die theoretisch van karakter is. Ook vragen over de historische ontwikkeling in de betekenis van een theoretische term betreffen de verandering van inzichten in de relaties van deze term tot andere, waarneembare of theoretische, termen. Hier wordt het wel erg belangrijk dat netwerk van relaties ook op schrift te stellen; een manier daartoe wordt in de volgende paragraaf gegeven.

Tenslotte: in de wetenschapsfilosofie wordt aan het onderscheid tussen waarneembare en theoretische termen niet meer

de betekenis gehecht die de empiristen eraan gaven (voor een overzicht: Suppe, 1977); dat wil echter niet zeggen dat het onderscheiden van waarneembare en theoretische termen voor de doeleinden waartoe het hier aangewend wordt dan ook minder zinvol zou zijn.

3.4 Onderlinge verknoptheid van termen

De betekenis van individuele termen ligt voornamelijk in de relaties van de term tot andere termen. Dat is al zo bij de waarneembare termen in een classificatie: met elkaar dekken zij alle verschijnselen in een bepaald veld. Dat is nog sterker het geval bij theoretische termen die hun betekenis uitsluitend ontleenen aan relaties tot andere theoretische termen en tot bepaalde waarneembare termen. Voor de leerstofinventarisatie betekent het dat een opsomming van relevante termen de stof slechts eenzijdig beschrijft. De inventarisatie is pas volledig wanneer daarin ook alle relaties tussen deze termen zijn opgenomen voor zover zij in het onderwijs aan de orde zijn gesteld of in de tekst besproken worden. Ik duid dat kort aan als het expliciteren van de onderlinge *verknoptheid* van de termen in de inventarisatielijst.

Om er zeker van te zijn dat geen enkele relevante relatie over het hoofd wordt gezien, kan voor iedere term telkens de hele lijst termen worden nagelopen op de vraag met welke van die termen er zo'n belangrijke relatie bestaat dat zij in het onderwijs behandeld wordt. Dat is omslachtig, en maar zelden nodig. In de natuurkunde bijvoorbeeld zijn de belangrijke relaties tussen theoretische termen gegeven in de behandelde wetten (formules). Wetenschappen die sterk gericht zijn op waarneembare verschijnselen hebben daar veelal goede classificaties of taxonomieën voor ontwikkeld, waarin de onderlinge relaties zijn vastgelegd.

De geïnventariseerde relaties kunnen vervolgens *schematisch* worden weergegeven, op dezelfde wijze als de inhoud van een stuk studietekst geschematiseerd kan worden. Voor dat schematiseren verwijs ik naar de in deze onderwijskundige reeks verschenen boeken van Breuker en van Mirande. De omgekeerde weg is ook begaanbaar: met de lijst van termen als uitgangspunt een schema maken waarin de termen in hun onder-

linge verband worden gezet. In de volgende paragrafen worden een paar eenvoudige voorbeelden gegeven.

Bij het schematiseren kan het handig zijn om onderscheid te maken tussen relaties tussen termen die van dezelfde soort zijn (alleen theoretische termen of alleen waarneembare termen) en relaties tussen termen van verschillend soort (tussen een theoretische en een waarneembare term bijv.), respectievelijk horizontale en verticale verknoottheid te noemen.

Verticale verknoottheid van termen. Rond de theoretische term intelligentie is een verticaal verknoot schema van termen weergegeven in fig. 3.1. Het ligt voor de hand om het schema in zo'n geval ook verticaal te structureren overeenkomstig het onderscheid naar waarneembare, abstracte, en theoretische termen.

Zo'n schema geeft een goed overzicht. De relaties in het schema worden in het onderwijs behandeld en zullen getoetst worden. Dat een relatie in het schema door een enkel lijntje is aangegeven, wil niet zeggen dat zo'n relatie niet erg complex kan zijn. Bijvoorbeeld de wederkerige relatie tussen intelligentie en testscore, die kan staan voor een overzicht van de ontwikkeling van het theoretische intelligentiebegrip in wisselwerking met in onderzoek verkregen empirische gegevens (de testcores). Zo'n complexe relatie is op zich ook weer in een schema uit te werken.

Een korte toelichting op de termen uit fig. 3.1 kan verduidelijken waarom ze benoemd zijn als theoretisch enz. Het is algemeen bekend wat een intelligentietest is. Het is een meetinstrument waarvan verschillende uitvoeringen bestaan, die in

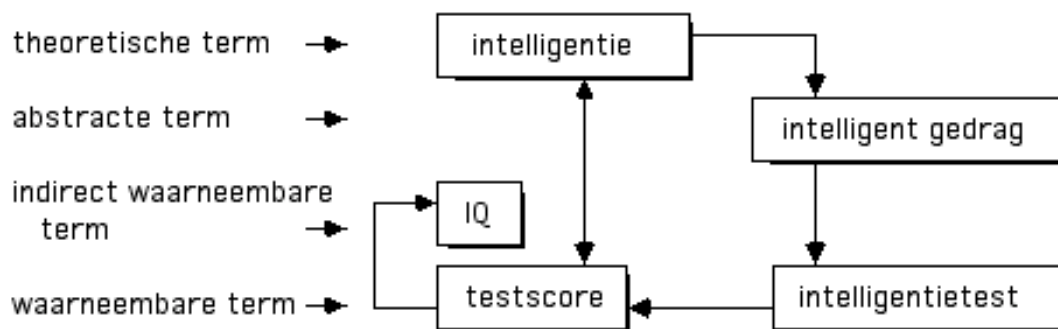


Fig. 3.1. Verticale verknoottheid van termen rond de term 'intelligentie'.

het testarchief te bekijken zijn en waar men in een practicumssituatie mee kan werken, hetzij als testafnemer, hetzij als proefpersoon. De testscore wordt volgens vastgestelde regels bepaald en het IQ is daar een bepaalde bewerking van. De test biedt een gestandaardiseerde situatie waarin bepaalde vormen van intelligent gedrag onderzocht worden. Intelligentie is een theoretische persoonlijkheidseigenschap, de verklarende factor achter uiteenlopende vormen van intelligent gedrag.

In verschillende intelligentietheorieën wordt het in fig. 3.1. gegeven geraamte anders in het vlees gezet: er worden andere accenten gelegd waar het gaat om de uitwerking van wat intelligent gedrag is, de tests verschillen dan ook qua opzet en inhoud van elkaar, en de achterliggende theorie is anders opgebouwd.

Overigens is het schema van fig. 3.1 in dezelfde vorm te gebruiken voor de relaties rond andere theoretische termen, niet alleen in de psychologie maar ook in andere wetenschappen. Bijvoorbeeld de theoretische term temperatuur: temperatuurverschillen in een bepaald voorwerp kunnen blijken uit een heel scala van uiteenlopende verschijnselen, warmtegedragingen zou je kunnen zeggen, naar analogie met intelligent gedrag. De theoretische eigenschap temperatuur is de daarvoor gedachte verklarende factor. Een geschikt meetinstrument voor temperatuur kan gevonden worden door een handige keuze uit de verschillende warmtegedragingen te doen en daarvoor een gestandaardiseerde onderzoekprocedure te ontwerpen. Bijvoorbeeld de wijze waarop de huisarts koortsige patiënten temperatuurt. De uitslag van het meetinstrument is een eenvoudige waarneembare term: deze uitslag wordt bij voorkeur vertaald naar een gebruikelijke temperatuurschaal waarop het instrument geijkt is: de aldus vertaalde uitslag is een indirect waarneembare term, ook al valt hij direct van het instrument af te lezen.

Voor andere theoretische eigenschappen valt dezelfde *schemavorm* te hanteren: agressiviteit, minderwaardigheidscomplex, massa, snelheid, lengte enz.

Horizontale verknooptheid van termen. Relaties tussen theoretische en waarneembare termen kunnen wel gemakkelijk door streepjes aangeduid worden, maar die streepjes staan nogal eens voor complexe, op zich weer afzonderlijk te sche-

matiseren relaties. Dat is anders bij horizontale verknooptheid; de relatie van de ene waarneembare term tot andere waarneembare termen is doorgaans simpel, terwijl wetmatigheden die de ene theoretische term verbinden met andere theoretische termen doorzichtig te formuleren zijn. Voorbeelden van horizontale verknooptheid van waarneembare termen zijn te vinden in classificaties of taxonomieën, waarvan het periodiek systeem van scheikundige elementen wel een heel fraai voorbeeld is. Verhalende teksten zijn met weinig moeite te schematiseren.

Bij theoretische termen zijn de onderlinge relaties van wetmatig karakter. Is de onderhavige theorie in een wiskundig jasje gestoken, dan geven de wetten, formules, of kernbetrekkingen de onderlinge relaties goed weer. Zie bijvoorbeeld het schema in fig. 3.2 van de theorie voor vallende lichamen. In het schema zijn *niet* de theoretische termen, maar de *wetten* in de theorie uitgezet; omdat iedere theoretische term met alle overige termen relaties heeft, zou een schema daarvan zinloos zijn. Daarom is in dit soort gevallen beter te kiezen voor het in schema zetten van de wetten of formules, waarbij de onderlinge relaties tussen de wetten bijvoorbeeld de volgorde van afleiding, historische ontwikkeling of 'logische' opbouw van de theorie kunnen zijn. Eigenlijk staan er in het schema van fig. 3.2 dan ook twee soorten relaties afgebeeld. Allereerst zijn er de relaties zoals door de formules weergegeven, en daarnaast zijn er de relaties tussen de formules. De formules in fig. 3.2 zijn uiteraard op meerdere manieren in schema te zetten; het gekozen schema vertrekt vanuit de eenvoudige formule dat snelheid een functie is van zwaartekracht en verstreken tijd, en omdat de snelheid na t seconden gelijk is aan de afgeleide van het verticaal afgelegde aantal meters h naar t , moet $h = 0,5 g t^2$ zijn (de afgeleide van $0,5 g t^2$ naar t is immers $g t$). De derde formule is een toegift; zij volgt onmiddellijk uit de tweede formule.

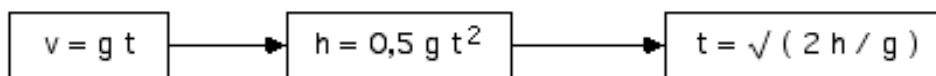


Fig. 3.2. Schema van de theorie voor vallende lichamen; v = bereikte snelheid, g = versnelling, t verstreken tijd, h = afgelegde verticale afstand.

De theorie over vallende lichamen is in het onderwijs op verschillende manieren te behandelen: studenten leren omgaan met de formules; de theorie leren toepassen op verbaal beschreven concrete situaties; in practicum-situaties of vanaf de toren van Pisa de theorie toepassen; de geschiedenis van de theorie bestuderen (Heeft Newton dat verhaal van die appel verzonnen?); de veronderstellingen van de theorie kritisch analyseren (bijv. dat de val in vacuüm plaatsvindt). Al naar gelang de wijze waarop de theorie in het onderwijs aan de orde is geweest, worden geschikte vragen geformuleerd. Met het in schema zetten van de termen waarop in het onderwijs de schijnwerpers gericht worden, is de leerstofinventarisatie volledig, althans wat betreft de opsomming van de onderwerpen uit de leerstof de termen en hun onderlinge relaties. Dan is nog niet aangegeven wat er over iedere term of relatie in de toets gevraagd kan en mag worden. Een poging om dat te formuleren is al bijna een poging om de vragen te formuleren die (binnen de onderwijsdoelen) over die bepaalde term of relatie gesteld zouden kunnen worden. Ik geef er daarom de voorkeur aan de leerstofinventarisatie niet verder uit te werken en beperkt te houden tot de opsomming van de termen en de schematisering van hun onderlinge relaties, om uitgaande hiervan direct over te gaan tot het schrijven van toetsvragen met behulp van de vuistregels die de volgende hoofdstukken aanreiken. Maar daaraan voorafgaand wil ik speciaal aandacht schenken aan de wijze waarop termen gedefinieerd of beschreven kunnen worden, omdat definities zo wel in het onderwijs als in de toets veelvuldig gebruikt worden en dat gebruik jammer genoeg niet altijd even adequaat is.

3.5 Varianten van 'definities'

'Het leren van nieuwe termen is geen alles-of-niets aangelegenheid - hoe meer we over een term weten, des te meer weten we over de relaties tussen deze term en andere termen. Er is geen glossarium in dit boek opgenomen, omdat ik het naïeve en foutieve idee niet wil aanmoedigen als zou je heel wat over een term weten wanneer je een definitie van één of twee zinnen kunt geven. Deze tekst onderwijst de betekenis van termen door ze telkens weer te gebruiken in uitspraken die

deze termen relateren aan andere termen.' Zo spreekt een vooraanstaand onderzoeker op het gebied van de cognitieve psychologie zijn lezers toe in het voorwoord van het door hem geschreven tekstboek (Wickelgren, 1979). Daarmee is zowel fraai samengevat wat in voorgaande paragrafen besproken is, als kort aangeduid wat de upshot is van de nu volgende behandeling van 'de definitie'.

De student die een correcte definitie geeft voor een gevraagde term, laat daarmee zien dat hij de betreffende term 'begrijpt' en op de juiste wijze weet te gebruiken. Een correcte definitie kunnen geven is te zien als een *criterium* voor het begrijpen van de betreffende term. Het is deze gedachtengang die docenten ertoe brengt zoveel definities in hun onderwijs en in de toetsing ervan te stoppen. Die neiging wordt nog versterkt door het didactisch nut van definities; een geschikte definitie maakt soms in enkele woorden duidelijk wat anders alleen op een omslachtige wijze te onderwijzen zou zijn. Laat er geen misverstand over bestaan dat ik het hier gestelde volledig kan onderschrijven. Maar er zijn wel een paar problemen die deze heldere soep vertroebelen.

Allereerst is het kunnen geven van een correcte definitie geen *voldoende* voorwaarde voor de conclusie dat de student de betekenis van de betreffende term kent. Bij verder onderzoek kan blijken dat ondanks een correct gegeven definitie de term vervolgens toch onjuist gehanteerd wordt. De student kan de definitie op domme wijze uit het hoofd geleerd hebben. Het laatste is heel wat ernstiger dan het op het eerste gezicht lijkt. Immers, wie in onderwijs en toetsing de nadruk legt op definities, zet een premie op het uit het hoofd leren van die definities. Hoe onbedoeld dat ook is, het is rampzalig voor de studenten omdat dit soort 'kennis' van nul en generlei waarde is (zie 2.6).

Daarnaast valt te bedenken dat het kunnen geven van een correcte definitie niet eens een *noodzakelijke* voorwaarde is voor het kennen van de betekenis van de betreffende term. Vraag een psycholoog om een definitie van de term intelligentie; of een wiskundige om een definitie van de term getal. Ik heb eerder al beklemtoond dat niet alle wetenschappelijke termen zich laten definiëren. Is het dan misschien zo, dat de docent zich maar beter kan matigen in het 'per definitie' introduceren van nieuwe termen?

De term definitie laat zich niet via kenmerken definiëren. De term definitie staat voor wat Wittgenstein een *family resemblance concept* noemt (ik kom daar nog op terug). Om de betekenis van de term definitie in te vullen, kan ik niets beter doen dan een aantal mogelijke vormen van definitie behandelen.

Definitie via kenmerken. Wat zou het niet prachtig zijn wanneer men bij iedere nieuwe term in de definitie ervan precies aan zou kunnen geven op welke objecten, gebeurtenissen of wezens de term van toepassing is. Dat betekent dat in de definitie een opsomming wordt gegeven van de kenmerken waaraan een object, gebeurtenis of wezen moet voldoen, wil de term erop van toepassing zijn. De opgesomde kenmerken moeten dan zowel noodzakelijk als voldoende zijn. Om de kenmerkenlijst in lengte te beperken, kan voor de definitie de bekende vorm worden gekozen dat erin wordt aangegeven om welke bekende soort het gaat om vervolgens via specifieke kenmerken aan te geven op welke objecten enz. van deze soort de gedefinieerde term van toepassing is (definitie *per genus et differentiam*). Een olifant is een zoogdier met flaporen; het soort: zoogdier, en specifiek kenmerk: flaporen. Het streven van veel docenten, studieboekschrijvers en filosofen is om alle termen op deze wijze te definiëren. Natuurlijk, dat streven loopt spaak op de onmogelijkheid alle termen op deze wijze te definiëren zonder in de definitie van een term gebruik te maken van termen die zelf gedefinieerd zijn met behulp van die term. Die moeilijkheid wordt omzeild door een of meer termen ostensief te definiëren (zie hieronder). Relevanter is de constatering dat het genoemde streven bijzonder gekunstelde resultaten oplevert, zelfs al zou het lukken om een wetenschappelijke theorie op deze wijze te 'definiëren' (zie Hempel 1952, of Stegmüller 1970).

Definitie door aanwijzen (de ostensieve definitie). 'Dit is rood' (wijzend op een boek met een rood omslag). Deze wijze van definiëren laat zich in het onderwijs uitstekend gebruiken. In het wetenschappelijk onderwijs wordt de student in het onderwerp, het te gebruiken instrumentarium of het veld van onderzoek ingeleid door rondleiding (in de kliniek), presentaties in het laboratorium, excursies in het veld enz. (Kaplan, 1964). Het laat zich raden dat het alleen de (indirect) waarneembare

termen zijn die zich op deze wijze laten definiëren, en misschien abstracte termen. Het is jammer dat de status van deze wijze van definiëren laag is, ook al is ze onmisbaar als (axiomatisch) uitgangspunt voor degenen die uitsluitend met definities via kenmerken zouden willen werken. Wittgenstein heeft veel aandacht besteed aan de ostensieve definitie (zie Baker en Hacker, 1980, hfdst. 2). Ook het aanwijzend kunnen definiëren van een term is een *criterium* voor het kennen van de betekenis van de term. In de toetsing kan dat blijken uit het aanstrepen van het juiste voorbeeld bij een meerkeuzevraag “Welke van de vijf alternatieven noemen we rood?”

Definitie door voorbeelden en niet voorbeelden. Een speciaal geval van de definitie door aanwijzen is het aanwijzen van voorbeelden en niet-voorbeelden.

In het werken met voorbeelden en niet voorbeelden wordt de nadruk gelegd op de begrenzing van het bedoelde begrip, op wat er nog wel en wat er niet meer onder bedoeld wordt. Het is niet zinvol om de betekenis van de term rood te definiëren door naast een boek met rode omslag ook boeken met allerlei andere kleuren omslag aan te wijzen. Het kan wel zinvol zijn wanneer het gaat om fijne nuances in de waarneming, zoals het onder de microscoop onderscheiden van gezond en ziek celweefsel. Deze wijze van definiëren leent zich eveneens goed voor de termen die door de student nogal eens te breed (overgeneralisatie) of te smal (ondergeneralisatie) gebruikt worden. De soort toetsvraag die hierbij past laat zich makkelijk denken: uit een lijst voorbeelden en niet voorbeelden deze als zodanig ook kunnen identificeren. Er zijn de laatste jaren nogal wat boeken verschenen die begripsmatig toetsen propageren, waarin deze wijze van definiëren nogal eenzijdig (als ‘ideaal’) naar voren geschoven wordt. Het probleem is dat slechts weinig termen zich bij uitstek voor deze wijze van definiëren lenen (zie o.a. Merrill en Tennyson, 1977).

Definitie door opsomming. Ook een variant op het definiëren door aanwijzen. In rechttoe-rechtaan gedaante kun je zo de letters van het alfabet definiëren door ze alle op te sommen.

Is deze wijze van definiëren ook bruikbaar voor de natuurlijke gehele getallen? Dat roept filosofische problemen op. Misschien kun je een *regel* te formuleren, waarmee in beginsel alle

natuurlijke gehele getallen geconstrueerd kunnen worden, dan zou zo'n regel ook als definitie door opsomming opgevat kunnen worden.

De *nominale definitie* introduceert een verkorte notatie, een naam of een nieuwe term voor een gegeven uitdrukking. De nieuwe term heeft dezelfde betekenis als de uitdrukking die hem definieert, wat soms aangegeven wordt door 'is per definitie gelijk aan' of '=Df'. Hempel (1952) geeft als voorbeeld: 'Americium =Df het element met 95 nucleaire protonen'. Nominaal gedefinieerde termen kunnen altijd zonder verlies aan betekenis vervangen worden door hun definiërende uitdrukking. De definitie via kenmerken is een speciaal geval van nominaal definiëren. De nominale definitie kan gebruikt worden uit het oogpunt van economie: een langere uitdrukking wordt door een enkel woord vervangen. Het kan echter ook zijn dat een bepaald soort verschijnsel, ding of wezen zich dermate profileert in de waarneming dat je kunt spreken van een nieuw begrip, waar dan ook een nieuwe term aan gekoppeld wordt.

De *reële definitie* is een soort tegenhanger van de nominale definitie; hij geeft alleen de kenmerken die de essentie van het bedoelde begrip aanduiden: 'Een gletsjer is een langzaam bewegende ijsmassa', 'Een stoel is een meubel om op te zitten'. Reële definities geven een bepaalde essentie van het bedoelde begrip weer, maar dekken daarmee niet volledig de betekenis van de term. Een zelfde term kan meerdere verschillende reële definities hebben. Het is niet altijd zinvol om in een tekst een term te vervangen door een reële definitie van die term. De student die een reële definitie van een term geeft, geeft daarmee aan ten minste deze wezenlijke betekenis van de term te kennen.

De *conditionele definitie* is een bijzonder soort definitie via kenmerken: het bijzondere zit in de gegeven specificatie van de *omstandigheden waaronder deze kenmerken waarneembaar zijn*. Bijvoorbeeld eigenschappen als elastisch, magnetisch, splijtbaar, meegaand. Deze wijze van definiëren, hoewel vaak onvermijdelijk, schept problemen omdat strikt genomen de definitie niet identiek gesteld kan worden aan het met

de term aangeduide begrip: dat is echter niet zozeer een praktisch als wel een filosofisch probleem (Stegmüller 1970).

De *operationele definitie* is weer een speciaal soort conditionele definitie, waarbij de nadruk valt op de omschrijving van de condities: de operaties of handelingen die nodig zijn om (in het kader van een experimenteel onderzoek) bepaalde condities te bewerkstelligen. Vooral in de sociale wetenschappen zijn hierover wel eens extreme standpunten ingenomen; algemeen bekend is de 'operationele' definitie van intelligentie als datgene wat door de intelligentietest gemeten wordt. De Groot kiest in zijn *Methodologie* een positie die daar nogal dicht bij zit (De Groot 1962, blz 88). Een operationele definitie kan soms in het jasje van een volledige opsomming gegoten worden: De Groot ziet de toetsvragenverzameling als operationele definitie van de doelstellingen van het onderwijs (in De Groot en Van Naerssen, 1973, blz. 38).

Men kan ook zeggen dat de operationele definitie een bijzondere vorm is van de

Stipulatieve definitie of definitie bij fiat. Deze definitie legt de betekenis van een term die in een college, tekstboek of onderzoek gebruikt gaat worden voor de duur daarvan vast en claimt geen algemene geldigheid.

Het hoeft niet altijd een gelegenheidsdefinitie te zijn: het kan ook gaan om een poging een term een nieuwe en beter omschreven betekenis te geven dan die waarin hij tot dan toe gehanteerd werd. Hempel (1952) noemt als voorbeelden de theorie van Frege-Russell (getallenleer), en Tarski's semantische definitie van waarheid. Het op deze wijze definiëren van termen die overigens nogal vaag in betekenis zijn, heeft alleen zin, wanneer er met de aldus gedefinieerde term vruchtbaar theoretisch werk verricht kan worden. Het kan en mag geen truc zijn om theoretische termen van een (pseudo-) definitie te voorzien. De definitie bij fiat is in haar meer bescheiden variant een werkdefinitie, die het mogelijk moet maken om de zo gedefinieerde term zinvol te gebruiken. Een dergelijke definitie behoort waarschijnlijk niet tot de onderwijsdoelen: het is niet de bedoeling dat de student deze definitie 'kent', alleen dat hij er voor bepaalde andere doeleinden mee kan werken.

Familiegelijkenis. In empirische wetenschappen wordt men nogal eens geconfronteerd met termen die zich niet lenen voor een definitie via kenmerken omdat de verschijnselen, dingen of wezens die eronder begrepen worden geen gemeenschappelijke kenmerken hebben. Wat is er bijvoorbeeld gemeenschappelijk aan de verschillende vormen van definitie die hierboven beschreven zijn? Wittgenstein heeft door zijn bespreking ervan in de *Filosofische onderzoekingen* de term familiegelijkenis befaamd gemaakt; hij geeft in de paragrafen 68 en 69 van zijn boek als voorbeelden de termen getal en spel. Verschillende spelen hebben niet een bepaalde set kenmerken met elkaar gemeen; er is wel een lijst kenmerken op te stellen waaraan tenminste sommige spelen voldoen. Je kunt dan zeggen dat verschillende spelen op elkaar lijken op dezelfde wijze waarop leden van eenzelfde familie op elkaar lijken: ze lijken allemaal op elkaar, maar niet telkens op dezelfde wijze, niet telkens op grond van dezelfde kenmerken. Hetzelfde geldt voor natuurlijke, rationele, complexe, kardinale enz. getallen. En voor nominale, reële, aanwijzende enz. definities. Gevraagd om een familiegelijkenis-term te 'definiëren', zal de student een aantal voorbeelden kunnen geven; een tweede student kan voor dezelfde term een geheel andere verzameling voorbeelden geven; beiden kunnen toch een goede 'definitie' van de term gegeven hebben.

Beschrijven. Termen kunnen hun betekenis in hoge mate of zelfs volledig (theoretische termen!) ontlenen aan hun relaties tot andere termen. Een beschrijving waarin een of meer van dergelijke relaties aangegeven worden, kan dan ook als een vorm van definiëren opgevat worden, al is het handiger om er de term beschrijven voor te blijven gebruiken. Een beschrijving kan correct zijn zonder de hele betekenis van de term te dekken. Vele termen hebben overigens niet een dermate scherp afgepaalde betekenis dat ze in een beperkt aantal van dergelijk beschrijvende uitspraken te vangen zouden zijn.

Kaplan (1964) geeft een paar aardige voorbeelden. Wil je erachter komen wat Marx met de term klasse bedoelt, dan is de enige begaanbare weg kennis te nemen van het hele werk van Marx. In de astronomie wordt een bepaald soort witte ster dwerg genoemd; de eigenschappen van witte dwergen zijn vastgelegd in een aantal verschillende astrofysische theo-

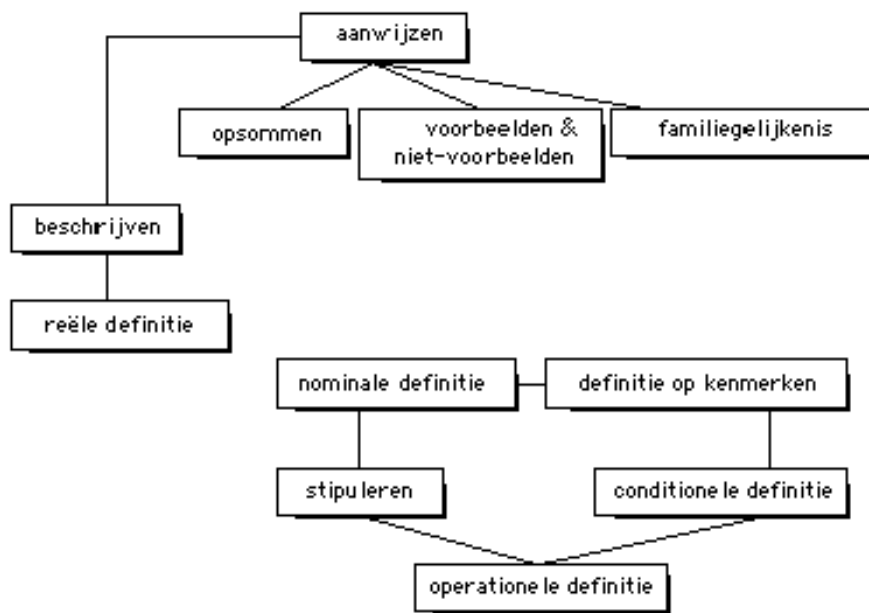


Fig. 3.3. Schema van besproken mogelijkheden voor het definiëren van termen.

riën, bijvoorbeeld die over de evolutie van de sterren. Aan deze theorieën kunnen talrijke beschrijvingen voor witte dwergen ontleend worden, maar geen behoorlijke definitie.

De gegeven lijst van mogelijke vormen van definities is niet uitputtend. Dat is ook niet nodig. Hopelijk is tussen de regels door ook duidelijk gemaakt dat dezelfde term op meerdere manieren gedefinieerd kan worden, ook binnen één tekstboek. Van verschillende mogelijke definities kan in het onderwijs didactisch handig gebruik worden gemaakt, en eventueel kan bij het schrijven van toetsvragen daar ook enige inspiratie uit geput worden.

Tenslotte is het goed er op te wijzen dat geen enkele definitie het probleem van de *grensgevallen* kan oplossen, dat zijn die gebeurtenissen, dingen enz, waarvan het niet duidelijk is of ze nog juist wel of net niet meer onder de term-zoals-gedefinieerd

vallen. Bij het schrijven van toetsvragen zijn juist deze grensgevallen mogelijke spelbrekers: de verleiding bestaat om voor wat 'moeilijker' vragen dicht in de buurt van een grensgeval te gaan zitten, en dat brengt het risico met zich dat vragen dubbelzinnig worden en zelfs door collega's niet meer correct beantwoord kunnen worden.

3.6 Literatuur

SMITH, E. E., & D. L. MEDIN, *Categories and concepts*, Cambridge, Massachusetts, Harvard University Press, 1981.

Geeft een overzicht van de cognitief-psychologische theorievorming rond het leren van en omgaan met (natuurlijke) categorieën en begrippen, ofwel het omgaan met de termen die deze categorieën en begrippen labelen. Een psychologische onderbouwing zou vooral op dit werk kunnen steunen.

HEMPEL, C. G., *Fundamentals of concept formation in empirical science*, London, The University of Chicago Press, 1972 (1952).

Er is in wetenschapsfilosofische hoek weinig aandacht voor zaken die slechts zijdelings betrekking hebben op vragen rond opkomst en ondergang van wetenschapsparadigma's. Hempel is een van de uitzonderingen: in dit kleine boekje behandelt hij de wijze waarop de termen in een wetenschappelijke theorie functioneren, hoe zij zich ontwikkelen in hun betekenis, definitieproblemen e.d.

KAPLAN, A., *The conduct of inquiry; methodology for behavioral science*, San Francisco: Chandler, 1964.

Naast het boek van Hempel is dit een uitstekende inleiding, vooral wat betreft het onderscheid tussen waarneembare, abstracte en theoretische termen.

STEGMÜLLER, W., *Probleme und Resultate der Wissenschaftstheorie und Analytische Philosophie. Band II: Theorie und Erfahrung*, Heidelberg: Springer-Verlag, 1970.

Geeft een zeer grondig overzicht van wetenschapstheoretisch werk met betrekking tot het definiëren van termen en de aard

van waarneembare, indirect waarneembare en theoretische termen. Bestrijkt daarmee hetzelfde terrein als Hempel (1952) en Kaplan (1964), maar is minder inleidend van aard.

BREUKER, J., *In kaart brengen van leerstof*, Aula 801, Utrecht: Het Spectrum, 1980.

Behandelt het schematiseren als methode om globale samenhangen in de leerstof in kaart te brengen. Geeft naast een 'minicursus schematiseren' ook enkele theoretische achtergronden bij de schematiseermethode.

MIRANDE, M. J. A., *Studeren door schematiseren*, Aula 805. Utrecht: Het Spectrum, 1981.

Presenteert het schematiseren als studietechniek. Gaat evenals het boek van BREUKER uit van een gegeven tekst. Een volledige leerstofinventarisatie behelst de leerstof zoals die onderwezen en getoetst wordt, en dat kan verder gaan dan wat er letterlijk in te bestuderen teksten staat; maar ook dan blijft de schematiseermethode bruikbaar om geïnventariseerde termen in hun relevante onderlinge relaties te kunnen afbeelden.

MICHENER, E. R., 'Understanding understanding mathematics', *Cognitive Science*, 1978, 2, 361-383.

Bespreekt het karakter van wiskundige leerstof, weliswaar in psychologische termen (begrijpen begrijpen). Een boeiende case study, met mogelijk ook voor andere vakgebieden relevante ideeën.

WITTGENSTEIN, L., *Filosofische onderzoekingen*, Meppel: Boom, 1976 (1953).

BAKER, G. P., & P. M. S. HACKER, *Wittgenstein: Understanding and meaning; an analytical commentary on the Philosophical Investigations*, (Volume 1), Oxford: Basil Blackwell, 1980.

Het (late) werk van Wittgenstein is direct relevant voor een aantal van de onderwerpen die in hoofdstuk 3 zijn aangesneden. Dat geldt vooral de vraag hoe termen gedefinieerd kunnen en mogen worden, maar ook de vraag waaruit blijkt of de betekenis van een bepaalde term 'begrepen' is (door de student bijvoorbeeld). Niet dat Wittgenstein hierop antwoorden geeft die je 'onderwijskundig' zou kunnen noemen: veeleer is het zo dat hij de grenzen afbakt waarbinnen de onderwijs-

kundige (en met hem de docent) zich zinvol kan bewegen. Het aforistische karakter van zijn werk, in dit geval de *Filosofische onderzoekingen*, maakt het erg toegankelijk, hoe prettig het op zich ook te lezen is. Baker en Hacker doen de eerste grootscheepse poging om het werk van Wittgenstein zijn systematische ordening te geven, juist en vooral in relatie tot het werk van Frege en Russell, en zijn eigen vroege werk, de *Tractatus*.

De hier genoemde filosofische literatuur wordt niet zonder reden vermeld. Hoofdstuk 3 bevat de grondslagen waarop vuistregels in hoofdstuk 4 e.v. rusten. Omdat het in het onderwijs gaat om de betekenis van termen en relaties, en hele theorieën, valt er niet aan enige (filosofische) reflectie te ontkomen voor wie zich voor de vraag gesteld ziet hoe het begrepen hebben van 'de betekenis' van termen enz. zinvol getoetst zou kunnen worden.

Dat is wat anders dan de hier niet aangehangen opvatting dat je pas zou kunnen beginnen met onderwijzen en toetsen wanneer het werk van Wittgenstein verwerkt en begrepen zou zijn!

4 TOETSVRAGEN SCHRIJVEN BIJ AFZONDERLIJKE TERMEN

‘De “termen” vormen het specifieke vocabulaire van het vak waarin de student zich heeft te verdiepen. Het gaat erom dat termen gedefinieerd kunnen worden, dat voorbeelden herkend kunnen worden, dat juist of onjuist gebruik van een term herkend wordt, of dat synoniemen herkend worden. Deze categorie vormt daarmee het laagste of eenvoudigste niveau van cognitieve beheersing in de Taxonomie.’ Aldus Bloom, Hastings en Madaus in hun *Handbook on formative and summative evaluation of student learning* (1971, blz. 119). Dit is een wijdverbreide en ongelukkige opvatting. De suggestie is dat het bij opgaven over afzonderlijke termen gaat om de bouwstenen van de leerstof, ‘basics’, en dat je dat stadium zo snel mogelijk achter je moet laten om met de echt belangrijke problemen te beginnen, zoals het oplossen van problemen. Daarmee wordt eraan voorbijgegaan dat aan groundbegrippen in iedere tak van de wetenschap nogal eens onvermoede en diepgaande problemen vastzitten. Vele wetenschappen hebben een grote worsteling gekend, of zitten daar nog middenin, om hun terminologie te ontwikkelen tot een helder, eenduidig, en werkzaam geheel; het zou kortzichtig zijn in het onderwijs die cognitieve worsteling als een vanzelfsprekendheid te beschouwen, en het voor te stellen alsof het voor de student makkelijk moet zijn zich de bestaande terminologie eigen te maken. Van Naerssen (1980, p. 193) drukt het zo uit: ‘Tenslotte hebben auteurs als Bloom et alii (1956) er hunne toe bijgedragen om door het creëren van een waardehiërarchie van items de mening te doen postvatten dat ordinaire kennisitems vervangen moeten worden door moeilijk te construeren vragen, die vooral inzicht, analyse, synthese en evaluatie zouden moeten meten. Maar met dit laatste meet men misschien juist weer hoofdzakelijk intelligentie, waar het niet om gaat bij de controle of de student zijn best gedaan heeft om de stof te bestuderen.’

Ik nodig de lezer hierbij dan ook uit om juist aan deze eenvoudige vragen over de stof ernstig aandacht te besteden. Niet alleen omdat ook kennisvragen met overleg en inzicht geschreven moeten worden (het zijn juist de misplaatste kennisvragen die het toetsen van ‘kennis’ in een ongunstig

daglicht hebben gesteld), maar ook omdat verschillende mogelijkheden voor het schrijven van vragen over termen in de volgende hoofdstukken telkens terugkomen.

De paragrafen zijn ingedeeld naar vertalen, definiëren, voorbeelden geven, voorbeelden herkennen en benoemen, herkennen en benoemen bij formeel gedefinieerde termen, en beschrijvende uitspraken.

4.1 VERTALEN

Het prototype van vertalen van woorden doet zich voor in het onderwijs in een vreemde taal: gegeven een woord in de eigen taal wordt een equivalent daarvan in de vreemde taal gevraagd, en omgekeerd. Het prototype van het geven van een synoniem doet zich voor in het onderwijs in de moedertaal. Het omzetten van een term in een symbool, en omgekeerd, kennen we bij uitstek in de scheikunde en in de wiskunde.

Het vertalen van een woord, term of symbool verschilt niet wezenlijk van het benoemen van voorbeelden, het geven van voorbeelden en het geven van een verbale definitie of beschrijving. Vertalen komt in enigerlei vorm in vrijwel alle onderwijs voor.

De betekenis van het kunnen vertalen is duidelijk: het gaat veelal om een basisterminologie die behoorlijk beheerst moet worden voordat op efficiënte wijze met de bestudering of oefening van de leerstof begonnen kan worden. Wat niet wegneemt dat sommige vertaalproblemen moeilijk of misschien in het geheel niet oplosbaar zijn, zodat ook bij vertaalopgaven doorgaans veel gradaties van moeilijkheid mogelijk zijn.

Een tamelijk fundamentele tweedeling in vraagsoorten is of een woord, term of symbool in *context* gegeven wordt of *zonder context*. Dat is het verschil tussen het vragen van de vertaling voor het woord zoals dat in een gegeven zin of tekst gebruikt wordt, en het vragen van de vertaling van datzelfde woord zonder meer.

Zonder context. Mogelijke varianten bij het vragen van vertalingen enz. van woorden die gepresenteerd worden zonder een context waarin ze gewoonlijk gebruikt worden zijn o.a.:

- *mondeling* presenteren van woorden of termen; gevraagd wordt de *spelling*
- *mondeling* presenteren van woorden; gevraagd wordt een *vertaling* in de eigen taal of in de vreemde taal
- *schriftelijk* presenteren van woorden waarvan een *vertaling* in de eigen of in een vreemde taal gegeven moet worden
- voor een gegeven woord of term wordt een *synoniem* gevraagd
- voor een gegeven term wordt de *formule* gevraagd (scheikundige termen!)
- voor een gegeven formule wordt de bijbehorende *term* gevraagd
- voor een gegeven term wordt het daarvoor gebruikte *symbool* gevraagd
- voor een gegeven symbool wordt de bijbehorende *term* gevraagd

Kortom, deze vraagsoorten zijn wel bekend. Let er op dat het aantal te geven antwoorden precies aangegeven wordt.

AFRADER

Geef zoveel mogelijk goede synoniemen voor 'ontaard'.

Deze vraag levert bij het scoren grote problemen op, omdat erop gerekend kan worden dat er heel veel antwoorden komen die op z'n minst twijfelgevallen zijn; dat mag men zelfs al verwachten wanneer er precies twee synoniemen gevraagd worden, en ook in dat geval zal het niet meevallen de vraag te scoren.

De voor de hand liggende vraagvorm is hier de aanvulvraag, omdat de opstelvraag hier niet passend is en de meerkeuzevraag doorgaans is af te raden. De redenen om geen meerkeuzevragen te gebruiken zijn: de afleiders zijn geen alternatieven waaruit de student ook zou kiezen wanneer de vraag in open vorm gesteld zou zijn; de student is niet onderwezen in het vermijden van het soort foute antwoorden dat als afleiders gebruikt wordt; het herkennen van een juiste vertaling is niet het (eerste) doel van het onderwijs, dat is het kunnen *produceren* van de juiste vertaling.

AFRADER

De juiste vertaling voor 'development' is:

- a. afgraving
- b. afdaling
- c. ontvluchting
- d. ontwikkeling

Alleen ingeval foutieve vertalingen in het onderwijs aan de orde zijn geweest, kan de meerkeuzevorm overwogen worden: het is immers de bedoeling dat de student de fouten in de afleiders ook als zodanig kent. Probleem blijft dan nog dat studenten die niet opgeleid worden tot corrector, dat het *herkennen* niet past op het onderwijsdoel dat vraagt om een correcte *produktie*. Bedenk dat bij deze vertaalopgaven de student inderdaad op de geboden alternatieven afgaat, en niet eerst het volgens hem of haar juiste antwoord zal bedenken om vervolgens te kijken of het ook een van de alternatieven is.

In veel gevallen kan een *complete* vragenverzameling gemaakt worden of tenminste een vragenverzameling die alle belangrijke gevallen 'dekt'. Denk aan de vreemde-woorden-lijst die studenten verondersteld worden te kennen, of aan de lijst van scheikundige stoffen waar de student aan de formule van moet kunnen uitschrijven.

VOORBEELD

Geef bij iedere formule de bijbehorende naam:

1. MgCO_3
2. CaF_2
3. $\text{Al}_2(\text{SO}_4)_3$

Een uitgebreide lijst van dergelijke scheikundige namen met hun formules is veel te omvangrijk om het lonend te maken alle koppels in beide richtingen uit het hoofd te leren: een naam of formule moet volgens eenvoudige regels *geconstrueerd* worden. Bij het leren van vreemde woorden ligt het iets anders: je zou haast zeggen dat het er daar juist om gaat dat zo'n lijst uit het hoofd wordt geleerd. In ieder geval lokt het toetsen van vreemde-woordenkennis *zonder context*, dus buiten zinsverband om, het uit het hoofd leren van woorden uit. Het is niet onmogelijk dat een onbedoeld gevolg daarvan is dat de student moeite zal hebben woorden in zinsverband correct

te vertalen, ook al worden die woorden perfect beheerst wanneer ze buiten zinsverband gevraagd worden. Het is geen eerlijke analogie, maar je kunt hier denken aan het beheersen van de schoolslag op het droge, en in het water toch zinken als een baksteen. Het hier bedoelde wel-weten-maar-niet-toe-kunnen-passen is een risico dat zich vooral kan voordoen bij woorden die een verschillende vertaling hebben afhankelijk van het zinsverband, van de context waarin ze gebruikt worden. In die gevallen is het een goede stelregel voor de student om de vertaling in zinsverband te oefenen, en voor de docent om de vertaling in zinsverband te vragen. Een en ander geldt *mutatis mutandis* ook voor vragen die synoniemen betreffen. Er zijn nogal wat leestests die synoniemen vragen voor woorden zonder context. Het is voor dergelijke tests een open vraag of wat zij meten ook geldigheid bezit voor de leesvaardigheid in praktische situaties, waar de leerling uit de betekenis en de syntax van een zin aanwijzingen krijgt over de betekenis van een bepaald woord (Farr, 1969, blz. 37).

In context. Wanneer het zinsverband van het woord in de toetsvraag wordt betrokken, veranderen de toetsingsmogelijkheden. Die mogelijkheden laten zich in tweeën delen, en dat ligt voor de hand: het vragen van een correcte vertaling of synoniem, of vragen een woord uit de eigen taal te 'gebruiken' in een zelf te geven zin in de vreemde taal. Deze door de student te bedenken zin mag best een heel simpele vorm hebben, en de student moet ook aangemoedigd worden om het zo eenvoudig mogelijk te houden wanneer het alleen om het juiste gebruik van het gevraagde woord te doen is.

Concrete uitwerking van dit soort toetsvragen hoeft geen problemen op te leveren. Wanneer een reeks van dergelijke vragen gesteld wordt, kan het bedoelde woord door onderstrepen worden aangeduid (wanneer vertaling naar de eigen taal gevraagd wordt). Het uitschrijven van toetsvragen wordt nu wel omslachtiger, omdat ook geschikte zinnen bedacht moeten worden. Bedenk dat er ook nogal wat van dergelijke vragen nodig zijn om in het onderwijs te gebruiken, als oefenmateriaal voor de student. Bij het schrijven van deze toetsvragen zal het doorgaans mogelijk zijn om handig te werk te gaan, bijvoorbeeld door geen concrete zinnen te schrijven, maar syntactische *rompzinnen* waarin bepaalde woorden ingevuld

kunnen worden, te kiezen uit een daartoe opgestelde woordenlijst. Die woordenlijsten kunnen natuurlijk gebruikt worden voor een aantal verschillende rompzinnen. Zo'n rompzin is een speciaal soort rompvraag (2.5). Dezelfde rompzin kan voor het toetsen van meer dan een woord gebruikt worden. Bij het vertalen van eigen in vreemde taal kan gebruik worden gemaakt van zinnen in de vreemde taal, waarin de plaats voor het te vertalen (en ook gegeven) woord is opengelaten.

Een toets met een aantal van dit soort zinnen kan een saaie toets worden. Er is niets op tegen om te werken met een samenhangende tekst, waaruit alleen van onderstreepte woorden de vertaling gevraagd wordt, respectievelijk in opengelaten plaatsen de vertaling van gevraagde woorden moet worden ingevuld. Bijna een proefvertaling, zou je kunnen zeggen, maar deze opgave kost de student aanzienlijk minder tijd dan het volledig uitschrijven van een vertaling van de tekst. Dat laat onverlet dat syntactisch juist vertalen bij voorkeur te toetsen is door de gewone proefvertaling.

In de laatste alinea's is alleen het vertalen aan bod gekomen, en ik heb aangenomen dat duidelijk is dat een en ander evenzeer geldt voor het vragen naar synoniemen. Maar ook de docent die te maken heeft met het vertalen naar *symbolen*, en terug, kan met context in de boven bedoelde zin te maken hebben. Is dat het geval, dan kan op analoge wijze gezocht worden naar toetsvragen waarin ook de context een rol speelt.

Ook bij contextafhankelijke toetsvragen is het aan te raden de meerkeuzevraag te vermijden. Eventueel kan de meerkeuzevraag gebruikt worden in die gevallen waarin de student geleerd heeft bepaalde vertalingen in een bepaald zinsverband juist *niet* te maken.

4.2 Definiëren

Een heel speciale vorm van vertalen is het geven van een definitie. Dat geldt voor alle soorten definities (zie 3.5), maar in het bijzonder voor de nominale definitie, waar een verkorte notatie, een naam of een nieuwe term gelijkgesteld wordt aan een gegeven uitdrukking. Op alle plaatsen in een tekst, waar de betreffende uitdrukking voorkomt, is deze te 'vertalen'

naar de bijbehorende term (en omgekeerd). Een vak als wiskunde is sterk in het op deze wijze vertalen en daardoor vereenvoudigen van uitdrukkingen. Bij het aanpakken van problemen is het een goede gewoonte te zoeken naar een formulering voor de probleemstelling waarin bepaalde onderdelen bekend zijn, en waar een naam, een term of een formule voor gesubstitueerd kan worden. Dat geldt ook voor het determineren van planten en dieren; men zou kunnen zeggen dat het zoeken is naar een dusdanige beschrijving van het specimen dat bekeken wordt, dat het probleem stapsgewijs tot de oplossing gebracht kan worden doordat soort, familie enz. 'herkenbaar' worden. Het op deze wijze kunnen omgaan met definities behoort voor sommige vakken tot de wezenlijke doelen van het onderwijs. Maar ook wanneer definities slechts een ondergeschikte rol hebben in het leerstofpakket, is het niet ongebruikelijk om in de toets relatief veel definitievragen op te nemen.

De student die een goede definitie voor een gevraagde term weet te geven, laat daarmee zien dat de betekenis van die term begrepen wordt. Toch kan bij verder onderzoek blijken dat de student, ondanks zijn kennis van de definitie, niet in staat is er in een concreet probleem mee te werken, of dat de definitie het enige is dat hij over de term weet. Worden definitievragen opvallend veelvuldig gesteld, dan zet de docent daarmee een premie op het uit het hoofd leren van definities, en dat gaat ten koste van de aandacht voor toepassingen en voor de bredere betekenis van termen. Definitievragen hoeven niet schadelijk te zijn, maar gebruik ze met mate.

Het is niet vanzelfsprekend dat de student voor alle behandelde termen een definitie moet kunnen geven: kies uit de lijst termen in de leerstofinventarisatie de termen waarvoor het kennen van een definitie wèl van belang is. Schrijf die definitie (eventueel meerdere definities voor dezelfde term) uit, en maak daar een lijst van waar studenten mee kunnen werken. Voeg daar ook voorbeelden en aanwijzingen aan toe over de wijze waarop naar deze definities gevraagd kan worden (zie hieronder).

Veel termen laten meerdere verschillende definities toe. In het onderwijs kunnen verschillende definities voor dezelfde term bruikbaar zijn. Didactisch kunnen ze worden uitgebuit door

bij de behandeling van een bepaald begrip te beginnen met ostensieve definities, voorbeelden, en vandaaruit toe te werken naar de meest geformaliseerde definitie die u wenst te gebruiken. Vragen in een eindtoets zijn meestal op die uiteindelijke, meest formele definitie gericht (dat meest formele niveau kan natuurlijk nog informeel zijn, bijvoorbeeld ostensief).

De aanvulvraag is heel passend voor definitievragen:

VOORBEELD

Wat is een pingo?

Antwoord: een pingo of vorstheuvel is een ijslens in de bodem, ontstaan onder omstandigheden van eeuwige vorst, waarbij de bovengrond opgeperst kan worden tot een enige meters hoge heuvel.

VOORBEELD

Hoe groot is de afstand tussen atomen in een molecuul?

Antwoord: de afstand tussen atomen is gelijk aan de som van de covalente stralen.

VOORBEELD

Geef de definitie van 'modus'.

Antwoord: de modus van een aantal getallen of metingen is de waarde die het meest voorkomt.

Het gebruik van aanvulvragen veronderstelt dat de student geleerd heeft definities te geven, en dat in het onderwijs behandeld is wat als adequate definitie beschouwd wordt. Aanvulvragen kunnen tot beoordelingsproblemen aanleiding geven wanneer de gevraagde definitie niet erg strikt is maar meer de kant van een beschrijving opgaat (zoals in het pingo-voorbeeld). Het kan aanbeveling verdienen dan als regel te hanteren dat minder essentiële toevoegingen niet tot een verhoogde score zullen leiden; bijvoorbeeld: de definitie van een pingo als 'een ijslens in de bodem' is voldoende.

Een enigszins gekunstelde vraagvorm is het omgekeerde: *gegeven een definitie, gevraagd de term die aldus gedefinieerd is*. In veel omstandigheden is dat toch prima bruikbaar. In sommige vakgebieden is dit definitiegebruik juist heel belang-

rijk, omdat op deze wijze teksten of probleemstellingen vereenvoudigd moeten worden.

VOORBEELD

Een door warmer klimaat gesmolten ijslens die een met water gevulde depressie in het landschap achterlaat noemen we

Antwoord: een pingoruïne.

Het risico van dit soort vragen is dat het abstractieniveau te hoog ligt (zie 2.6). Daaraan zou men in dit voorbeeld iets kunnen verbeteren door ook een behandeld voorbeeld van een pingoruïne te vragen. bijvoorbeeld het Esmeer bij Norg (Drente). Een pingoruïne is niet direct als pingoruïne waarneembaar, zodat het vrijwel niet mogelijk is om nieuwe voorbeelden te vragen (sommige vennen en dobben zijn pingoruïnes, door onderzoek te bepalen).

Meerkeuzevragen zijn niet aan te raden. Het is zelden een onderwijsdoel dat inadequate, onjuiste, of zelfs geheel willekeurige definities als onjuist afgestreept kunnen worden.

AFRADER

Een synaps is het best te beschrijven als

1. een klont of laag protoplasma met veel nucleï maar zonder onderscheidbare celgrenzen
2. geheugenverlies veroorzaakt door onvoldoende bloedtoevoer naar de hersenen
3. het paren van chromosomen van mannelijke en vrouwelijke zijde bij het rijpen van de kiemcellen
4. het lange cilindrische deel van een axon
5. het punt waar de zenuwimpuls van het ene op het andere neuron overgaat

(Bloom c.s., 1956, blz. 79.)

Dat de student moet weten dat alternatief 5 een beschrijving (nominale definitie) van een synaps is, zou een onderwijsdoel kunnen zijn; maar stel dan liever een aanvulvraag, laat een synaps schetsen, of laat bij de schets van een zenuwcel de onderscheiden onderdelen benoemen. Keuzevragen zoals deze afrader zijn op zijn best denkbaar als onbenullige toetsvraag (6.1).

Er is wel een variant op de meerkeuzevraag die hier soms bruikbaar is: de matchingvraag waarbij gevraagd wordt de gegeven termen op de juiste wijze te koppelen aan de eveneens gegeven definities, bijvoorbeeld voor vier of vijf paren tegelijk. Fraai is het niet, maar het is beter dan de gewone meerkeuzevraag.

Een mogelijkheid die eveneens afgeraden wordt, is het laten verbeteren van een foutieve, gegeven definitie bij een gegeven term. Ook dat past doorgaans niet bij de onderwijsdoelen, tenzij het een opleiding tot corrector van definities betreft.

Er zijn doorgaans wel mogelijkheden om een *gedeelte van een definitie* te vragen, zoals deze invulvraag.

VOORBEELD

Een pingo is een in de bodem.

Niet alleen zijn vragen naar een gedeelte van de definitie gemakkelijker, maar ze zullen doorgaans ook sneller te beantwoorden zijn en eenduidiger beantwoording toelaten (een gemak bij het nakijken). Daar staat dan weer de geringere informatieve waarde tegenover.

Vanzelfsprekend liggen de beste vraagmogelijkheden bij de *toepassingen*. Maar daarmee kom ik op het terrein dat in volgende paragrafen en hoofdstukken bestreken zal worden.

VOORBEELD

Bereken de C-Cl-afstand in het CCl₄-molecuul.

Antwoord: de afstand tussen atomen is gelijk aan de som van de covalente stralen (in een tabel gegeven). Zodat: afstand C-Cl is $R_C + R_{Cl} = 0,77 + 0,99 = 1,76 \text{ \AA}$.

(Pierce & Smith, *General chemistry workbook*, Freeman, 1971, blz. 94.)

Definitievragen kunnen ook gericht worden op het speciale karakter van de definitie die aan een gegeven term is toegekend, en dat is zeker een geldige wijze van vragen wanneer daaraan in het onderwijs of het tekstboek aandacht is besteed. Ik verwijz hiervoor terug naar 3.5.

Omgekeerd geeft het karakter van een bepaalde definitie ook richting aan wat erover gevraagd kan worden. Termen die alleen ostensief gedefinieerd zijn, daarover is moeilijk iets an-

ders te vragen dan een gegeven voorbeeld te benoemen of zelf een nieuw voorbeeld te produceren. Om met de toetsvraag in te haken op het speciale karakter van een definitie, is het gemakkelijk fig. 3.3 (te doorgronden en) bij de hand te hebben. Vergeet niet dat vragen rond een bepaalde definitie ook over andere zaken kunnen gaan dan de inhoud ervan, bijvoorbeeld over de historische ontwikkeling. 'De wisselvalligheid van wetenschappelijke definities: wat vandaag proefondervindelijk geldt als een begeleidend verschijnsel bij fenomeen A, wordt morgen gebruikt voor de definitie van "A".' (Wittgenstein, 1953, par. 79).

Definitievragen zijn vrijwel altijd ongepast wanneer ze uit nood geboren zijn, bijvoorbeeld om maar aan zoveel mogelijk toetsvragen te komen of omdat de ideeënstroom voor behoorlijke toetsvragen is opgedroogd. Ik heb in deze paragraaf geen aanwijzingen willen geven die deze verleiding nog groter maken. Gebruik ze met mate, en alleen wanneer een enkele mogelijkheid herkend wordt als goed aansluitend op de onderwijsdoelen.

4.3 Voorbeelden geven

Wat versta ik nu onder een voorbeeld? Alle verschijnselen, objecten en dergelijke die we met een en dezelfde term aanduiden, zijn te beschouwen als voorbeelden van (de betekenis van) die term, of als voorbeelden van het begrip dat met die term wordt aangeduid. In het volgende zal ik, zoals ook in hoofdstuk 3 gedaan is, op elliptische wijze over 'voorbeelden van een term' spreken, zolang daardoor geen misverstanden kunnen ontstaan. Voorbeelden zijn ruwweg op twee manieren te vragen: er kan gevraagd worden voorbeelden te *herkennen of benoemen*, of er kan gevraagd worden een voorbeeld te geven. Herkennen en benoemen wordt in 4.4 besproken. Het vragen van voorbeelden kan er ongecompliceerd aan toe gaan.

VOORBEELD

Geef een nieuw voorbeeld van een ecosysteem.

Mogelijke antwoorden: wei; vijver; berm; tuinkas; pingoruïne.

Een goed antwoord zou ook kunnen bestaan uit het *aanwijzen* van een ecosysteem in de omgeving (Zoek een voorbeeld van een ecosysteem). Zeggen dat een vijver een voorbeeld van een ecosysteem is, en het aanwijzen van een vijver als een voorbeeld van een ecosysteem is bijna hetzelfde gedrag. Je ogen over de omgeving laten gaan om te zien of er ergens een voorbeeld van een ecosysteem te vinden is, verschilt niet zoveel van in gedachten of verbeelding hetzelfde doen. In beide gevallen gaat het erom dat de student ecosystemen heeft leren *zien*, zowel in overdrachtelijke zin als in directe *waarneming*. Beheersing van het begrip ecosysteem houdt in dat de zintuiglijke waarneming op een bepaalde wijze georganiseerd is: een vijver wordt niet zomaar meer als een pittoreske plas water gezien, maar als een ecosysteem, en die vijver heeft daardoor een rijker betekenis gekregen.

'Er is een manier van kijken naar elektrische machines en installaties (dynamo's, radio zendstations, en dergelijke) waarbij deze objecten gezien worden als een ensemble van koper, ijzer, rubber etcetera in de ruimte, zonder enig begeleidend begrijpen'. (Wittgenstein, *Zettel*, nr. 711). De ordening in zo'n samenstel van materialen zien, en in die ordening ook de functies herkennen, is precies wat ik meer prozaïsch aanduid als het waarnemen van voorbeelden. Dat waarnemen kan vaak als vanzelfsprekend beschouwd worden, het kan in veel gevallen bijna ongemerkt geleerd worden, juist omdat het zo'n menselijke, alledaagse wijze van leren is. Daarnaast zijn er in ieder vakgebied wel onderwerpen waar het bewuste aandacht en training vereist, of een ruime ervaring die alleen maar in de praktijk kan worden opgedaan, om ze adequaat te leren waarnemen, of om bepaalde verbanden te leren zien. De wetenschap probeert dan wel ordening in de verschijnselen aan te brengen, maar natuurlijke verschijnselen laten zich niet altijd even makkelijk ordenen en van elkaar onderscheiden (zie bijv. Smith en Medin, 1981).

'In het dagelijks leven zijn slecht-omschreven begrippen de regel, niet de uitzondering. Het visuele onderscheid tussen honden en katten, of tussen schoonheid en lelijkheid, is slecht omschreven evenals de begripsmatige verschillen tussen creatieve wetenschap en broddelwerk, of tussen gezondheid en neuroticisme. Hetzelfde geldt voor EEG-patronen die samengaan met een bepaald stadium in de slaap, röntgenschaduwen die

wijzen op een tumor, de schilderstijl die een Picasso karakteriseert, of de gelaatstreken die over de jaren heen het gezicht van een vriend kenmerken' (Neisser, 1967 blz 58).

Het waarnemen wordt in het onderwijs nogal eens stiefmoederlijk behandeld. In de cognitieve taxonomie van doelstellingen van Bloom c.s.(1956) is er geen plaats voor ingeruimd; dat levert in ieder geval bij de kunstzinnige vakken onmiddellijk problemen op, en we zien dan ook in het hoofdstuk 'Art Education' in het handboek van Bloom, Hastings en Madaus (1971) een toevoeging van het waarnemen aan de taxonomie, die in dat verband voor zichzelf spreekt (voor kennis en begrip). Wilson, de auteur van dat hoofdstuk, merkt over het waarnemen op 'In de kunstzinnige vorming heeft het waarnemen betrekking op het vermogen om dingen en gebeurtenissen te zien op een wijze die verder gaat dan de alledaagse waarneming en het louter herkennen' (o.c., blz. 517). 'Kunstzinnige vorming onderscheidt zich hierin echter niet of nauwelijks van andere vakken, of van een denksport als go of schaken, waar de amateur en de meester dezelfde stelling op heel andere wijze "zien" (De Groot, 1978).

Ik wil hiermee benadrukken dat het leren waarnemen van de wereld op een wijze die adequaat is voor de beoefening van de onderwezen discipline geen geringe prestatie hoeft te zijn, en bij de toetsing niet vergeten mag worden, al was het alleen maar omwille van een goede evaluatie van wat met het gegeven onderwijs is bereikt. Aan deze wens zal niet altijd voldaan kunnen worden, omdat er soms speciale technieken nodig zijn om het waarnemen goed te kunnen toetsen. Met inschakeling van audiovisuele media, waaronder bijvoorbeeld reproductietechnieken voor afbeeldingen, is echter veel te bereiken. Voor enkele voorbeelden in 'art education' zie Wilson, in Bloom c.s. (1971).

Tot zover over het 'zien' en aanwijzen van nieuwe voorbeelden, wat niet hetzelfde is als, maar dicht staat bij, het herkennen en benoemen van voorbeelden die de student worden aangereikt.

De meerkeuzevraag is niet geschikt om voorbeelden te vragen, omdat deze vraagvorm mikt op het herkennen van een in de alternatieven voorkomend voorbeeld. Dat is niet slecht,

maar is het onderwerp van 4.4. In de regel zijn alleen aanvulvragen geschikt om naar nieuwe voorbeelden van een bepaalde term te vragen. Maar vat dat niet te strikt op, want er zijn goede tussenvormen tussen de meerkeuze- en de aanvullende vraagvorm die wèl bruikbaar zijn.

VOORBEELD

Wijs op de afgedrukte kaart aan waar zich een bergpas bevindt.
(Of: vliegveld, kerncentrale, uraniummijn, waterval enz.)

Wanneer gevraagd wordt naar een *oud* voorbeeld, een voorbeeld dat in de leerstof al gegeven is, is de vraag een eenvoudige reproductievraag. Hoe zit dat nu wanneer een *nieuw* voorbeeld wordt gevraagd? In theorie is het mogelijk dat sommige studenten nieuwe voorbeelden hebben bedacht bij de voorbereiding op de toets en deze eenvoudig uitschrijven, reproduceren tijdens de toetsing. In de praktijk zal het meestal zo zijn dat het aantal mogelijk te vragen voorbeelden te groot is om ze alle op deze wijze te kunnen prepareren. Dat geldt zeker wanneer gevraagd wordt naar voorbeelden die moeten beantwoorden aan opgegeven specificaties van plaats, tijd, omvang en dergelijke, voorbeelden die veeleer geconstrueerd dan bedacht moeten worden.

Een interessante mogelijkheid is (een deel van) de toetsvragen op te geven als huiswerk: bedenk thuis een opgegeven aantal nieuwe voorbeelden bij opgegeven termen, en lever die voor de toets in. Een gedeeltelijk take-home-tentamen, met een enigszins practicumachtige inslag. Het zijn opgaven die de student in hoge mate kunnen verzekeren van een positieve beoordeling. Het samenwerken van studenten is niet af te raden, het kan misschien zelfs semi-geïstitutionaliseerd gebeuren. Deze mogelijkheid is alleen reëel wanneer het ongebruikelijk is dat studenten het werk van voorgangers overschrijven of wanneer gevraagde nieuwe voorbeelden zo nauw omschreven kunnen worden dat voor iedere toetsgelegenheid een nieuwe set opgaven verstrekt wordt.

Sommige termen zijn namen, ofwel begrippen waar maar één 'voorbeeld' van bestaat: de maan, de Mona Lisa, het getal pi. Het zou ongepast zijn om een voorbeeld van de Mona Lisa te

vragen. Wat wel mogelijk is: vragen om dat ene bestaande voorbeeld aan te wijzen (de Mona Lisa in het Louvre, de maan aan de hemel, maar ook: een afbeelding van de Mona Lisa in een boek over het werk van Da Vinci, de maan in een plaat met afbeeldingen van hemellichamen in ons zonnestelsel).

De grondvorm voor het vragen van voorbeelden is deze:

VOORBEELD

Geef een voorbeeld van een (term).

(Mogelijke termen: een metafoor, vogelsoort, ontdekkingsreiziger, priemgetal, zeeslag uit de Tachtigjarige Oorlog enz.)

Uit deze grondvorm zijn een groot aantal varianten af te leiden door variëren van de gevraagde presentatie ('geef'), het aantal gevraagde voorbeelden, een oud of een nieuw voorbeeld, het soort voorbeeld, abstractie van het voorbeeld, opgegeven randvoorwaarden of specificaties, terwijl het ook nog mogelijk is de vraag te beperken tot aspecten (eigenschappen bijv.) van het bedoelde begrip in plaats van 'complete' voorbeelden. De aard van de gevraagde presentatie wordt veelal gedicteerd door de aard van het gevraagde voorbeeld (een tekening, naam, constructie enz.).

Enkele mogelijkheden zijn:

- geef - wijs aan
- maak - demonstreer
- teken - speel
- construeer - noem

Gebruik liever geen opdrachten als 'bedenk', omdat de student daar zelf bij moet bedenken dat zijn bedenken niet voldoende is, en kennelijk ook verwacht wordt dat wat bedacht is opgeschreven wordt. Het lijkt overdreven, maar dit soort zorgvuldigheid in toetsopdrachten betaalt zich ongetwijfeld terug.

De variaties in het *aantal* gevraagde voorbeelden behoeven geen opsomming. Wanneer niet een exact aantal voorbeelden gevraagd wordt, wordt de verwachting gewekt dat meer voorbeelden meer punten voor de vraag opleveren. Dat kan soms een te rechtvaardigen bedoeling zijn, maar in

het algemeen is het aan te raden precies aan te geven hoeveel voorbeelden worden gevraagd. Anders is het onduidelijk wat er getoetst wordt: vakkennis, het creatieve vermogen een groot aantal voorbeelden te construeren of het lief om niet meer dan een enkel voorbeeld te geven.

Het *soort* te vragen voorbeelden laat een grote variatie toe. In 2.5 werd een uitgebreide lijst gegeven; samengevat zijn het deze mogelijkheden voor nieuwe, eventueel ook voor oude, voorbeelden:

- typisch, veel voorkomend, meest voorkomend
- gewoon, normaal
- buitengewoon, bijzonder
- grensgeval (waarbij bestreden kan worden of het wel een voorbeeld is)
- uitzondering

Wanneer een bepaald soort voorbeeld gevraagd wordt, moet dat in de vraagstelling wel duidelijk uitkomen, en soortgelijke voorbeelden moeten in het onderwijs behandeld zijn.

Het gevraagde voorbeeld kan nader bepaald zijn doordat het binnen opgegeven *randvoorwaarden* moet vallen, doordat het aan bepaalde *specificaties* moet voldoen enz.:

- een bepaalde eigenschap moet aanwezig zijn (niet vliegende vogelsoort)
 - een bepaalde eigenschap is kwantitatief gespecificeerd (driehoek met een hoek van 15 graden)
 - een randvoorwaarde is opgegeven (ontwerp voor een stoel waarbij de materiaalkosten binnen een bepaalde grens moeten blijven)
 - een bepaald tijdsbestek (een hedendaags auteur); nationaliteit (een Nederlands auteur); een bepaalde plaats of streek (een in Parijs wonend auteur); combinaties daarvan (een hedendaags, in Parijs wonend, Nederlands auteur)
 - bepaalde materialen moeten gebruikt worden (een amalgaamvulling)
- enzovoort.

Er kan variatie gebracht worden in het *niveau van abstractie* waarop het voorbeeld geleverd moet worden (zie ook 2.6):

- concreet, lijfelijk

- symbolische of schematische representatie
- verbale beschrijving
- naam (term)
- afbeelding, schets.

‘Voorbeelden’ zijn niet altijd voorbeelden: bij *abstracte termen* (landsbestuur, intelligent gedrag, sociale klasse) spreken we niet van voorbeelden, maar van uitingen, aspecten, onderdelen en dergelijke. Je kunt wel zeggen dat het gebruik van voorwerpen als instrument een voorbeeld van intelligent gedrag is, maar van het wekelijks kabinetsberaad zeg je dat het een onderdeel van het landsbestuur is. Besteed bij abstracte termen enige zorg aan de omschrijving van de uiting, het onderdeel enz. die gevraagd worden.

Bepaalde termen laten zich niet in voorbeelden vertalen of definiëren: *theoretische termen* lenen zich dan ook niet tot deze voorbeeldvragen. Wie toch iets voorbeeldachtigs rond theoretische termen wil vragen, komt terecht bij het vragen van voorbeelden die betrekking hebben op een bepaalde *relatie* van de theoretische term tot een andere (theoretische) term. Zie daarvoor hoofdstuk 5.

Een heel speciaal voorbeeld is nog niet genoemd: het *niet-voorbeeld*. Bij termen die een tamelijk scherpe definitie kennen in termen van voldoende en noodzakelijke eigenschappen, kan het zinvol zijn om expliciet aandacht te schenken aan gevallen die *niet* onder de term vallen omdat er één kritische eigenschap ontbreekt. Dergelijke gevallen worden wel niet-voorbeelden genoemd. Wanneer dat zinvol is, kan dan ook in de toets gevraagd worden bij een bepaalde term een niet-voorbeeld te geven dat in alle kritische eigenschappen op één na voldoet aan de definitie van de gegeven term. Omdat er noch in de wetenschap, noch in het onderwijs veel termen zijn die zich strikt laten definiëren in voldoende en noodzakelijke eigenschappen, en omdat het dan kunnen bedenken van niet-voorbeelden maar zelden relevant zal zijn, is deze variant maar beperkt bruikbaar. De bruikbaarheid van de omgekeerde variant is iets groter, omdat de docent dan zelf de constructie van het niet-voorbeeld in de hand heeft (zie 4.4). Enige verwantschap met het laatste onderwerp heeft de vraag naar voorbeelden van *grensgevallen*. In plaats van voorbeelden

te vragen, zou ook overwogen kunnen worden om een opstelachtige uiteenzetting te vragen over de problematiek van de vage afbakening van een bepaald begrip (medische diagnostiek, biologische taxonomie, psychopathologie).

Een laatste vraagvariant is deze: gegeven een aantal voorbeelden van een niet met name aangeduide term, wordt een nieuw voorbeeld van de zo aangeduide term gevraagd. Dit is een combinatie van herkennen en produceren van voorbeelden, ik zal daar in de volgende paragraaf op terugkomen.

Wanneer de student voorbeelden in hoge mate beschikbaar heeft, ze gemakkelijk en snel kan produceren, vergroot dat het vermogen om actief zowel als passief met de wetenschappelijke terminologie om te gaan. Enige nadruk op deze voorbeeldvragen in onderwijs en toetsing is daarom de moeite waard. Daar is aan toe te voegen dat voorbeelden meer waarde hebben naarmate ze concreter zijn, levendiger voorgesteld kunnen worden, of appelleren aan eigen ervaringen (bijv. in het laboratorium of de kliniek).

4.4 Voorbeelden herkennen en benoemen

In 4.3 was het onderwerp: geef een voorbeeld bij een gegeven term. Omgedraaid levert dat het nu te bespreken onderwerp: gegeven een voorbeeld, herken of benoem het. Ik gebruik benoemen voor het zelf produceren van de bijbehorende term, en herkennen voor het koppelen van het gegeven voorbeeld aan de eveneens bijpassende term. Dat correspondeert met respectievelijk aanvulvragen en keuzevragen.

Het belang van het kunnen benoemen of herkennen van voorbeelden heeft nauwelijks toelichting: zonder bij voortduring onze zintuigindrukken te ordenen en te rubriceren zou normaal leven onmogelijk zijn. Wat voor het dagelijks leven geldt, strekt zich ook uit tot het voortdurend rubriceren van al die zintuigindrukken die beroepsmatig relevant zijn. In de praktijk dienen voorbeelden zich ongevraagd aan en moeten 'spontaan' waargenomen en herkend worden. In de praktijk kom je geen 'voorbeelden' tegen met een label eraan 'ik ben een te rubriceren voorbeeld'. Er is een zekere spanning tussen wat er in de praktijk gevraagd wordt aan waarnemen en

thuisbrengen van voorbeelden, en de wijze waarop in een toets situatie voorbeelden min of meer kant en klaar ter benoeming aangeboden worden. Vragen waarin gegeven voorbeelden gerubriceerd moeten worden, zijn niet helemaal 'geldig' of representatief voor het rubriceren in de beroepspraktijk.

Een gebrek waar voorbeeldvragen erg gevoelig voor zijn, is dat de gegeven voorbeelden vaak op een te hoog abstractieniveau staan. Vragen waarbij de voorbeelden bij naam gepresenteerd worden zijn in dit opzicht al gauw verdacht.

VOORBEELD

Wei, vijver, aquarium en berm zijn voorbeelden van

Antwoord: ecosystemen.

In het gegeven voorbeeld kan de student zich bij de gegeven termen waarschijnlijk wel een goede voorstelling maken, en kan deze abstracte presentatie in dit geval aanvaardbaar zijn. Een besliste afrader is het gebruiken van een voorbeeld dat in de meest algemene termen beschreven is, eigenlijk een beschrijving of definitie van de bedoelde term is (zie 2.6 voor voorbeelden).

Een speciaal aandachtspunt bij deze voorbeeldvragen is of het benoemen van een gegeven voorbeeld wel redelijk eenduidig kan gebeuren.

Een gegeven voorbeeld kan ook een voorbeeld van een helemaal niet door de docent bedoelde term zijn. In kenmerkende en tevens extreme vorm doet dit probleem zich in de wiskunde voor:

VOORBEELD

Geef de vijfde en de n-de term bij de rij 1, 16, 81, 256.

Antwoord: 625 en n^4 .

Maar ook 601 en $10n^3 - 35n^2 + 50n - 24$.

'Alleen maar het, geven van een eindig aantal termen van een rij is niet voldoende om een unieke n-de term te definiëren. In feite zijn een oneindig aantal verschillende getallen voor die n-de term mogelijk' (opgave en citaat uit Spiegel, *Advanced Calculus*, Schaum Series, 1963, blz 44).

Wat voor bepaalde wiskundige opgaven geldt, geldt ook voor

voorbeeldvragen in andere vakgebieden. Het is niet altijd mogelijk door het geven van meer termen onbedoelde juiste antwoorden uit te sluiten:

AFRADER

Mulisch, Hillenius, Wolkers en Vinkenoog zijn

Antwoord: schrijvers. Maar ook: mannen, Nederlanders, Amsterdammers enz.

Wie alle onbedoelde juiste antwoorden wil uitsluiten, staat voor een soms immense opgave. In de schrijversvraag moet bijvoorbeeld geslacht, nationaliteit en iedere andere 'irrelevante eigenschap' stelselmatig gevarieerd worden. Maar zo zwartgallig hoeft het probleem nu ook weer niet bekeken te worden: niet alleen kan een goede variatie in de enkele aangeboden voorbeelden voldoende duidelijk maken welke term gevraagd wordt, maar ook zal vanuit de impliciete context van het gegeven onderwijs duidelijk te maken zijn dat bepaalde onbedoelde antwoorden, hoewel taalkundig juist, toch redelijkerwijs niet goed gerekend kunnen worden. Dat Harry Mulisch een Nederlander is, is op zich niet onjuist, maar is evident niet het antwoord dat met deze vraag bedoeld is.

AFRADER

Een vijver is een (ecosysteem)

VOORBEELD

Vijver, wei, aquarium, berm, zijn (ecosystemen)

Wees voorzichtig met het fout rekenen van evident niet bedoelde, maar op zich correcte antwoorden; wanneer tevoren geen duidelijke afspraken gemaakt zijn, is het beter andere mogelijkheden te onderzoeken om de vraag een betere 'definitie' te geven. Een mogelijkheid daartoe is het in de vraag formuleren welk soort antwoord er verlangd wordt.

VOORBEELD

Van welk soort toetsvraag is het voorbeeld hierboven een voorbeeld? (aanvulvraag)

Evenwel, door een extra aanwijzing in de vraag op te nemen,

verschuift de vraag al enigszins van het benoemen naar het herkennen.

VOORBEELD

De hierboven gegeven toetsvraag is een

1. aanvulvraag
2. keuzevraag
3. opstelvraag

In deze en de volgende paragraaf is het herkennen en benoemen van voorbeelden aan de orde, voor zover het daarbij niet gaat om het onderscheiden tussen verschillende termen (zie daarvoor 5.2). 4.5 is speciaal gewijd aan termen die strikt definieerbaar zijn '*per genus et differentiam*', in voldoende en noodzakelijke eigenschappen.

Er zijn twee grondvormen als basis voor varianten: 'benoem dit voorbeeld' en 'herken dit voorbeeld'. De varianten op de eerste grondvorm zijn te verkrijgen op dezelfde wijze als in de voorgaande paragraaf, maar let op: het gaat niet om louter een herhalingsoefening.

De enige varianten op 'benoemen' zijn synoniemen daarvan ('hoe heet'), en de soms noodzakelijke, soms als hint werkende, nadere aanduiding van wat er gevraagd wordt.

VOORBEELD

Welke penseeltechniek is hier gebruikt?

(Zonder het noemen van de penseeltechniek als het gevraagde, gaat het niet.)

VOORBEELD

Welke plantesoort is dit?

(Hier zou volstaan kunnen worden met het geven van een afbeelding, en de opdracht de plant te benoemen.)

VOORBEELD

Van welke Franse schilder is dit doek?

(Hint: de nationaliteit van de schilder.)

In plaats van de toepasselijke term of naam te vragen, kan een nieuw voorbeeld, of ten minste een ander voorbeeld, ge-

vraagd worden. Dan ontstaat een vraagvorm die een *combinatie* is van het benoemen van een gegeven voorbeeld met het geven van een voorbeeld bij een gegeven term. Een andere combinatiemogelijkheid is het *beschrijven*: gegeven een of meer voorbeelden bij een (niet genoemde) term, vragen naar een wetenswaardigheid met betrekking tot die term (zie ook 4.6). Toetsvragen waarbij het benoemen wordt overgeslagen, zijn in veel vakken ongebruikelijk, vooral wanneer het om toepassingsvragen gaat. In de wiskunde bijvoorbeeld gaat het er niet om dat het type opgave correct benoemd kan worden, maar of er de juiste of beste oplostechniek op toegepast wordt. De variaties in het *aantal gegeven voorbeelden* zijn louter getalsmatig: een, twee, of meer. De reden om meer dan een voorbeeld te geven kan zijn om zo het verlangde soort antwoord scherper te specificeren; zorg er dan voor dat de gegeven voorbeelden heterogeen zijn op bijkomstige eigenschappen (schrijvers van verschillend geslacht, leeftijd, nationaliteit, wanneer het erom gaat dat de student de gegeven namen van het label 'schrijver' moet voorzien). Een andere reden om meer voorbeelden te geven: de vraag gemakkelijker maken.

Het maakt meestal enig verschil of een gegeven voorbeeld een *oud*, een bekend voorbeeld is, of een *nieuw*, nog niet eerder ontmoet voorbeeld. Het kunnen benoemen van oude voorbeelden zal doorgaans niet tot de onderwijsdoelen behoren; deze wijze van stof terugvragen is misschien geschikt voor de onbenullige toets, waarbij het alleen maar de bedoeling is vast te kunnen stellen of de student de literatuur onder ogen heeft gehad (zie 6.1).

Een rijke bron van variaties ligt in het *soort* voorbeeld dat gegeven wordt: typisch, gewoon, buitengewoon, grensgeval, uitzondering (uitgebreider in 2.5). Met grensgevallen moet heel behoedzaam worden omgesprongen. Het is eigenlijk alleen maar zinvol om met grensgevallen te werken wanneer er bijzondere waarde wordt gehecht aan het herkennen van grensgevallen 'als grensgevallen', omdat het gevolgen kan hebben in de toepassings sfeer. Bijvoorbeeld: diarree bij pasgeborenen is door de huisarts moeilijk als onschuldig of niet onschuldig te herkennen, wat kan leiden tot de stelregel altijd tot ziekenhuisopname over te gaan.

Een andere mogelijkheid tot variëren is gelegen in de keuze van *tijdsbestek*, *gebied* en dergelijke waaruit het voorbeeld

afkomstig is, of door het afwisselen van tamelijk ondergeschikte kenmerken (niet alleen schilderijen, maar ook tekeningen van Picasso; en plant in bloei, naast dezelfde plant met vrucht; gave klokbekers of scherven).

Vanzelfsprekend kan het abstractieniveau van het voorbeeld gevarieerd worden (met voorkeur voor een zo concreet mogelijk niveau).

Dan is er ook nog de mogelijkheid om een *niet-voorbeeld* te geven, maar dat vraagt een kleine aanpassing in de vraagstelling. Immers, wanneer gevraagd wordt te benoemen, impliceert dat dat het gegeven voorbeeld inderdaad een voorbeeld is, en geen niet-voorbeeld. Wil men nu ook niet-voorbeelden gaan gebruiken, dan moet de vraagstelling in tweeën gebeuren: eerst vragen of het gegeven een 'voorbeeld' is, en wanneer dat niet geval is, vragen het te benoemen.

VOORBEELD

Is bij het gegeven preparaat sprake van gezond weefsel, en zo nee, van welk soort afwijking is er dan sprake?

Ik vermoed dat het werken met niet-voorbeelden beperkt is tot gevallen waarin het niet-voorbeeld behoort tot dat wat 'normaal' is, en voorbeelden behoren tot de afwijkingen van het normale: een pathologie, overtreding, of defect.

Een heel belangrijke voorbeeldvariant is het versluisde voorbeeld, het voorbeeld waar eerst enige bewerking op gepleegd moet worden om het in een vorm te krijgen die zich laat benoemen. Ik heb er hierboven al op gewezen dat in opgaven van wiskundige aard het vaak voorkomt dat gegevens omgewerkt moeten worden tot bekende vormen, al is dat niet in de eerste plaats om die te kunnen benoemen, maar om er vervolgens bekende oplostechnieken op te kunnen gebruiken. Dit is als het ware het omgekeerde van het geven van een extra aanwijzing: er wordt nu geen kant en klaar, direct te benoemen voorbeeld gegeven, maar een 'vuil' voorbeeld dat eerst gewassen en voorgebakken moet worden voordat het te benoemen of te herkennen is. Ik zal dat voorbeelden noemen die opgeschoond moeten worden. Dat *opschonen* kan op verschillende manieren gebeuren:

- anders rangschikken van de gegevens;

- in kaart brengen van de gegevens, een tekening maken;
- analyse: welke kenmerken zijn belangrijk, en welke doen er niet toe;
- redeneren: stapsgewijze het veld inperken waarbinnen het antwoord te vinden moet zijn (determineren);
- doorbreken van een verkeerde 'instelling', door het gegeven voorbeeld expliciet vanuit andere invalshoeken te gaan analyseren;
- hypothetisch-deductief te werk gaan: potentiële antwoorden opperen, en aan de hand van de verstrekte gegevens de juistheid ervan onderzoeken;
- zoeken naar impliciete context (bijvoorbeeld: op welk onderdeel van de bestudeerde stof heeft dit waarschijnlijk betrekking);
- gissen naar de bedoelingen van de docent (een onbedoelde vorm van zoeken naar impliciete context).

De variant waarbij potentiële mogelijkheden geopperd en op hun juistheid onderzocht worden, is van speciaal belang. Dat 'onderzoek' kan zich beperken tot het op een bepaalde wijze herschikken van het gegeven materiaal, het kan er eenvoudig uit bestaan dat nagegaan wordt of een bepaald kritisch kenmerk aanwezig is, maar het kan ook een laboratoriumonderzoek zijn. Vragen waarbij deze toetsende aanpak van de student gevraagd wordt, kunnen bekort worden door slechts te vragen welke potentiële mogelijkheden de student zou onderzoeken, of door te vragen welke test of tests de student uit zou kunnen voeren om tot een correcte benoeming of diagnose te komen. Uit de mogelijkheid dat er eerst een test uitgevoerd moet worden, voordat een gegeven voorbeeld benoemd kan worden, volgt de omgekeerde mogelijkheid dat een opgave vergemakkelijkt wordt door het testresultaat al bij de gegevens te verstrekken: dat is zowel een heel sterke hint als een bekorting van het zoekproces.

Een laatste variant is het vragen naar de *reden* voor de gegeven benoeming; op grond van welke kenmerken, waarom andere mogelijkheden verworpen zijn, en dergelijke. Of dat een zinvolle toevoeging is, hangt weer af van de aard van de stof en de onderwijsdoelen.

Dan moet nu de andere grondvorm nog besproken worden: *herken dit voorbeeld*. Herkennen is eigenlijk niet zo'n geschikte term om te gebruiken, want het verwijst naar een proces dat zich in het hoofd van de student, en dus niet direct waarneembaar, afspeelt. Ik bedoel met herkennen het (correct) koppelen van een gegeven voorbeeld aan een eveneens gegeven term. 'Herken een voorbeeld' is altijd een keuzevraag. De mogelijkheden zijn de meervoudige ja/nee-vraag met meerdere voorbeelden en niet-voorbeelden als alternatieven; de matchingvraag, bijvoorbeeld een rij gegeven namen invullen in een gegeven kaart of afbeelding; en de meerkeuzevraag met verschillende termen als alternatieven. Ieder van deze heeft eigen voor en nadelen, en daarmee verschillende toepassingsmogelijkheden.

De meervoudige ja/nee-vraag is heel geschikt om te toetsen of de student bij een gegeven term in staat is om voorbeelden en niet-voorbeelden te identificeren, normaal en afwijkend te onderscheiden en dergelijke. Ook wanneer voorbeelden van verschillende termen van elkaar onderscheiden moeten kunnen worden, is deze vraagvorm bruikbaar (zie 5.2). Waar het om gaat is of de student de gegeven term correct weet toe te passen. Er zijn twee mogelijkheden voor de student om dat ideaal te missen: hij kan de term te ruim gebruiken, ook voor niet-voorbeelden, dat is *overgeneralisatie*, of hij kan de term te voorzichtig gebruiken, enkele voorbeelden verslijten voor niet-voorbeelden, dat is *ondergeneralisatie*. De meervoudige ja/nee-vraag vormt een (grove) toets op zowel over als onder generalisatie.

Een simpele variant is de *enkelvoudige ja/nee-vraag*. Er wordt dan afgezien van het meer uitdrukkelijk toetsen van over- en ondergeneralisatie. Een reden daarvoor kan zijn dat het belangrijker is een wat groter aantal onderwerpen te toetsen, dan op enkele onderwerpen wat dieper in te gaan met meervoudige ja/nee-vragen. Ook de enkelvoudige ja/nee-vraag is alleen goed te gebruiken bij termen waarbij het erom gaat voorbeelden van niet-voorbeelden te kunnen onderscheiden.

Omdat het in veel (toegepaste) wetenschappen gaat om natuurlijke categorieën die geen scherpe afbakening toelaten, moet er speciaal op gelet worden dat voorbeelden en niet-voorbeelden zo gekozen worden dat er (tussen deskundigen)

geen verschil van mening bestaat over hun karakter als voorbeeld, respectievelijk niet-voorbeeld (en dat moet gecontroleerd worden, zie hoofdstuk 8).

AFRADER

Is dit een meubelstuk?

1. hi-fitoren ja/nee
2. antieke Friese staande klok ja/nee
3. stoof ja/nee
4. asbakstandaard ja/nee

De *matchingvraag* is te gebruiken wanneer een aantal verwante paren van voorbeelden en termen correct aan elkaar gekoppeld moeten kunnen worden. Een typisch voorbeeld is het invullen van gegeven namen in een kaart (Engelse plaatsnamen), of het invullen van gegeven namen in een afbeelding (onderdelen van een biologische cel). Ook het in hoofdstuk 2 gegeven voorbeeld van auteurs en titels uit hun oeuvre. Minder fraai is het combineren van paren voorbeelden en termen die weinig of niets met elkaar te maken hebben; wanneer zo'n vraag toch nog moeilijk blijkt te zijn, is het niet duidelijk wat ermee getoetst wordt behalve misschien het feit of al dan niet kennis is genomen van de opgegeven literatuur (onbenullige toetsvraag, zie 6.1).

De matchingvraag is ook geschikt in die gevallen waarin het kunnen produceren van namen niet zo belangrijk is. Bijvoorbeeld: het kan belangrijker zijn dat de leerling een goed overzicht heeft van de kaart van Engeland, hetgeen te toetsen is met een matchingvraag, dan dat de leerling alle plaatsen kan benoemen. Een matchingvraag is gemakkelijker dan de equivalente benoemvraag: het is gemakkelijker Engelse plaatsen met hun namen te combineren, dan de namen te noemen. De matchingvraag kan dan ook gebruikt worden als een gemakkelijke variant op de aanvulvraag waarin gevraagd wordt te benoemen.

Voor het toetsen van over- en ondergeneralisatie zijn matchingvragen niet geschikt, maar weer wel voor het toetsen van het kunnen onderscheiden tussen verwante begrippen (zie 5.2).

De *meerkeuzevraag* is te gebruiken in gevallen waarbij ervan

uitgegaan mag worden dat voor het benoemen van een gegeven voorbeeld de student een keuze doet uit een klein aantal alternatieven die hem of haar goed bekend zijn.

Een eenvoudig voorbeeld:

VOORBEELD

De hoofdstad van Gelderland is:

1. Arnhem
2. Apeldoorn
3. Nijmegen

Evenals de matchingvraag is de meerkeuzevraag geschikt voor toetsen op het kunnen onderscheiden tussen verwante begrippen (zie 5.2), wanneer dat onderscheiden zo moeilijk is dat daarop geoefend is.

Het belang van het kunnen benoemen van nieuwe voorbeelden bij de belangrijke termen van het vak, en in mindere mate het herkennen ervan, hoeft nauwelijks nog onderstreept te worden. Het vormt de basis voor iedere verdere activiteit, van het met anderen kunnen spreken over het vak tot aan het oplossen van problemen en het wetenschappelijk bezig zijn. Toch wordt aan benoemen en herkennen van nieuwe voorbeelden nogal eens te weinig aandacht besteed, zowel in het onderwijs als in de toets, ten gunste van extra tijd en plaatsruimte voor minder concrete en meer theoretische zaken. Een kenmerk daarvan is de schaarste aan voorbeelden die we in de meeste tekstboeken tegenkomen: er wordt dikwijls volstaan met het geven van slechts een enkel voorbeeld bij een geïntroduceerde term, en soms is zelfs dat ene voorbeeld maar een mager scharminkel van het in de meest algemene termen gestelde soort. Uitzonderingen vinden we bij de exacte vakken waar de nadruk ligt op het *actief* met de stof kunnen omgaan, en bij de sociale wetenschappen voor wat het vak 'cognitieve psychologie' betreft, waarin auteurs meer en meer de principes uit het eigen vakgebied ook op hun eigen studieboek leren toepassen. John R. Anderson, de schrijver van *Cognitive psychology and its implications*, zegt in zijn voorwoord: 'De vormgeving van de tekst berust op enkele "informed guesses" over hoe het studenten gemakkelijker gemaakt kan

worden belangrijke informatie onthouden. Het is niet de pretentie van dit boek ieder resultaat uit de cognitieve psychologie te bespreken; ik heb geprobeerd enkele honderden van de meest belangrijke feiten uit het vakgebied te identificeren, en heb daar alle aandacht op gericht. Er wordt echter wel een uitgezocht aantal details gegeven om de tekst enige redundantie te geven (dat bevordert het onthouden), en om de hoofdzaken belangwekkend en geloofwaardig te maken. Daarnaast heb ik geprobeerd te laten zien hoe deze algemene punten verband houden met specifieke gevallen door het gebruiken van voorbeelden, in het bijzonder voorbeelden die dicht bij het dagelijks leven staan.'

4.5 Herkennen en benoemen bij formeel gedefinieerde termen

In de voorgaande paragraaf over het herkennen en benoemen van voorbeelden heb ik niet gesproken over die bijzondere manier van benoemen of herkennen die gebruik maakt van formele definities. Aan het goedmaken van dat verzuim is deze hele paragraaf gewijd. Redenen voor afzonderlijke behandeling zijn onder andere dat het formele karakter ertoe uitnodigt de toetsing daar ook op te richten en sommige auteurs (Merrill en Tennyson, 1978; Tiemann en Markle, 1978) er geheel aan voorbijgaan dat de meeste termen uit leerstof niet formeel gedefinieerd zijn en soms ten onrechte formeel gedefinieerd zijn.

Formele definities zijn er in soorten en ze worden voor verschillende doeleinden gebruikt:

- in het onderwijs, om *didactische* redenen: een nieuwe term wordt geïntroduceerd, en daar worden dan *definiërende eigenschappen* bij gegeven, als de kortste weg om de betekenis van de term aan te geven;
- soms bij *natuurlijke categorieën*, om toch tot een scherpe afbakening van deze vaag omliggende begrippen te komen; dergelijke definities worden ook wel *technische* definities genoemd, en hebben vaak betrekking op eigenschappen die indirect waarneembaar zijn;
- voor *constructs*; dat is niet verwonderlijk: de definitie moet het construct volledig karakteriseren, en kan zodoende wei-

nig anders dan formeel zijn;
- als *stipulatieve* definitie, de definitie bij fiat.

In 3.5 is er uitvoerig op gewezen dat er naast formele definities (definities door kenmerken, door opsomming, door stipuleren, en de nominale definitie) andere definitiemogelijkheden zijn, die niet minder bruikbaar hoeven te zijn. Ik wijs daar nog eens op, omdat het zo makkelijk is bij het behandelen van deze formele definities te 'vergeten' dat er ook andere manieren van definiëren zijn en dat niet alle termen zich formeel laten definiëren.

Formele definities hebben een logisch karakter: toepassen ervan komt neer op logische operaties. Vandaar dat bij het schrijven duidelijk moet zijn of het daar gaat om een formeel gedefinieerde term of niet: de aard van de vraagstelling past zich daar bij aan. Ik heb al even enkele soorten formele definities genoemd. Een korte opsomming:

1. De meest bekende is de *definitie via kenmerken* (*per genus et differentiam*), waarbij een term gedefinieerd wordt aan een of meer andere termen en aan een of meer eigenschappen.
2. De *nominale* definitie stelt de gedefinieerde term gelijk aan een bepaalde uitdrukking (omgekeerd eigenlijk, ten minste meestal). Denk aan hoe je 'grootvader' zou definiëren: dat gaat niet in termen van eigenschappen, maar van bepaalde relaties.
3. de *technische* definitie is een definitie via kenmerken, of een nominale definitie, waarbij meestal gebruik wordt gemaakt van indirect waarneembare eigenschappen (waar instrumenten bij nodig zijn of een onderzoek e.d.).
4. De *stipulatieve* definitie is een definitie die de betekenis van een term bij afspraak vastlegt, en dat is meestal een nogal strikte afspraak.

Wat is nu het belang van formele definities? Overal waar de leerstof *constructs* behelst, zal een formele definitie doorgaans niet ver te zoeken zijn. En dat is goed omdat in zekere zin de betekenis van een construct geheel door zijn definitie bepaald wordt. Zo vinden we in de wiskunde, statistiek, en logica een belangrijke rol weggelegd voor (formele) definities. En hetzelfde is het geval bij al die wetenschappen die een enigszins geformaliseerde theorievorming kennen. In alle andere

gevallen, bij alle andere onderwerpen die in het onderwijs besproken worden, kan de vraag gesteld worden of formele definities wel op hun plaats zijn, dan wel of ze wel zo gelukkig zijn. Technische definities bijvoorbeeld hebben een nuttige functie in de uitoefening van het vak, maar zijn vaak specialistisch van karakter, en daarom zal de student er maar zelden mee lastig gevallen hoeven worden. Didactisch gezien kan het soms handig zijn een (pseudo-) formele definitie voor een bepaalde term te gebruiken, maar misschien is het desondanks beter om het toch maar niet zo doen. Vermoedelijk heeft het te maken met een bepaalde opvatting wat 'echte' kennis is, die een heel eerbiedwaardige geschiedenis heeft, maar die in het onderwijs meer kwaad dan goed kan doen. Baker en Hacker (1980, blz. 668):

'In zijn meer Socratische dialogen geeft Plato een negatieve karakterisering van begrijpen. Van ondervraagden die niet in staat zijn om bevredigende algemene definities te geven van zulke termen zoals "deugd", "gerechtigheid", "vroomheid", wordt gezegd dat ze deze termen niet begrijpen of dat ze niet weten wat deugd, gerechtigheid en vroomheid zijn. De zending waartoe Socrates zichzelf geroepen had, was om de mensen te overtuigen van hun onkunde over deze zaken. Zijn methode bestond eruit het onvermogen om een woord te definiëren te presenteren als bewijs van het niet begrijpen van het woord. Wittgenstein noemt dit argument "Plato's methode", en hij bekritiseert het als bedrieglijk. (...). De "Socratische drogreden" wordt ontmaskerd door het feit dat er andere criteria voor het begrijpen van een woord zijn dan het geven van zijn definitie; hoe iemand het woord gebruikt toont zijn begrip ervan, onafhankelijk of hij daar een definitie van geeft. Daar komt bij dat het toeschrijven van begrip geenszins belemmerd wordt doordat op verzoek een verkozen *vorm* van uitleg niet gegeven kan worden. Niettemin blijft het idee dat onvermogen om als antwoord op een Socratische vraag een definitie te geven, en schandelijk bewijs van onwetendheid is, en dit is de kracht achter de aanhoudende vraag naar definities via kenmerken (Merkmal-definities) voor termen die voor enig wetenschappelijk doel gebruikt worden.'

Wittgenstein was niet de eerste om de "Socratische drogreden" te signaleren. Baker en Hacker wijzen op een uitspraak van

Whewell, uit 1845, dat het vals is om te veronderstellen dat de wetenschap erom vraagt dat natuurlijke categorieën door essentiële kenmerken gedefinieerd worden. In tegenstelling daarmee laat onderzoek van de ontwikkelde takken van de natuurlijke historie het succesvolle gebruik zien van klassen die bepaald zijn door gelijkenis met voorbeelden ('typen'). Deze dienen het doel van het generaliseren; zij kunnen dat zelfs wel eens beter doen dan termen die met definiërende kenmerken en al ingevoerd worden (Baker en Hacker, 1980, blz. 325).

In scherp contrast met het bovenstaande zijn de pogingen in onderwijskundige boek om een didactiek voor begripsmatig leren van de grond te krijgen, nogal expliciet gebaseerd op het ideaal van formeel gedefinieerde begrippen. Bijvoorbeeld Merrill & Tennyson's *Teaching concepts; an instructional design guide* (1977, blz. 3 en 25):

'Een begrip is een verzameling specifieke objecten, symbolen, of gebeurtenissen die samen gegroepeerd worden op grond van gemeenzame kenmerken, en die met een bepaalde term of symbool aangeduid kunnen worden.' 'Een begripsdefinitie is een uitspraak die ieder van de kritische eigenschappen noemt, en aangeeft hoe deze eigenschappen gecombineerd zijn.'

Het zal duidelijk zijn dat deze aanpak maar op een klein aantal van de in onderwijs en wetenschap gebruikte termen zinvol toepasbaar is. Zodra deze aanpak als voorbeeldige didactiek gepresenteerd wordt, functioneert hij als een te strak keurslijf waar te veel zaken alleen maar verwrongen ingeperst kunnen worden. Op zijn best is deze aanpak eenvoudig onuitvoerbaar voor alle termen in een bepaald curriculum (kritiek van Roid en Haladyna, 1980, blz. 306), op zijn slechtst is het opleggen van dit paradigma aan het onderwijs rampzalig (de boze kritiek van Freudenthal, 1978, blz. 98 e.v.). Maar dat neemt allemaal niet weg dat met de nodige voorzichtigheid formele definities voor bepaalde soorten leerstof zinvol te gebruiken zijn.

De mogelijkheden voor toetsvragen zijn qua *vorm* dezelfde als in de voorgaande paragraaf behandeld, zowel wat benoemen als wat herkennen betreft, en daar verwijs ik dan ook naar. Wat er toegevoegd wordt, is dat voor het *herkennen* het ge-

ven voorbeeld of niet-voorbeeld *getoetst* wordt aan de formele definitie: zijn alle definiërende kenmerken aanwezig? Bij het *benoemen* gebeurt hetzelfde, maar daar zal veelal eerst veronderstellenderwijs een term geopperd moeten worden, op zijn toepasbaarheid onderzocht, en eventueel ingeruild voor een nieuwe veronderstelling enz.

De wijze waarop de student het voorbeeld toetst aan de definitie is in overeenstemming met de relatie die meerdere definiërende eigenschappen tot elkaar hebben. De meest voorkomende relaties zijn eigenschappen die gezamenlijk aanwezig moeten zijn, en eigenschappen waarvan er tenminste een aanwezig moet zijn. Wat voor eigenschappen geldt, geldt natuurlijk ook voor relaties, wanneer de definitie niet in termen van eigenschappen, maar in termen van relaties is gesteld.

Is een formele definitie bruikbaar, dan zal er vaak sprake zijn van bepaalde eigenschappen die in de definitie voorkomen (*definiërende* eigenschappen), eigenschappen die de term onderscheiden van een verwante term (*kritische* eigenschappen) en eigenschappen die niet van belang zijn voor de definitie van de term, maar mogelijk bruikbaar zijn om individuele objecten enz. te beschrijven (*variabele* eigenschappen). Dit is de indeling die in onderwijskundige publikaties vaak te vinden is, en daar moet de waarschuwing bij dat hier impliciet is aangenomen dat de definiërende eigenschappen uniek zijn. Het probleem is dit: vaak zal er voor het definiëren van een term gekozen kunnen worden uit meerdere daarvoor in aanmerking komende eigenschappen. Dat maakt de niet-gekozen eigenschappen niet tot variabele eigenschappen.

Anders gezegd: er zijn soms meer mogelijkheden om voorbeelden correct te benoemen of te herkennen dan alleen controle uit te voeren op eigenschappen die in de definitie zijn opgenomen. Een formele definitie wordt gekozen op haar eenvoud; wanneer twee eigenschappen voldoende zijn, zal een derde eigenschap er niet extra bij opgenomen worden. Vooral in het onderwijs moet er op gelet worden dat essentiële eigenschappen veronachtzaamd worden, alleen omdat ze niet noodzakelijk zijn om een term formeel te definiëren.

Implicaties voor het onderwijs en de toetsing kunnen zijn:

- Gebruik in het onderwijs een set voorbeelden die zo heterogeen mogelijk is wat betreft de variabele eigenschappen.
- Soms kunnen niet-voorbeelden gebruikt worden, waarvoor

geldt dat ze alle definiërende kenmerken min een hebben.

-Bij benoemen en herkennen kan gevraagd worden op grond van welke eigenschappen een term is gekozen (benoemen) of een koppeling is gemaakt (herkennen).

- Misvattingen van studenten bestaan er soms uit dat een variabele eigenschap wordt aangezien voor een definiërende eigenschap.

4.6 Beschrijvende uitspraken

Bij de meeste termen zoals ze in het onderwijs gepresenteerd worden valt heel wat meer te vertellen dan alleen een definitie en een aantal typische voorbeelden. Slechts een deel van de betekenis wordt door die voorbeelden gedekt, allerlei andere zinvolle informatie komt slechts zijdelings of in het geheel niet in die voorbeelden tot uitdrukking. Bijvoorbeeld het onderwerp *inenting tegen mazelen*. Daarbij horen een aantal beschrijvende uitspraken die antwoord geven op de volgende vragen:

- Op welke leeftijd wordt deze inenting gegeven?
- Zijn er contra-indicaties?
- Welke risico bestaat er voor ernstige blijvende schade aan het zenuwstelsel?
- Zijn er speciale maatregelen te nemen bij kinderen met chronische astma en bronchiectase?
- Is er een gunstig effect bij verzwakte kinderen?
- Wat zijn de risico's van mazelen?

Vergelijkbare lijstjes *beschrijvende uitspraken* kunnen gegeven worden voor andere medische ingrepen en voor geneesmiddelen. Het gaat daarbij niet om extra informatie die bedoeld is om wat meer kleur en fleur te geven aan de presentatie in het onderwijs, integendeel, het is informatie die essentieel is voor de juiste toepassing in praktische situaties. Dat hoeft niet altijd zo te zijn: Er zijn veel onderwerpen die voor de student tot leven gebracht kunnen worden door naast voorbeelden ook achtergrondinformatie te geven over historische ontwikkeling, plaats binnen een theorie, alternatieve betekenissen en dergelijke.

Beschrijvende uitspraken zijn in een bepaald studieboek door-

gaans te vinden op de plaatsen die de index aanwijst. Wanneer het om meer dan incidentele beschrijvende uitspraken gaat komt men al gauw terecht bij hele stukken tekst die aan de behandeling van een bepaald begrip (term) gewijd zijn; in dat geval zit er meestal ook een bepaalde structuur in de gegeven beschrijving, waarbij het kennen van die structuur tot de doelen van het onderwijs hoort (zie ook hoofdstuk 6). Wat het schrijven van toetsvragen betreft, stellen deze beschrijvende uitspraken geen nieuwe problemen. Het voorbeeld over inenting tegen mazelen heeft al laten zien dat toetsvragen daarbij op voor de hand liggende wijze te formuleren zijn. Het risico van het op deze wijze vragen naar kennis rond een bepaalde term of een bepaald onderwerp is natuurlijk dat daarmee alleen de verbale beheersing wordt getoetst, het kunnen reproduceren of herkennen, terwijl het uiteindelijk toch de bedoeling zal zijn dat de student deze informatie in concrete situaties moet weten te produceren, en ook moet weten toe te passen. Wees daarom terughoudend met het stellen van louter 'kennisvragen' betreffende beschrijvende uitspraken.

5 Toetsvragen schrijven bij relaties tussen termen

Zelden of nooit is leerstof zo ongestructureerd van karakter dat volstaan kan worden met het behandelen van afzonderlijke termen, zonder dat *relaties* tussen die termen op enigerlei wijze van belang zijn. Het leren van rijtjes vreemde woorden komt daar nog het dichtst bij in de buurt, maar ook dan zullen er altijd relaties van gelijkenis, overeenkomst en verschil zijn, die het leren kunnen vergemakkelijken. Onderlinge verknooptheid van leerstof is zo vanzelfsprekend dat het moeite kost om afzonderlijke termen daaruit te isoleren. Zo zijn in het voorgaande hoofdstuk over termen een aantal relaties eveneens in de behandeling binnengeslopen. Denk aan definities, waarin vaak al sprake is van relaties tot andere termen. In dit hoofdstuk over relaties zullen afzonderlijke termen toch uitdrukkelijk aanwezig kunnen zijn, bijvoorbeeld in de vorm van bekend veronderstelde kennis over de afzonderlijke termen in een bepaalde relatie. De diverse soorten relaties hebben met elkaar een familiegelijkenis: ze hebben geen gemeenschappelijke set kenmerken, al kunnen ze wel enigszins gegroepeerd worden in gelijkgeaarde clusters. In iedere discipline zal het merendeel van de nu op te sommen relaties voorkomen, sommige meer geprononceerd dan andere.

-*Relaties van identiteit, overeenkomst of verschil.* 5.2 (Onderscheiden) behandelt het schrijven van toetsvragen bij dergelijke relaties. 5.3 geeft een bespreking van classificaties.

-*Relaties van samengaan, covariëren, oorzakelijkheid of volgorde.* Het zijn vooral deze relaties waaruit wetenschappelijke theorieën zijn opgebouwd. 5.5 (Wetmatige relaties) geeft hier de meest relevante vraagmogelijkheden bij. Zie ook de hoofdstukken 6 en 7.

-*Relaties van grammaticale of idiomatische aard, logische of wiskundige relaties en afbeeldingsrelaties* tussen verschillende beschrijvings- of representatiewijzen (verbaal, symbolisch, algebraïsch, picturaal): zie 5.1 en ook 5.5.

-*Handelingsvolgorde* (starten van een auto, staartdeling); zie 5.4. Relaties die *deterministisch* of *stochastisch* (wet-

matigheden), *dwingend of veroorlovend* (rechtswetenschap) zijn; zie ook 5.5 en 6.3.

- Andere specifieke relaties (Wolkers is de schrijver van *Een roos van vlees*, de relatie 'is onderdeel van' en dergelijke).

Het onderscheid dat iedere docent in een inventarisatie van relaties in de leerstof zal aanbrengen, is dat naar belangrijkheid. Net zoals termen te onderscheiden zijn in meer centrale, sterk met andere termen verknoopte, en meer perifere, met slechts een enkele andere term verknoopte termen, zo kunnen ook relaties op deze wijze gerangschikt worden. Belangrijke relaties komen in een groot aantal wetten telkens terug of het zijn simpelweg de relaties tussen centrale termen. Een andere wijze waarop de ene relatie belangrijker kan zijn dan de andere, is de praktische bruikbaarheid bij het oplossen van opgaven. Ieder (geformaliseerd) vak kent wel een kleine verzameling formules, wetten of betrekkingen, waarmee de meeste problemen kunnen worden opgelost. Soms worden die wel aangeduid als *kernbetrekkingen*. Andere wetten of betrekkingen zijn veelal af te leiden uit kernbetrekkingen.

Relaties zijn doorgaans een stap abstracter dan de afzonderlijke termen die erdoor gerelateerd worden. Bij het schrijven van toetsvragen zal er dan ook wat meer aandacht geschonken moeten worden aan het niveau van abstractie van de toetsvraag (zie 2.6). Hoe abstract sommige relaties ook mogen zijn, meestal is het mogelijk om toetsvragen op een redelijk concreet niveau te stellen. Wanneer het gaat om wetmatigheden die in algebraïsche vorm zijn uitgedrukt, zoals $E = mc^2$ of $F = ma$, dan moet er speciaal op gelet worden dat de toetsvragen niet (louter) de *algebra* betreffen; dat is een heel ander vak, op z'n best een hulpvak. De onderwijskundige reden voor deze aanbeveling is nog steeds dat een abstractie pas als abstractie wordt beheerst, wanneer de abstractie toegepast kan worden op concrete situaties. Terwijl de pragmatische kant van deze aanbeveling weer is dat een meer concreet niveau van vraagstelling het mogelijk maakt een grote variëteit van toetsvragen te schrijven rond een zelfde onderwerp (relatie).

Sommige relaties vormen samen een bepaalde *structuur*. Toetsvragen over de structuur, als zijnde iets anders of iets meer dan haar afzonderlijke relaties, kunnen dan geschreven worden zoals alle andere toetsvragen over een bepaald onderwerp. Afhankelijk van wat er over de structuur wordt gevraagd, kan de bijpassende paragraaf uit hoofdstuk 4 of 5, maar vooral ook uit hoofdstuk 6 (tekst) of 7 (aanpakken van problemen) geraadpleegd worden. Het is mij niet gebleken dat vragen over structuren zo bijzonder zouden zijn dat afzonderlijke behandeling nodig is. Dat geldt ook voor de speciale structuren als de classificatie, taxonomie of typologie, waar wel een afzonderlijke paragraaf (5.3) aan gewijd is, maar waar geen specifieke vraagmogelijkheden bij gegeven worden.

Uit deze uiteenzetting over relaties zal al duidelijk geworden zijn dat de vraagmogelijkheden die in hoofdstuk 4 voor afzonderlijke termen gegeven zijn, ook toepasbaar zijn voor relaties. Voor toetsvragen op basis van concrete voorbeelden maakt het geen verschil of het voorbeelden van een term danwel van een relatie zijn. Raadpleeg daarom voor vraagmogelijkheden bij voorbeelden van relaties ook hoofdstuk 4.

5.1 Vertalen en afbeelden

De groep relaties die het onderwerp van deze paragraaf vormen zijn de grammaticale en idiomatische relaties, de logische en wiskundige relaties, en vertalingen van afbeeldingen (representaties) van de ene in de andere vorm. Waar het bij deze relaties in het bijzonder om gaat is het kunnen formuleren, vertalen of in een andere afbeelding omzetten. Ofwel grammaticaal spreken of vertalen, en verbale uitspraken om zetten in symbolische of picturale.

Grammaticale relaties vormen met elkaar de grammatica van eigen of vreemde taal. Het gaat hier om de betekenis van uitspraken of zinnen, zoals bepaald door de onderlinge rangschikking van de gebruikte woorden en de vervoeging en verbuiging van de woorden waarvoor dat van belang is. Daarmee

is tegelijk aangegeven dat twee punten belangrijk zijn: 1. Is een gegeven zin op zich grammaticaal correct? En 2. Zo ja, is de betekenis van deze zin overeenkomstig de bedoelde betekenis? Het onderscheid zal iedereen aanspreken die zich bij gelegenheid moet bedienen van een niet geheel vertrouwde vreemde taal. Onbedoelde en zelfs absurde betekenissen kunnen gemakkelijk ontstaan door verkeerde woordkeus of door een ongeschikte grammaticale vorm.

Grammaticale regels (relaties) worden heel expliciet onderwezen; er bestaat een lange traditie in en het schrijven van toetsvragen zal weinig problematisch zijn. Immers, de traditie is een gelukkige, omdat de vragen zich op een heel concreet niveau zullen bevinden: gewoon *doen*. Dat doen kan verschillende vormen aannemen:

- dictee, waarbij naast de juiste spelling van woorden (zie 4.1) ook de juiste vervoeging van werkwoorden en de juiste verbuiging van andere woorden van belang zijn;
- opstel (voor zover dat tevens of uitsluitend op grammatica beoordeeld wordt);
- proefvertaling (van eigen taal in vreemde taal);
- uitspraken, of afzonderlijke zinnen, in vreemde taal vertalen;
- aanvulvragen, waarbij de uitgang van een woord moet worden ingevuld of een heel woord (deze wijze van vragen maakt het mogelijk meer gericht op grammaticale deelvaardigheden te toetsen, ongehinderd door een eventueel tekortschietende woordenkennis van de student).

VOORBEELD

She has [to bear] twelve children and the youngest was born yesterday.

Answer: borne

- keuzevragen, alleen dan te gebruiken wanneer de student geleerd heeft een bepaalde fout (bepaalde fouten) te vermijden; in dat geval kan (kunnen) deze als afleider(s) gebruikt worden.

Naast deze concrete vragen, toepassingsvragen kan men zeggen, kunnen meer abstracte vragen in aanmerking komen, bijvoorbeeld wanneer het van belang is dat de student kan aange-

ven waarom een bepaalde zin grammaticaal (on)juist is, of dat hij kan aangeven welke regel in een bepaald geval is toegepast en dergelijke. Wanneer deze vaardigheden tot de onderwijsdoelen horen, hetgeen voor lerarenopleidingen heel goed denkbaar is, dan kunnen dergelijke vragen ook als 'concreet' beschouwd worden. Er wordt immers met concrete voorbeelden (grammaticaal mogelijk onjuiste zinnen) gewerkt, waarbij de student de geschonden grammaticale regel moet kunnen noemen (wanneer de zin inderdaad onjuist is).

Vragen naar de vervoeging van een bepaald werkwoord (swear-swore-sworn bijv.) is wel echt abstract vragen, en voor dergelijke vragen lijkt in het onderwijs geen plaats naast de goede concrete vragen die hier als alternatief voor beschikbaar zijn. Het vragen van 'rijtjes' is doorgaans in eindtoetsen onterecht. Het (tijdelijk) beheersen of op kunnen zeggen van bepaalde 'rijtjes' heeft op zijn best een didactisch nut in een tussenstadium in het onderwijs.

Het vragen van nieuwe voorbeelden bij een bepaalde grammaticale regel past niet in een toets op grammaticaal correct taalgebruik. Dat wordt anders wanneer een actieve beheersing van de grammatica als systeem het doel van het onderwijs is. Bijvoorbeeld wanneer grammatica zelf het onderwerp van studie is, of wanneer de student grammaticale taalvaardigheid moet leren onderwijzen.

De mogelijkheden voor het gebruik van keuzevragen zijn heel beperkt. Ze kunnen wel geschreven worden, en in de Verenigde Staten worden ze op grote schaal gebruikt, maar het probleem is dat er geen verantwoording gegeven wordt (en ook moeilijk of niet te geven is) voor het gebruik van onzinnige afleiders.

AFRADER

I don't know *as I occur with your decision to try and run for office.*

- 1) that I concur in your decision to try to
- 2) as I concur in your decision to try and
- 3) that I concur with your decision to try to
- 4) as I concur with your decision to try to
- 5) no change

Dit is een vraag over idioom, de student moet het idiomatisch beste alternatief aankruisen of alternatief 5. De meeste studenten die de test afleggen zullen alle alternatieven bestuderen. Het probleem is dat het herkennen van niet-idiomatische zinsdelen misschien niet eens tot de onderwijsdoelen hoort, en hoogstwaarschijnlijk niet de vaardigheid is die met deze vraag *bedoeld* wordt te toetsen. Vragen als deze komen voor in de Amerikaanse New High School Equivalency Examination, een grote landelijke toets.

Logische en wiskundige relaties zijn onderling nauw verwant. Dat ik ze hier onderscheid heeft ook een 'didactische' reden: Wanneer er al vakken zijn waarin wiskundige hulpmiddelen in het geheel niet gehanteerd worden, dan zullen er zeker nog logische relaties een rol in spelen, al zou dat slechts zijn in de vorm van een 'argumentatiekunde'.

Het accent ligt op het 'vertalen' van gegeven tekst in symbolen en/of in symbolische relaties, en niet zozeer op het volgens de regels kunnen werken met formules of strings van symbolen (zie daarvoor o.a. 5.4). Het schematiseren van een tekst kan een voorbeeld zijn, met het weer terug weten te vertalen van de relaties die in zo'n schema heel abstract zijn aangeduid (zie 3.4).

Waar gaat het nu om bij het vertalen in logische of wiskundige relaties? Een beschreven situatie of gebeurtenis wordt er door teruggebracht naar een abstracte vorm. Dat is een eerste stap bij het maken van opgaven of het oplossen van problemen. Er wordt geabstraheerd van alle specifieke, concrete bijzonderheden. Al naar de aard van de beschrijving moet daarbij gelet worden op:

- getalsmatige kwantificering (schaalwaarde, aantal);
- woorden of uitdrukkingen 'vertalen' in gangbare symbolen;
- beschreven verbanden 'vertalen' in symbolische relaties (logische, wiskundige);
- bij kwantificering expliciet letten op de eenheden waarin gekwantificeerd wordt;
- expliciteren van impliciete gegevens (stilzwijgende veronderstellingen);
- gevolgtrekkingen waar de tekst aanleiding toe geeft.

Omdat het een voorstadium van het maken van opgaven of het oplossen van problemen betreft, zijn dit soort vragen voor eindtoetsen alleen geschikt wanneer het de bedoeling is bijzondere aandacht te geven aan het kunnen vertalen in logische of wiskundige relaties. Dat kan het geval zijn wanneer dat vertalen zo moeilijk is, of althans voor de studenten zo is, dat vragen van grotere omvang al in het stadium van het vertalen fout gaan; uit overwegingen van onder andere toets efficiëntie is het dan beter om het vertalen in symbolische vorm afzonderlijk te toetsen.

Afbeelden van de ene representatie in een andere, in het bijzonder het afbeelden van relaties in een nieuwe representatie. Een prototype van deze wijze van vertalen is het omzetten van algebraïsche relaties in meetkundige. Een ander bekend voorbeeld is het gebruik van het Venn-diagram bij de verzamelingenleer en de waarheidstabel in de logica. De behandelde vertaling van concrete situaties in abstracte, wiskundige symbolen en relaties is ook een voorbeeld. Het bouwen van *modellen* is er eveneens onder te begrijpen; dat kunnen dan zowel wiskundige modellen als schaalmodellen zijn.

Kaarten zijn afbeeldingen van een bepaald landschap vanuit een specifieke interesse (verkeer, grondstoffen) bekeken. Grafieken en tabellen zijn afbeeldingen van bestanden van gegevens, gericht op het met meer gemak kunnen aflezen van gewenste gegevens, samenhangen enz.

Ieder vak kent wel eigen specifieke afbeeldingsvormen, met daarnaast het gebruik van meer algemene, zoals de algebraïsche en meetkundige. Het doel van dergelijke afbeeldingen is in het algemeen het verkrijgen van een beter overzicht (bijna ergonomisch), en het daardoor gemakkelijker aan kunnen pakken van problemen. Niet zelden is het maken van een bepaalde afbeelding een onmisbare stap bij het oplossen van problemen. Soms is het maken van een schets niet beslist noodzakelijk, maar kan dit het vinden van oplossingen wel enorm vergemakkelijken (thermodynamica).

Meestal hebben bepaalde afbeeldingstechnieken wel een duidelijke eigen plaats in het onderwijs (de leerstof) gekregen, en vormt het kunnen afbeelden een vanzelfsprekend onderdeel

van de toets. Sommige vakken zouden vruchtbaar gebruik kunnen maken van bepaalde afbeeldingstechnieken die nu nog niet gebruikelijk zijn. Een voorbeeld is het leren schematiseren van leerstof (Mirande, 1981), het meer dan voorheen afbeelden van abstract gegeven stof in concrete voorbeelden, van verbaal aangeboden stof in visuele vorm.

Het maken van een afbeelding is een constructieve vaardigheid; het is dan weinig zinvol vragen te bedenken waarbij de student foutieve afbeeldingen als zodanig moet herkennen of verbeteren. Toetsvragen zijn vrijwel altijd van het type aanvul- of opstelvragen (constructievragen).

5.2 Onderscheiden

Een veel voorkomende relatie tussen termen is wat hen van elkaar *onderscheidt*. Bij onderscheidingen-zonder-meer, het onderwerp van deze paragraaf, gaat het om verschillen die niet op enigerlei wijze gesystematiseerd zijn (zoals een taxonomie of classificatie, zie 5.3). Dergelijke onderscheidingen zijn er in twee soorten: het *onderscheiden van gemakkelijk te verwarren termen* en de *juiste keuze van instrumenten, technieken en dergelijke*. Termen die gemakkelijk verward worden met andere termen worden als zodanig meestal al aangeduid: de student krijgt de waarschuwing 'niet te verwarren met' of er wordt gewezen op een voor de hand liggende misvatting. Bij de juiste *keuze* gaat het om het volgende probleem: een aantal verschillende technieken zijn ieder afzonderlijk behandeld en geoefend en de student die iedere techniek afzonderlijk beheerst, wordt gevraagd bij een gestelde opgave uit deze beschikbare technieken de juiste te kiezen. Ik heb het vermoeden dat veel auteurs en docenten zich onvoldoend bewust zijn van de ingrijpende omkering die hier aan de orde is. In het onderwijs oefent de student het correct toepassen van een *gegeven* techniek, terwijl later opgaven gesteld worden waarbij het allereerst van belang is de juiste techniek te *kies*en. Een algemeen bekend voorbeeld hiervan is de statistische hypothesetoetsing: een tamelijk groot aantal verschillende technieken wordt in het onderwijs afzonderlijk geoefend, in toepassingen op de specifieke problemen die met een bepaalde techniek aan te pakken zijn, maar bij de

afsluitende toets krijgt de student opgaven voorgeschoteld waarbij hij met de handen in het haar zit omdat hij niet getraind is in het vinden van de juiste techniek voor een dergelijke opgave.

VOORBEELD

Docent Jensen geeft een college over de mate waarin verschillen in intelligentie erfelijk, dan wel milieubepaald zijn.

Voorafgaand vullen studenten de enquêtevraag in of zij denken dat erfelijkheid, dan wel milieu de grootste bijdrage heeft. In zijn college wijst Jensen erop dat bij rigide sociale structuren, slecht onderwijs, en slechte gezondheidszorg, een sterke milieubijdrage is te verwachten; verbeteren deze omstandigheden, dan zullen nog overblijvende verschillen in intelligentie relatief in hogere mate erfelijk bepaald zijn. Na dit college krijgen de toehoorders dezelfde enquêtevraag weer te beantwoorden. Jensen wil nagaan of zijn betoog de meningen heeft beïnvloed. Welke statistische toets (eventueel meerdere) kan hij op zijn gegevens gebruiken?

- 1) Fischers exacte-waarschijnlijkheidstoets ja/nee
 - 2) tekentoets ja/nee
 - 3) Cochran-Q toets ja/nee
 - 4) McNemar toets voor significantie van veranderingen ja/nee
 - 5) chi-kwadraattoets voor een enkele steekproef ja/nee
- (Alternatieven 3 en 4 zijn goed.)

Voor een doorzichtige toetsing is het een eerste vereiste dat de student met dit type toetsvraag al tijdens het onderwijs vertrouwd is geraakt. Hij moet kennis hebben gemaakt met aanwijzingen of criteria op grond waarvan de passende techniek gekozen kan worden. In het gegeven voorbeeld is het voor de keuze van de juiste techniek van belang te weten wat het meetniveau van de verzamelde gegevens is (ordinaal, nominaal, interval), hoe groot het aantal steekproeven is (een, twee, of meerdere) en of deze steekproeven onafhankelijk van elkaar zijn. Door deze criteria langs te lopen, kan de juiste techniek worden gevonden. Dat kan gebeuren aan de hand van een tabel, of misschien moet de student de betreffende beslissingsstappen zo goed beheersen dat de keuze zonder een dergelijke tabel gemaakt kan worden. Om achteraf enig inzicht te ver-

krijgen in de overwegingen van de student, kan bovendien naar deze gegevens worden gevraagd (zonder dat het stellen van deze vragen de student extra aanwijzingen verschaft).

VOORBEELD

Het meetniveau van de enquêteresultaten is

1. Nominaal
2. Ordinaal
3. Interval

Het aantal steekproeven bedraagt

1. Een
2. Twee

Zijn er twee steekproeven, zijn deze dan

1. Afhankelijk van elkaar?
2. Onafhankelijk van elkaar?

Deze vraagmogelijkheid kan in veel vakken gebruikt worden. Ik geef een paar voorbeelden, in algemene termen geformuleerd:

- Welke van de aangegeven mogelijke oorzaken is het meest waarschijnlijk bij deze storing?
- Welke van de genoemde rechtsregels of wetsartikelen kunnen in het gegeven casus van belang zijn?
- Welke van de genoemde oplostechieken zijn voor dit probleem bruikbaar?
- Welke van de gegeven variantie-analytische modellen past bij de beschreven onderzoekopzet?

De vraagmogelijkheden bij dit soort keuzeproblemen bieden een aantal interessante varianten. Het eerder gegeven voorbeeld is opmerkelijk op deze punten:

- a. Het gegeven probleem is *concreet*, heeft een hoge werkelijkheidswaarde, ook al is het verbaal beschreven.
- b. De keuze-alternatieven zijn expliciet aangegeven.
- c. *Alleen* de keuze wordt gevraagd, *niet* de verdere berekening.

Bij dit soort vraagstelling is de verleiding aanwezig problemen op te geven in abstracte vorm.

AFRADER

Wanneer nominale data van twee afhankelijke steekproeven beschikbaar zijn, kan (kunnen) voor toetsing van de nulhypothese dat de gemiddelden niet van elkaar verschillen, gebruikt worden:

- 1) Fischers exacte-waarschijnlijkheidstoets ja/nee
- 2) de tekentoets ja/nee
- 3) Cochran-Q-toets ja/nee
- 4) McNemar-toets voor significantie van veranderingen ja/nee

Merk op dat in de abstracte vorm precies die informatie weggegeven wordt die nodig is om de juiste keuze te kunnen maken. Zo'n vraag kan op het gepaste moment didactisch nuttig zijn (leren omgaan met een gegeven tabel), maar heeft weinig of niets te maken met het vermogen van de student om bij een bepaalde set gegevens uit een gegeven onderzoek de geschikte statistische toets te vinden.

Alleen de keuze wordt gevraagd, en dat ligt soms niet al te zeer voor de hand, vooral niet wanneer het grootste deel van de onderwijstijd juist is gestoken in de rekenkundige bewerkingen die voor het uitvoeren van een gegeven statistische toets nodig zijn. Maar wanneer het belangrijker is dat de student de juiste toets weet te *kiezen*, is het wenselijk daar ook in de toetsvragen (en in het onderwijs) de nadruk op te leggen. Het is immers mogelijk om alleen de keuze te vragen, zonder verdere rekenkundige uitwerking. De omgekeerde mogelijkheid kan ook worden uitgebuit (alleen berekening vragen, geen keuze). Er zijn nog meer mogelijkheden om een *deel* van de oplossing van een complexe opgave te vragen, en dat is met name van belang voor wie bij voorkeur van aanvul- of keuzevragen gebruik wil maken:

- vraag naar relevante overwegingen (zijn de data nominaal? zijn de steekproeven onafhankelijk?);
- laat de student aangeven welke relevante informatie nog *ontbreekt* om een keuze (diagnose enz.) te kunnen maken;
- laat de student aangeven welke gegeven informatie *irrelevant* is voor de te maken keuze;
- laat de student een aantal mogelijk bruikbare technieken opsommen, waaronder de juiste te vinden moet zijn;
- laat de student hypothesen formuleren (bijv. over moge-

lijke diagnoses) en laat hij aangeven hoe de juistheid daarvan verder te onderzoeken is.

In het algemeen valt te overwegen om uit en bekend *oplossingsalgoritme* (zie 5.4) afzonderlijke stappen te isoleren voor afzonderlijke toetsvragen die kort en snel te beantwoorden zijn.

Het derde punt: *meerkeuze* of niet. Vaak zullen meerkeuzevragen goed te gebruiken zijn, omdat het aantal alternatieve antwoordmogelijkheden maar klein is en de student heel goed weet welke die mogelijkheden zijn. Daarmee is tevens bepaald wanneer meerkeuzevragen informatie 'weggeven': wanneer de geboden alternatieven er slechts enkele uit een veel groter in aanmerking komend aantal zijn of wanneer het niet vanzelfsprekend is dat studenten de alternatieve mogelijkheden op zich ook goed kennen. In die gevallen zijn aanvulvragen beter te gebruiken.

De overige vraagmogelijkheden hebben betrekking op wat in hoofdstuk 4 al uitgebreid besproken is:

- is het gegeven voorbeeld een oud dan wel een nieuw voorbeeld;
- is het gegeven voorbeeld 'typisch', betreft het een uitzondering enz.;
- het aantal gegeven voorbeelden;
- moet het gegeven voorbeeld (of de probleemstelling) eerst bewerkt (opgeschoond) worden voordat de keuze bepaald kan worden;
- is het voorbeeld concreet of abstract gepresenteerd.

Voor *gemakkelijk te verwarren* termen liggen de toetsingsmogelijkheden in lijn met die welke gegeven zijn in 4.3, 4.4 en 4.5. Daar valt aan toe te voegen dat bij te verwarren termen het gebruik van *keuzevragen* vaak op vanzelfsprekende wijze mogelijk is, door als alternatieven de termen waarmee verward kan worden (respectievelijk de voorbeelden van die termen) op te nemen. Het spreekt bijna vanzelf dat dergelijke opgaven alleen in en toets gebruikt worden wanneer het onderscheiden van deze gemakkelijk te verwarren termen in het onderwijs is behandeld.

'Bijna' vanzelfsprekend, want een veel voorkomend *misbruik van de meerkeuzevraag* is nu juist het vragen van onderscheidingen die in het onderwijs *niet* behandeld zijn en die evenmin

in de onderwijsdoelen thuisshoren. De moeilijkheid om goede 'afleiders' voor meerkeuzevragen te bedenken, leidt nogal eens tot vondsten die onbedoeld van de vraag een onderscheidingsvraag maken: de student moet het beste of het juiste antwoord vinden door alle geboden alternatieven te bestuderen en uit de onderlinge vergelijking zijn keuze te bepalen. Ook bij vragen waarbij de docent van de 'goede' student verwacht dat deze zonder omwegen het juiste alternatief weet te kiezen, kan zich dit zelfde probleem voordoen omdat maar zelden een alternatief in absolute zin juist of onjuist zal zijn. Anders gezegd: bij meerkeuzevragen kan zich voor ieder afzonderlijk alternatief hetzelfde probleem voordoen waaraan ja/nee-vragen nogal eens laboreren: dat het betreffende alternatief onder bepaalde (niet genoemde) condities juist, dan wel onjuist gemaakt kan worden. Dit onjuiste gebruik van de meerkeuzevraag is zo veelvuldig, dat iedere docent die zelf meerkeuzevragen schrijft er wel een aantal van in zijn verzameling zal hebben. Een verkeerde wijze van toetsen, met bovendien als bijzonder hinderlijk gevolg dat studenten verward kunnen raken. Wie zich niet heeft kunnen voorbereiden op vragen van het verwarbare-termen-type zal tijdens de toets pas met de verwarring geconfronteerd worden. Het vergt dan goede zenuwen, en een wat grotere slimheid of een beter geheugen dan anderen hebben, om die vragen goed te maken.

Gemakkelijk te verwarren termen moeten dus op enigerlei wijze onderscheiden leren worden. Dat kan soms louter op basis van vergelijkende *waarneming* gebeuren: in dat geval zullen toetsopgaven ook waarnemingsopgaven moeten zijn (concrete objecten, foto's, dia's, film, geluid e.d. gebruiken). Het verschil tussen waarneembare termen zal doorgaans ook verbaal te beschrijven zijn en zo'n verbale beschrijving kan voor de student betekenis krijgen wanneer deze gekoppeld is aan vergelijkende waarneming. In die gevallen is het mogelijk toetsvragen te baseren op de verbale beschrijving van het onderscheid, maar het risico is daarbij dat zo'n vraag nogal op een abstract niveau zit (er wordt niet gewerkt met voorbeelden, maar met die ene verbale beschrijving). Wanneer het onderscheid kan worden aangegeven in termen van kritische kenmerken wordt het onderscheid naar een meer formeel niveau getild. Toetsvragen zijn dan gemakkelijker te constru-

eren, ook al is het abstracte vragen naar onderscheidende kenmerken niet aan te raden (zie ook 4.6).

Voor gedetailleerde vraagmogelijkheden kan in 4.3, 4.4 en 4.5 gekeken worden en is ook het laatst gegeven rijtje mogelijkheden, voor problemen van juiste keuze van instrumenten, van toepassing. Juist omdat het om een relatie tussen gemakkelijk te verwarren termen gaat, is ook de enkelvoudige vraag om een voorbeeld te benoemen impliciet een vraag naar het onderscheiden van beide (of meerdere) termen die met elkaar verward kunnen worden; de meerkeuzevraag ligt dan wel voor de hand, maar de gewone aanvulvraag kan het ook prima doen. Hetzelfde geldt voor het vragen van een nieuw voorbeeld bij een van de termen die verward kunnen worden.

Tot slot twee opmerkingen. Vergeet niet dat er ook termen zijn die wezenlijk aanleiding geven tot verwarring, omdat ze niet objectief zuiver van elkaar te onderscheiden zijn. Wie probeert om onderscheidingsregels op te stellen voor dergelijke termen is waarschijnlijk op de verkeerde manier met de stof bezig. Veel van de natuurlijke categorieën, die toch bij uitstek het voorwerp van wetenschappelijk onderzoek zijn, hebben geen duidelijk ten opzichte van elkaar afgebakende grenzen, en in die gevallen is het beter de nadruk te leggen op de meer (proto)typische voorbeelden, en hoe die van elkaar te onderscheiden zijn.

De tweede opmerking is in zekere zin het omgekeerde van de opmerking hiervoor: er zijn ook termen die in het geheel niet moeilijk van elkaar te onderscheiden zijn. Toch kan het van belang zijn dat de student zich van de verschillen bewust is en dat het onderscheid ook expliciet (en niet als vanzelfsprekend) gemaakt kan worden.

5.3 Classificaties

Een aparte paragraaf is gewijd aan een apart onderwerp: classificaties. Een classificatie is een ordening die opgelegd wordt aan de verschijnselen die in een bepaalde tak van wetenschap bestudeerd worden. Een aantal voorbeelden: Mendelejevs periodiek systeem van scheikundige elementen; de taxonomie voor planten, respectievelijk dieren, in de biologie; typolo-

gieën zoals in de persoonlijkheidsleer wel gehanteerd worden: min of meer logische indelingen zoals in de rechtsgeleerdheid worden gehanteerd; en classificaties zoals door geneeskundigen gebruikt worden. Een classificatie is meestal een *categorische* ordening (biologische taxonomieën); de verschillende categorieën overlappen elkaar niet en ieder verschijnsel of object hoort in een en niet meer dan een categorie thuis. Wanneer de categorieën elkaar overlappen, er veel grensgevallen zijn en het lidmaatschap van een categorie bepaald wordt door het in meerdere of mindere mate hebben van bepaalde eigenschappen, wordt in plaats van een classificatie veelal over een *typologie* gesproken. Het opstellen van een typologie is niet zinvol wanneer de verschijnselen niet op een of andere wijze geclusterd zijn: anders worden verschijnselen niet meer gekarakteriseerd door ze te 'typeren' maar door ze een bepaalde waarde op een schaal voor een bepaalde eigenschap (meerdere eigenschappen) toe te kennen. Een voorbeeld: windsterkte uitgedrukt in kilometers per uur, in plaats van op de schaal van Beaufort.

Ik heb het onderwerp classificaties een 'apart' onderwerp genoemd. De reden is deze: bij classificaties zijn geen toetsvragen te schrijven die specifiek zijn voor classificaties. Omdat classificaties echter een heel geprononceerde plaats in de leerstof kunnen hebben, wil ik er toch een bespreking aan wijden, ongeveer op de wijze waarop in hoofdstuk 3 over definities is gesproken. Deze uiteenzetting moet de docent voldoende inzicht verschaffen in de aard van een classificatie zoals die in de eigen leerstof behandeld wordt, om een juiste keuze te kunnen doen met betrekking tot de aard van wat in dit geval op relevante wijze over deze classificatie gevraagd kan worden. Waaruit wordt die keuze gemaakt? Ik kan nu al een opsomming van de meer voor de hand liggende mogelijkheden geven:

1. reproduceren van de classificatie (biologie: de taxonomie 'klasse orde familie soort species');
2. aangeven op welk principe (welke principes) een classificatie is gebaseerd;
3. aangeven waarin een bepaalde categorie in de classificatie zich onderscheidt van andere categorieën (waarin onderscheiden zoogdieren zich van andere klassen van dieren);
4. aangeven waarin twee categorieën zich van elkaar onder-

- scheiden, respectievelijk waarin zij overeenkomen;
5. determineren van verschijnselen of objecten;
 6. aangeven op grond van welke overwegingen een verschijnsel of object is gedetermineerd.

Het is niet zinvol deze vraagmogelijkheden hier tot in detail uit te werken, omdat daarvoor naar andere paragrafen in dit boek verwezen kan worden. Vraagmogelijkheden 1. en 2. behelzen het terugvragen van informatie (zie hoofdstuk 6, vragen over tekst), 3. en 4. verwijzen naar wat in 5.2 over *onderscheiden* is gezegd, terwijl 5. en 6. ofwel een vorm van eenvoudig benoemen betreffen (hoofdstuk 4), ofwel een stappenvolgorde of *algoritme*, en dat wordt in 5.4 behandeld.

Classificaties zijn er in soorten, en je zou dan ook een classificatie van classificaties kunnen opstellen. Hoewel dat niet noodzakelijk tot een paradox hoeft te leiden, geef ik er de voorkeur aan enkele vormen van classificatie op een meer informele wijze te bespreken. Eenvoudige classificaties zijn gebaseerd op een logische ordening of een stipulatieve ordening (bij afspraak). Voorbeeld: temperaturen boven, respectievelijk bene den nul graden Celsius. Deze temperatuurindeling heeft een zeker aantrekkelijkheid vanwege de eenvoud van het gekozen principe: het vriespunt van een voor de mens zo belangrijke stof als water, maar het blijft een tamelijk willekeurige indeling. In de geneeskunde komen dergelijke classificaties wel voor, wanneer er met administratieve of wettelijke normen wordt gewerkt. In de rechtswetenschap is er een behoefte aan eenduidige normen, ook waar die niet op natuurlijke wijze bij de betreffende verschijnselen of objecten passen; het vak kent dan ook veel indelingen die bij afspraak (bij wet, via jurisprudentie, in de dogmatiek) zijn vastgelegd. De zwakte van classificaties op deze basis, bij afspraak, is dat de classificatie geen of vrijwel geen theoretische betekenis heeft, of met een anglicisme aangeduid: geen systemische betekenis heeft. De classificatie is gericht op een heel direct en praktisch doel en heeft daarbuiten geen nut. In het onderwijs is het dan ook niet raadzaam om aan deze 'kunstmatige' classificaties veel aandacht te besteden. Inleidingen in de rechtswetenschap kenmerken zich nogal eens door een overmaat van juridisch jargon dat op deze wijze is ingedeeld. Bijvoorbeeld: 'Het staats-, administratief en privaatrecht bevatten tal van regels,

hoe de mensen zich in de maatschappij horen te gedragen. Deze rechtsnormen kan men onderscheiden in *verplichtende* en *veroorlovene* normen; verplichtende normen weer in normen die gebieden iets te doen, en normen die, omgekeerd, een bepaald handelen *verbieden*. De overgangen zijn 'vloeiend' (Enschedé, 'Strafrecht en strafvordering', in: Bakels, 1971). Of het van belang is dat studenten deze terminologie leren hanteren, zal vooral moeten afhangen van het feit of de terminologie betekenis heeft boven wat zij zelf al uitdrukt.

Classificaties worden pas echt interessant wanneer de variabele of variabelen op grond waarvan de indeling gebeurt, samengaan met andere belangrijke variabelen. Verdeel je de mensheid in vrouwen en mannen op grond van primaire geslachtskenmerken, dan is de mensheid ook op een breed scala van andere kenmerken langs dezelfde lijn verdeeld. Of anders gezegd: de informatie op grond waarvan de indeling kan worden gemaakt, maakt ook gevolgtrekkingen mogelijk over kenmerken die niet gegeven zijn, Hempel (1952, blz 53): 'De rationele kern van het onderscheid tussen natuurlijke en kunstmatige classificaties zit in de overweging dat bij de zogenaamde natuurlijke classificaties de bepaalde kenmerken meestal geassocieerd zijn met andere kenmerken waarvan ze overigens logisch onafhankelijk zijn. (...) taxonomische categorieën als soort, species, etc., zoals gebruikt in de biologie, bepalen klassen waarvoor geldt dat organismen diverse biologische kenmerken met elkaar gemeen hebben anders dan die welke de klasse in kwestie definiëren; vaak geven de zo bepaalde groeperingen ook relaties weer die de phylogenetische oorsprong betreffen.' De biologische taxonomie heeft dan ook theoretische of systemische betekenis, gaat veel verder dan het louter op beschrijvende kenmerken indelen van de verschijnselen. In de geneeskunde kunnen we classificaties tegenkomen die gebaseerd zijn op prognose (bijv. leukemie en multiple sclerose), op behandelingsaanpak (bijv. astma, artritis) of op beschrijvende kenmerken (bijvoorbeeld syndromen, vele huidziekten, de meeste hartziekten) (Murphy, 1976, blz. 104). 'Binnen de medische taxonomie is de diagnose het brandpunt van het denken bij de behandeling van de patiënt; terug naar pathogenese en etiologie, voorwaarts naar prognose en behandeling. Zodoende verschaffen diagnostische categorieën de locaties waar klinici de waarnemingen uit

hun klinische ervaring opslaan, en bepaalt de diagnostische taxonomie de patronen waarin klinici waarnemen, denken, zich zaken herinneren, en handelen.' (Mezzich & Solomon, 1980, blz. 2).

De rijkste classificaties zijn die welke zijn gebaseerd op essentiële kenmerken. Hecht niet te veel betekenis aan het woordje 'essentieel', dat leidt tot filosofische problemen. Denk aan het periodiek systeem van Mendelejev, dat gebaseerd is op aantallen protonen in de atoomkern en op de daarmee corresponderende aantallen elektronen. In de geneeskunde is te denken aan classificatie gebaseerd op het achterliggende 'mechanisme' van de ziekte (bijv. tuberculose, de bloedziekten). Zo'n classificatie is het brute beschrijvende niveau overstegen en kan veeleer gelijkgesteld worden aan een wetenschappelijke theorie.

Het ordenend beschrijven van de verschijnselen is veelal een eerste stap op de lange weg naar een werkbare theorie (dat is een vruchtbare theorie). Geen wonder dat een afzonderlijke hulpwetenschap aan het ontstaan is, gericht op methoden en technieken voor het ordenen van verschijnselen (zie de aanbevolen literatuur aan het eind van dit hoofdstuk).

Classificaties hebben in veel disciplines een belangrijke plaats, ook in de leerstof die studenten wordt aangeboden. Desondanks zijn er geen vraagmogelijkheden die specifiek zijn voor classificaties. Dat betekent dat de docent zich moet afvragen op welke wijze de student met een bepaalde classificatie moet weten om te gaan, en dat al naar gelang het daarop gevonden antwoord de vraagmogelijkheden in een ander deel van dit boek te vinden zijn. Voor (medische) diagnostiek bijvoorbeeld zal hoofdstuk 7 (problemen stellen) in zijn geheel van belang zijn.

5.4 Stappenschema's, algoritmen, routines

Ieder vak kent wel een aantal technieken, methoden, rekenprocedures en dergelijke, zonder welke het vak niet beoefend zou kunnen worden. In algemene bewoordingen gaat het om een volgorde van handelingen die zo goed beheerst moet worden dat ze probleemloos uitgevoerd kan worden wanneer dat

in de uitoefening van het vak, bij het maken van ingewikkelde opgaven of bij het oplossen van problemen te pas komt. Soms moeten routines zo perfect beheerst worden dat zij foutloos kunnen verlopen ook zonder dat aan het uitvoeren van de routine bewust aandacht wordt geschonken. Dat is bijvoorbeeld het geval met de vanzelfsprekende vaardigheid de moedertaal grammaticaal correct te spreken. Deze vaardigheid is niet bewust aangeleerd en slechts weinig mensen die hun grammatica correct hanteren zouden de grammaticale regels ook kunnen verwoorden. Een ander voorbeeld van algoritmen die bijna blindelings beheerst worden zijn de eenvoudige rekenkundige bewerkingen. Deze zijn wel expliciet geoefend en wie ze gebruikt is eveneens in staat uit te leggen wat hij doet en waarom. Autorijden is eveneens een 'routine' waarbij de ervaren bestuurder geen expliciete aandacht meer hoeft te geven aan de vraag welke handelingen hij wanneer moet verrichten, en hoe hij dat moet doen. Het aanpakken van wetenschappelijke problemen is voor de onderzoeker meestal zo vanzelfsprekend geworden dat het moeilijk valt om alle ondernomen 'denkstappen' aan een ander duidelijk te maken (een verschijnsel dat versterkt wordt doordat sommige stappenreeksen in een keer gemaakt worden op basis van 'ervaring', een ervaring die voor anderen en zeker voor studenten geenszins vanzelfsprekend is). Naast routines die dergelijke gedetailleerde (denk)-handelingen betreffen, zijn er ook stappenschema's die een globaal karakter hebben, die meer een geheugensteuntje zijn bij het uitvoeren van complexe opdrachten (oplossen van problemen, plannen van een scriptie of opzetten en uitvoeren van een onderzoek) (zie hoofdstuk 7).

Een vak dat bijna een verzameling van algoritmen genoemd kan worden, waarbij onderzoekers gespitst zijn op het ontwikkelen van nieuwe en betere algoritmen, is de wiskunde. Voorbeelden te over: van het eenvoudige optellen of het maken van een straatdeling tot het ontwerpen van een computerprogramma's, terwijl er ook specialiteiten zijn zoals het leveren van het bewijs dat een bepaalde klasse van problemen geen algoritme kent, dat voor ieder probleem een oplossing oplevert. Om het onderwerp te concretiseren geef ik als voorbeeld van een lineaire (1e graads) vergelijking in een onbekende.

VOORBEELD

Los op: $8x + 3 = 3x + 18$

Beschrijf het gebruikte algoritme.

Antwoord. Gevraagd wordt de waarde van x , of: $x = \dots$

stap 1: een constante in het linkerlid kan verwijderd worden door zowel links als rechts deze constante af te trekken.

Resultaat: $8x = 3x + 18 - 3 = 3x + 15$

stap 2: de factor x in het rechterlid kan verwijderd worden door deze zowel links al rechts af te trekken.

Resultaat: $8x - 3x = 5x = 15$

stap 3: de coëfficiënt van x kan gelijk 1 gemaakt worden door links en rechts door deze coëfficiënt te delen.

Resultaat: $5x / 5 = 15 / 5$, ofwel $x = 3$.

Zo'n algoritme kan op verschillende manieren geformuleerd worden of in een schema uitgebeeld. Doorgaans is de stappenvolgorde kritisch, een andere volgorde zou een ander (onjuist) resultaat opleveren. In het gegeven voorbeeld zou het uitvoeren van stap 3 voor stap 2 tot problemen leiden; stap 1 en 2 zijn onderling wel verwisselbaar.

Het gegeven voorbeeld is een eenvoudig algoritme waarin geen logische beslissingen voorkomen, waarin niet aan specifieke voorwaarden voldaan hoeft te zijn om de volgende stap te kunnen uitvoeren (anders dan dat de voorgaande stap is uitgevoerd). De meeste algoritmen of stappenschema's kennen wel van die logische beslispunten of 'tests', met daaruit resulterende 'loops' naar een eerdere stap of *short cuts* naar een stap verderop in het algoritme. In het volgende voorbeeld is dat geïllustreerd, en dat is tegelijk een voorbeeld van een meer globaal stappenschema, in tegenstelling tot een specifiek algoritme zoals hierboven gegeven.

VOORBEELD

Voor het oplossen van een juridische casuspositie kan de student het volgende stappenschema hanteren:

stap 1: breng de casus in kaart

stap 2: geef voorlopige juridische vertaling van feiten in casus

stap 3: selecteer op basis van resultaten van 2 relevante rechtsregels

stap 4: bepaal op basis van resultaat van 3 voorwaarden waaronder de vertaling in stap 2 steek houdt

stap 5: is aan voorwaarden uit 4 niet voldaan in 1, dan terug naar stap 2 (een 'test' stap)

stap 6: pas regels uit 3 toe op casus en bepaal gevolg(en)

stap 7: zijn gevolgen niet aanvaardbaar, dan terug naar stap 2 ('test' stap)

stap 8: formuleer uitspraak

(Een heel vrije bewerking van een stappenschema gegeven door Crombag, De Wijkerslooth en Van Tuyll van Serooskerken, 1972).

Naast deze stappenschema's met een voorgeschreven volgorde zijn er ook handelingsvoorschriften die geen bepaalde volgorde kennen, maar wel volgens strikte regels bepaald zijn. Dat worden wel 'technieken' genoemd. Een toetsvraag gericht op zo'n techniek is:

VOORBEELD

Maak een tekening van zes figuren. Plaats ten minste een van de personen op de voorgrond, evenzo ten minste een persoon op de achtergrond, en tenminste een in een tussenpositie. De relatie tussen de figuren moet de regels van het *rechtlijnig perspectief* weerspiegelen.

(Wilson, in: Bloom, Hastings en Madaus, 1971, blz. 550)

Laten we voor het gemak de term 'algoritme' gebruiken voor het hele scala van handelingsvoorschriften: zowel stappenschema's, technieken en routines, als algoritmen zelf. Het is van enig belang om de vraag te beantwoorden welke rol algoritmen in het onderwijs hebben en speciaal in de eigen discipline. Meestal krijgen algoritmen in het onderwijs een heel duidelijke plaats toegewezen: de studenten worden er uitgebreid mee in kennis gebracht en in geoefend. Maar is dat ook voldoende? Aan het gebruik van een algoritme zitten meer kanten dan alleen het kunnen toepassen. Een heel belangrijk punt is wel dat de student in staat moet zijn het algoritme toe te passen *wanneer de situatie daar om vraagt*. Daarnaast zou het niet zo gek zijn wanneer de student voldoende inzicht zou hebben in de achtergronden van het algoritme om ook in staat te zijn dezelfde principes in ongewone situaties toe te passen (of

gewoon in het dagelijks leven, waar ook heel wat problemen op te lossen zijn). Dat vraagt niet alleen een goede routinematige beheersing, maar vooral het bewustzijn van het hoe en waarom van het algoritme.

Daarmee ontstaat een indeling van twee categorieën toetsvragen rond algoritmen: het correct *uitvoeren* van het algoritme, en bij een opgegeven probleem het juiste algoritme weten te *kiezen*.

De meest voor de hand liggende vraagmogelijkheden behoeven nauwelijks de verheldering aan de hand van extra voorbeelden:

- Vragen gericht op het juist kunnen *uitvoeren* van het algoritme. Om zo specifiek mogelijk op dat uitvoeringsaspect te mikken, kan in de formulering van de vraag erop gewezen worden dat een met name genoemd algoritme voor het oplossen te gebruiken is. Daarmee wordt voorkomen dat studenten die het algoritme wel kunnen *uitvoeren* de vraag missen doordat ze het juiste algoritme niet weten te *kiezen*.

- Vragen gericht op het weten te *kiezen* van het juiste algoritme. Dat is een bekend soort vraag: wanneer de student moet kiezen tussen meerdere in aanmerking komende algoritmen, kan het een onderscheidingsvraag zijn (zie 5.2). Meerkeuzevragen zijn dan mogelijk bruikbaar. Let op een goede keuze van het probleem dat in de vraagstelling voorkomt: het probleem moet bij voorkeur in concrete termen gesteld zijn en moet vallen binnen het bereik van verschillende problemen waarop de student het bedoelde algoritme heeft leren toepassen. Met name is het onjuist in de eindtoets opgaven te gebruiken die net een slag moeilijker zijn dan in het onderwijs behandelde opgaven.

- Wanneer een specifieke stap uit het algoritme erg moeilijk blijkt, kunnen vragen geformuleerd worden om alleen deze stap te doen uitvoeren.

- Wanneer het van belang is dat studenten zich bewust zijn van de aard van het algoritme, de reden voor bepaalde stappen weten, een afleiding kunnen geven, weten op welke algemene principes het algoritme berust, dan kunnen daar ook vragen over gesteld worden.

Omdat algoritmen uiteindelijk bedoeld zijn om bepaalde problemen op het eigen vakgebied tot een oplossing te (helpen)

brengen, zullen ze in een eindtoets vaak voorkomen als onderdeel van andere opgaven. Wanneer dergelijke opgaven vaak fout gemaakt worden op de keuze, dan wel de uitvoering van het algoritme, is het efficiënter om uitgebreidere opgaven minder afhankelijk van de kennis van algoritmen te maken, en de laatste afzonderlijk te toetsen.

Stappenschema's spelen vaak een rol bij het gestructureerd aanpakken en oplossen van problemen, zoals ook de bedoeling van het als voorbeeld gegeven stappenschema voor het aanpakken van juridische casus is. Zie verder vooral hoofdstuk 7.

5.5 Wetmatige relaties

Het schrijven van toetsvragen rond bijvoorbeeld natuurkundige wetten is niet lastig of problematisch (zolang niet ten onrechte van meerkeuzevragen gebruik wordt gemaakt). Het sterk geformaliseerde karakter van deze relaties, het nauwe verband tussen (theoretische) relatie en (concrete) verschijnselen en een lange traditie in het formuleren en gebruiken van vragen dragen daar alle toe bij. Voorbeelden van toetsvragen zijn voor de docenten uit deze vakgebieden zo vanzelfsprekend dat ik ze hier niet nog eens hoeft te geven. Ik zal niettemin toch enkele voorbeelden uitwerken, omdat het voor docenten uit in dit opzicht minder bevoorrechte vakgebieden toch aardig is om te zien hoe een wat strakkere formalisering als bijkomend voordeel heeft dat het schrijven van toetsvragen vereenvoudigd wordt.

Mag het schrijven van toetsvragen hier dan al min of meer vanzelfsprekend zijn, dan neemt dat toch niet weg dat er wel een risico bestaat dat de vragen die de docent gewoon is te gebruiken een zekere eenzijdigheid hebben, dat bepaalde aspecten van de leerstof onvoldoende of in het geheel niet in de toetsvragen gerepresenteerd zijn. Er zijn ook enkele vraagmogelijkheden, zie 6.3 en 6.4 (analyse en inferentie) en hoofdstuk 7 (bedenken en oplossen van problemen).

Er zitten heel wat aspecten aan een enkele wetmatige relatie (verder *wet* genoemd). Wetten zijn nogal eens genoemd naar hun 'ontdekker' en/of hebben een speciale benaming zoals

‘eerste hoofdwet van de thermodynamica’. De wet is in verbale vorm weer te geven, zoals stellingen in de wiskunde, en in symbolische (meestal een wiskundige) vorm. De meeste wetten betreffen relaties tussen twee of meer theoretische termen. De *afleiding* van de wet is van belang; afleiding uit andere wetten, uit empirische waarnemingen of de historische ontwikkeling. Toepassingen van de wet zijn het meest interessant, ook in het onderwijs, en dat onderwerp vergt een afzonderlijke uitwerking. Dan zijn er de relaties tot andere wetten binnen een en dezelfde theorie en de toepassingen door het combineren met andere wetten. En tenslotte de plaats die de wet binnen de theorie in haar geheel inneemt en hoe zij afhangt van bepaalde veronderstellingen waarop de theorie gebouwd is.

VOORBEELD

Newtons tweede bewegingswet. De naam geeft de opsteller van de wet de eer die hem toekomt en geeft het karakter van de wet aan. De wet luidt dat de netto kracht die op een lichaam inwerkt proportioneel is aan de massa van het lichaam en aan zijn versnelling, waarbij de richting van de kracht gelijk is aan de richting van de versnelling. De wet relateert aan elkaar: kracht, massa en versnelling (werkend op, resp. Van een lichaam), en moet zo verstaan worden dat de versnelling van het lichaam *veroorzaakt* wordt door de netto kracht die op het lichaam inwerkt.

De algebraïsche formulering van de wet is

kracht = *massa* * *versnelling*, in symbolen: $F = ma$.

De vergelijking ziet eruit als een algebraïsche, maar dat is niet helemaal correct: verondersteld wordt dat de grootheden uitgedrukt zijn in bij elkaar passende *eenheden*, bijvoorbeeld de kracht in newtons (N), de massa in kilogrammen (kg), en de versnelling in meters per seconde-kwadraat (m / s^2). Bij bewerkingen op de formule $F = ma$ is het verstandig daarbij de eenheden ook ‘mee te nemen’.

Voor de afleiding van de wet kan Newton zelf geraadpleegd worden (*Principia*, volume 1, *The motion of bodies*) of de wet kan afgeleid worden uit waarnemingen verzameld in een geschikt laboratoriumexperiment enz.

Toepassingen komen straks afzonderlijk ter sprake.

Er zijn o.a. verbanden met de eerste en derde bewegingswet, en met Newtons wet van de universele zwaartekracht.

De wet veronderstelt niet dat luchtweerstand te verwaarlozen is (dat is een van de krachten die de netto kracht bepalen). Er wordt verondersteld dat het lichaam rechtlijnig beweegt wanneer er geen netto kracht op inwerkt (Einstein heeft het daar moeilijk mee). Newton presenteerde zijn bewegingswetten overigens als 'axiomata of bewegingswetten', en in het Latijn 'Mutationem motis proportionalem esse vi motrici impressae, et fieri secundum lineam rectam qua vis illa imprimatur.'

Er zijn talrijke vraagmogelijkheden die betrekking hebben op deze verschillende aspecten, naast de toepassingen; maar dat zijn precies de vraagmogelijkheden die in hoofdstuk 4 al behandeld zijn.

Wetmatige relaties kunnen gebruikt worden om verschijnselen te *verklaren* of te *voorspellen*. Speciale vormen van voorspellen zijn; het schatten of berekenen van gevraagde onbekende grootheden en het opstellen van hypothesen en/of het ontwerpen van een onderzoek opzet. Een bijzondere vorm van verklaren waar de student nog wel eens mee te maken krijgt: verklaren waarom een onderzoeker een bepaalde onderzoek opzet kiest of een bepaalde handeling verricht. Dat moet niet opgevat worden als een psychologische vraag over de onderzoeker: bedoeld wordt dat de student verband legt tussen de acties van de onderzoeker en bekende wetten in dat vak (dat hij kan formuleren wat de werkhypothese van *een ander* is).

De voor de hand liggende 'toepassingsvragen' zijn nu:

- een nieuw voorbeeld van het toepassen van een bepaalde wet laten geven;
- toepassing van een bepaalde wet ook als zodanig laten herkennen;
- algoritmische bewerkingen: onbekende variabele berekenen uit de verstreekte gegevens door toepassen van de wet, eventueel door toepassen van een bepaalde wet in combinatie met een of meerdere andere wetten;
- een opgave laten herkennen als opgave die met behulp van een bepaalde wet is op te lossen (eventueel in combinatie met één of meer andere wetten);

- een opgave laten vertalen in wiskundige vorm (zie 5.1);
- en probleem dat in alledaagse bewoordingen of waarnemingen is geformuleerd laten omzetten in een meer vakmatige formulering, waaruit duidelijk kan zijn welke wet hier tot een oplossing zou kunnen leiden.

VOORBEELD

Geef een voorbeeld van toepassing van de tweede bewegingswet van Newton.

Antwoord. Op een lichaam met een massa van 1 kilogram wordt een netto kracht uitgeoefend van 1 Newton. De versnelling is dan $a = F / m = 1 \text{ N} / 1 \text{ kg} = 1 \text{ m} / \text{s}^2$.

Het gegeven antwoord is niet meer dan een beschrijving van de wet in haar theoretische termen. Is het de bedoeling dat de student een voorbeeld geeft uit het dagelijks leven in concrete termen, vraag dat dan ook.

VOORBEELD

Een auto met een massa van 1000 kg verhoogt zijn snelheid in 5 seconden van 15 m/s tot 20 m/s. Welke netto kracht wordt er op de auto uitgeoefend?

(Hint: Bereken eerst de versnelling, en gebruik dan Newtons tweede bewegingswet.)

Het aantal vraagvarianten voor de tweede bewegingswet van Newton is eindeloos, en daar komen dan nog de varianten bij die ontstaan door combinatie met andere wetten (combinaties waarmee de student door uitgebreide oefening ook vertrouwd is geraakt).

Bij deze toetsvragen over wetten is het van belang zicht te houden op de afstemming van wat in feite gevraagd wordt op de onderwijsdoelen. Wanneer het uitvoeren van wiskundige bewerkingen in de doelen een ondergeschikte plaats bekleedt, is het zaak de toetsvragen zo te schrijven dat wiskundige bewerkingen daarin geen extra moeilijkheidsfactor vormen. Dat heeft ook pragmatische betekenis: wanneer berekeningen tijdrovend zijn en slechts van marginaal belang voor de onderwijsdoelen valt er een hoop toetstijd op betere manier te besteden dan door berekeningen te vragen. Het zijn niet alleen opgaven rond natuurkundige wetten die vatbaar zijn voor re-

kenkundige 'overkill', dat kan ook maar al te gemakkelijk gebeuren bij de vakken statistiek en methodologie in de sociale wetenschappen. Zoek een goede balans in de toetsvragen en de verschillende al eerder genoemde aspecten: formuleren van het probleem, vertalen van het probleem, keuze van toepasselijke wetten bij het gegeven probleem, wiskundige bewerkingen om de (een) oplossing te verkrijgen. Wanneer het mogelijk is toetsvragen zo te schrijven dat ze op een of twee van deze aspecten zijn toegespitst, dan heeft dat voor de toetsing grote voordelen. Dan komt het niet meer voor dat een grote opgave fout gemaakt wordt alleen omdat in de vertaling van het probleem een vergissing is gemaakt.

Varianten op een zelfde opgave ontstaan door in de opgave gegevens op te nemen die niet ter zake zijn (en door de student als zodanig 'ontmaskerd' moeten worden), door de probleemformulering concreter dan wel abstracter te maken, waardoor de student meer dan wel minder moeite moet doen om probleem in exacte termen te vertalen. Een mogelijkheid tot variëren ligt in het aanbieden van te weinig gegevens, met de opdracht aan de student om zelf te vragen om extra gegevens die nodig zijn (bijv., bij juridische casus, bij medische diagnostiek). Denk ook aan de invloed van *impliciete context* op de moeilijkheid van de opgaven: tijdens het onderwijs zijn de opgaven nogal direct gekoppeld aan het hoofdstuk waarin de bijbehorende theorie werd gegeven, waarmee impliciet gegeven is welke wetten waarschijnlijk toegepast moeten worden. Wanneer in een eindtoets opgaven van hetzelfde soort voorkomen, kunnen die veel moeilijker blijken te zijn omdat die impliciete context dan is weggefallen.

Wat de te gebruiken vraagvormen betreft, zal het duidelijk zijn dat hier bij uitstek open vragen zoals de opstelvraag en de aanvulvraag passen. De meerkeuzevraag kan soms bruikbaar zijn, bijvoorbeeld wanneer gevraagd wordt welke wet toepasbaar is op een gegeven probleemstelling en de goed geïnformeerde student zo'n vraag beantwoordt door te kiezen uit een kleine set van alternatieven. Dan kunnen die zelfde alternatieven gebruikt worden om er een meerkeuzevraag mee te maken.

Omdat vragen rond wetten en relaties al snel het karakter van problemen krijgen, is er een soepele overgang van deze paragraaf naar hoofdstuk 7 over het stellen van problemen

5.6 Literatuur

MACDONALD-ROSS, M., 'Scientific diagrams and the generation of plausible hypotheses: an essay of the history of ideas, *Instructional Science*, 1979, 8, 233-234.

De rol van wetenschappelijke diagrammen in de structuur van de wetenschappelijke communicatie, hun waarneming en interpretatie door de lezer.

BERTELS, K., & D. NAUTA, *Inleiding tot het modelbegrip*, Bussum, De Haan, 1969.

NAUTA, D., *Logica en model*, Bussum, De Haan, 1970.

Geeft samen met BERTELS & NAUTA (1969) een inleiding op vertaal- en representatieproblemen van de 'werkelijkheid' naar een model voor of theorie over die 'werkelijkheid'.

AITCHISON, J. & I. R. DUNSMORE, *Statistical prediction analysis*, Cambridge, Cambridge University Press, 1975.

GREGSON, R. A. M., *Psychometrics of similarity*, London, Academic Press, 1975.

SNEATH, P. H. A., & R. R. SOKAL, *Numerical taxonomy, the principles and practice of numerical classification*, San Francisco, Freeman, 1973.

JARDINE, N., & R. SIBSON, *Mathematical taxonomy*, London, Wiley, 1971.

MEZZICK, J. E., & H. SOLOMON, *Taxonomy and behavioral science*, London, Academic Press, 1980.

De vijf laatstgenoemde titels betreffen de 'hulpwetenschap' die aan het ontstaan is rond het onderscheiden en classificeren in de wetenschapspraktijk (zowel theorieontwikkeling als praktische toepassingen).

6 Toetsvragen schrijven bij tekst

In sterkere mate dan in het voorafgaande onderwijs kent het hoger onderwijs toetsen die over omvangrijke teksten handelen: 'literatuurtentamens.' De vraagmogelijkheden uit de voorgaande hoofdstukken passen daar niet zonder meer op. Daar komt bij dat juist het literatuurtentamen de docent voor problemen plaatst bij het schrijven van toetsvragen. Een afzonderlijk hoofdstuk dus, waarin de eigenaardigheden en geheel eigen mogelijkheden voor het schrijven van toetsvragen bij tekst centraal staan.

Het is best mogelijk om bij een tekst toetsvragen te schrijven zonder enig voorbereidend werk gedaan te hebben. Verstandig lijkt dat niet, Een leerstofinventarisatie is wel de minste voorbereiding. Leerstofinventarisatie langs de lijnen zoals in hoofdstuk 3 uiteengezet, deelt de stof echter op in afzonderlijke partjes, en daarmee zouden belangrijke bedoelingen van het literatuurtentamen gemist kunnen worden. Daarom biedt dit hoofdstuk enkele aanvullende alternatieven op die leerstofinventarisatie.

Een globale indeling van mogelijke toetsvragen is in vragen over de *letterlijke* tekst, vragen over de *globale* inhoud, *analyse* van de tekst, vragen naar *inferenties* bij de tekst, vragen naar actieve beheersing van de informatie uit de tekst.

1. *Letterlijk*. Kunnen reproduceren van de hele tekst of bepaalde gedeelten (dramatische vakken). Herkennen van uitspraken (en opzettelijk gemutileerde uitspraken ook als zodanig onderkennen); komt heel vaak in toetsvragen voor, terwijl het toch zelden of nooit de bedoeling van de docent is dat passages uit de tekst uit het hoofd geleerd worden. In dit hoofdstuk wordt deze wijze van vragen stellen over tekst verder genegeerd.

2. *Globaal*. De eenvoudigste doelstelling bij een literatuurtentamen is wel dat de student de teksten *gelezen* heeft, niets meer, niets minder. Dan kan de student saillante informatie reproduceren of tenminste herkennen. Een iets actievere kennisname van de opgegeven literatuur leidt tot kennis van de hoofdlijnen, de structuur van een betoog e.d.

3. *Analyse*. Analytische vragen behelzen het thema van een tekstonderdeel, de interne logica ervan, de ordening, de ordening van de geboden informatie en dergelijke.

4. *Inferentie*. Vragen naar conclusies die uit de tekst getrokken kunnen worden. Daar horen ook bij de mogelijke verbindingen van informatie in de tekst met voorkennis, vakkennis of kennis uit het dagelijks leven.

5. *Actief*. De termen en relaties uit de literatuur moeten actief beheerst en toegepast kunnen worden. De student moet bijvoorbeeld nieuwe voorbeelden van termen en/of relaties kunnen herkennen of produceren. (Voor vraagmogelijkheden zie de hoofdstukken 4 en 5). Of informatie uit de literatuur moet gebruikt kunnen worden bij het aanpakken of oplossen van nieuwe problemen (zie hoofdstuk 7).

De categorieën 2, 3 en 4 komen in dit hoofdstuk aan de orde. Het is nu al duidelijk dat er heel uiteenlopende vraagmogelijkheden zijn, en dat toont aan dat het voor een student geenszins vanzelfsprekend is op welk soort vragen hij zich moet voorbereiden wanneer daar geen goede voorlichting over wordt gegeven.

6.1 Participatiecontrole

Het is niet vanzelfsprekend dat de toets tot doel heeft de diepte van bereikte inzichten te peilen. Bij het literatuurtentamen kan het doel ertoe beperkt zijn dat de student bepaalde informatie opneemt en kan reproduceren; toetsvragen horen dan ook tot die reproductie beperkt te blijven. Het kan zelfs gebeuren dat het doel beperkt is tot het simpelweg vaststellen of opgegeven literatuur gelezen is of niet. In dit geval heeft de toets de rol van presentielijst toegewezen gekregen. De analogie is met het practicum waar het er alleen om gaat of de student aan de geplande activiteiten heeft deelgenomen. De toets kan gebruikt worden om op indirecte wijze vast te stellen of de student aan zijn verplichtingen tot het doornemen van de literatuur heeft voldaan en *niet meer dan dat*. de toets als participatiecontrole dus, een idee afkomstig van Hofstee (1973).

De technische uitwerking die Hofstee geeft, is onnodig be-

perkt tot tweekeuzevragen die door bijna iedereen die geparticipeerd heeft goed beantwoord kunnen worden. Keuzevragen kunnen geschikt zijn door hun 'herkenningskarakter', maar dat mag het gebruik van aanvulvragen niet uitsluiten. Gemakkelijke vragen liggen inderdaad voor de hand, maar ook moeilijke vragen, die niet iedere participant goed zal maken, kunnen gebruikt worden voor participatiecontrole. Voorwaarde is dat de vragen duidelijk 'beter' beantwoord worden door participanten dan door niet-participanten; eventueel kan dat met een proefgroep gecontroleerd worden.

Ook toetsing als participatiecontrole vraagt om zorgvuldige voorbereiding: een goede leerstofinventarisatie en toetsvragen geschreven volgens de vraagmogelijkheden die passen bij het doel 'participatiecontrole'. Om met de leerstofinventarisatie te beginnen: deze is te verrichten door het screenen van de opgegeven literatuur op al die informatie die zeker na een keer aandachtig lezen bij de student zal blijven hangen. Doorgaans zal er een geprononceerd onderscheid zijn tussen informatie die *geproduceerd* kan worden, zoals belangrijke thema's, en informatie die mogelijk niet geproduceerd, maar zeker wel *herkend* kan worden, zoals namen en getalsmatige gegevens. Het onderscheid geeft al aan waar aanvulvragen te gebruiken zijn, en waar in ieder geval een beperking tot keuzevragen nodig is. In de leerstof inventarisatie kunnen ook 'vanzelfsprekende' gevolgtrekkingen meegenomen worden. Omdat het in ieder geval om 'saillante' informatie gaat, moet een dergelijke inventarisatie in korte tijd te maken zijn. De vorm die ervoor gekozen kan worden: een lijst (vergeet niet bij ieder item de vindplaats aan te geven, bron met bladzijdennummer) of simpelweg aanstrepen in de tekst. Schematiseren (zie 3.4) kan zeker geen kwaad en is aan te raden wanneer dezelfde onderwerpen op meerdere plaatsen in de opgegeven literatuur ter sprake komen.

VOORBEELD

De 'toets als participatiecontrole' is een gedachte van

- 1) De Groot
- 2) Hofstee
- 3) Van Kemenade.

VOORBEELD

Wat is bedoeld met 'toetsing als participatiecontrole'?

VOORBEELD

Tweekeuzevragen zijn ongeschikt als 'onbenullige toetsvraag.' juist/
onjuist

VOORBEELD

Het laatste voorbeeld is een 'onbenullige toetsvraag.' ja/nee

Onbenullige toetsvragen blijven tamelijk dicht bij de informatie zoals die in de bestudeerde literatuur al gegeven werd. Bedenk dat het ook voor onbenullige toetsvragen mogelijk kan zijn met nieuwe voorbeelden te werken (het laatst gegeven voorbeeld illustreert dat). Wanneer gevraagd wordt nieuwe voorbeelden te produceren, te benoemen of te herkennen moet er wel op gelet worden of van de studenten verwacht wordt dat ze zich op dergelijke vragen actief voorbereiden of niet. In het laatste geval zijn dergelijke vragen alleen te gebruiken wanneer er voldoende aanwijzingen of hints in verwerkt worden om de vraag tot een echt onbenullige vraag te maken: wie de stof gelezen heeft kan de vraag beantwoorden, wie dat niet gedaan heeft blijft het antwoord schuldig (of moet raden). De reden hiervoor is dat er anders strijd met het doorzichtigheidsbeginsel zou ontstaan en dat anders niet uitgesloten kan worden dat de vragen veeleer een beroep doen op intelligentie.

6.2 Thema's en hoofdpunten

Een iets minder passief verwerken van de tekst dan het alleen maar gelezen hebben, is het weer kunnen geven of herkennen van de hoofdpunten, de kern gedachten, de grote lijnen, de belangrijke thema's of verstreckte gegevens. Dat zijn al die onderwerpen die in de tekst zelf al aangeduid worden als belangrijke punten. Wanneer het de bedoeling is dat de literatuur tamelijk *globaal* bestudeerd wordt, kan niet van de student verlangd worden dat hij de literatuur te lijf gaat op een meer actieve wijze zoals in de volgende paragrafen besproken wordt. Geen kwestie van principe, maar van simpele overeenstemming met een expliciet gekozen onderwijsdoel. Dat bete-

kent voor teksten waarbij het niet zonder meer duidelijk is welke de belangrijke onderwerpen zijn, dat de docent een inventarisatie van die onderwerpen maakt en deze lijst aan de studenten beschikbaar stelt (als onderdeel van de stof voor de toets). Advies: ga er niet al te snel vanuit dat studenten deze onderwerpen zelf wel uit de literatuur kunnen selecteren. De student kijkt met heel andere ogen naar de tekst dan de docent met enkele tienduizenden uren ervaring op zijn vakgebied. Wanneer er ook maar enige twijfel is, kies dan de zekere weg en *maak* zo'n onderwerpenlijst. Of doe een experiment, bijvoorbeeld door enkele collega's helemaal onafhankelijk van elkaar de belangrijke onderwerpen op een rijtje te laten zetten. Het is wel zeker dat de onderlinge overeenstemming veel minder dan perfect is en dat sommige lijsten heel wat meer 'belangrijke onderwerpen' bevatten dan andere over dezelfde literatuur. Het helpt de student natuurlijk ook enorm wanneer hij aan een ruime verzameling ter beschikking staande toetsvragen kan zien en door oefening kan ervaren, op welke wijze en wat er over de literatuur gevraagd wordt.

Probeer onder woorden te brengen op welke manier er over deze hoofdpunten gevraagd zal worden. Schrijf de toetsvragen in overeenstemming met deze bedoeling. En geef de beschrijving als een soort instructie ook aan de studenten.

VOORBEELD

De hoofdpunten uit de opgegeven literatuur zijn met een enkel trefwoord aangeduid op bijgaande onderwerpenlijst. De toets zal over deze hoofdpunten vragen bevatten waarbij van u een korte aanduiding in eigen woorden van de essentie van een gedachtengang of theorie e.d. gevraagd kan worden. Daarnaast zullen aanvulvragen en keuzevragen gesteld worden, die te beantwoorden zijn met dezelfde kennis waar mee u anders de vraag naar het aangeven van de essentie goed zou beantwoorden.

Let er bij het maken van de inventarisatie van hoofdpunten op dat daar ook alle belangrijke relaties tussen hoofdpunten *afzonderlijk* in genoemd worden. Het zou niet fair zijn de student tijdens de toetsing te verrassen met vragen die de relatie tussen twee hoofdpunten betreffen, terwijl de hem bekende lijst onderwerpen slechts beide hoofdpunten en niet hun

onderlinge relatie bevat in de volgende voorbeelden is dat de relatie (het kenmerkende verschil) tussen de klassieke en de keynesiaanse economische school. De volgende vragen passen bij een tekstfragment uit Pen en Van Gernerden *Macro-economie* (Aula 620), ook afgedrukt in Mirande (1981, blz. 161 e.v.).

VOORBEELD

Wat is volgens Pen en Van Gernerden de kern van de keynesiaanse economie? (Gebruik niet meer dan ongeveer 50 woorden.)

Antwoord: 'Typerend voor het keynesiaanse denken is dat het nationale inkomen in handen komt van de consumenten, die het uitgeven, en zo komt het inkomen weer terug bij de bedrijven. Er bestaat zoiets als een kringloop van inkomens en besteding die kringloop moet worden geanalyseerd.'

Het antwoord in dit voorbeeld is de letterlijke tekst uit Pen en Van Gernerden; in redelijkheid zal ieder antwoord goedgekeurd worden waarin het bestaan van een kringloop van inkomens en besteding wordt genoemd, en ook daar zijn heel wat uiteenlopende bewoordingen voor te kiezen (zie het volgende voorbeeld). De vraag slaat op een onderwerp dat in de lijst waarvan hierboven sprake was aangeduid kon zijn als 'de kern van de keynesiaanse economie', en niet als 'de keynesiaanse economie' want daar kan van alles en nog wat mee bedoeld zijn, bijvoorbeeld de historische ontwikkeling van deze economische school.

VOORBEELD

Volgens Pen en Van Gernerden is de kern van de keynesiaanse economie de analyse van een bepaalde kringloop. Welke kringloop is dat?

Mogelijke antwoorden:

- van inkomens en besteding
 - van inkomens en besteding van consumenten
 - de geldstroom van bedrijven naar consumenten en van consumenten naar bedrijven
 - van lonen en van hun besteding
- enzovoort.

Het is voor veel studenten een uitstekende geheugensteun wanneer zo'n theoretisch punt van de keynesiaanse economie niet alleen in verbale vorm onthouden wordt, maar wanneer er ook een andere representatie van gemaakt wordt, bijvoorbeeld schematisch. Een stimulans voor het op deze wijze verwerken van de gegeven hoofdpunten is dan het terugvragen van de essentie ervan in de vorm van zo'n schema. In dit geval al heel eenvoudig, een blokje voor respectievelijk bedrijven en consumenten, en twee pijlen met daarbij de aanduiding dat het om geldstromen gaat (inkomens en besteding).

Keuzevragen zijn soms mogelijk, maar daarom nog niet aan te raden.

VOORBEELD

Bedrijven betalen lonen uit die in de vorm van bestedingen weer naar de bedrijven terugvloeien. Van welke economische school is dit de kerngedachte?

- 1) de klassieke
- 2) de keynesiaanse
- 3) de monetaire

Pen en Van Gernerden behandelen slechts drie scholen, dus is het iedere lezer van dat boek duidelijk waartussen gekozen moet worden. De keuzevraag geeft geen extra informatie weg en levert geen raadkans op waar die er in de vorm van een aanvulvraag niet zou zijn. De formulering van de vraag is een parafrase van de formulering zoals die in het boek voorkomt. Gebruik bij voorkeur parafrasen en geen letterlijk uit de opgegeven literatuur overgenomen uitspraken. Dat voorkomt dat studenten kapitaliseren op uit het hoofd leren of op domme herkenning van uitspraken. Het gebruik van parafrasen bevordert een betere verwerking en een beter onthouden van de leerstof (zie o.a. Rickards, 1979). Voorwaarde daartoe is dan wel dat studenten hun stofbeheersing oefenen op dergelijke geparafraseerde vragen.

Een voorbeeld van een vraag over de relatie tussen twee hoofdpunten, waarbij de relatie met zoveel woorden ook in de lijst van hoofdpunten opgenomen hoort te zijn:

VOORBEELD

Wat is het belangrijke verschil tussen de klassieke en de keynesiaanse economie?

Antwoord: 'De keynesiaanse macro economie houdt zich vooral bezig met de factoren die de vraag bepalen. Daar zit het grote verschil met de klassieke theorie die let meer op de produktie, dus op het aanbod.' Of: 'De totale vraag past zich in de klassieke denkwijze aan bij het aanbod (...), maar bij Keynes is het andersom als er onvoldoende vraag is, daalt de produktie.'

De antwoorden zijn passages uit Pen en Van Gernerden. Men kan de student niet verbieden letterlijk te citeren, maar een parafrase is natuurlijk prima. De letterlijke passages zouden als 'modelantwoorden' gebruikt kunnen worden bij het nakijken van de antwoorden. Dat nakijken kan overigens wel lastig zijn, want studenten kunnen met sterk uiteenlopende formuleringen voor de dag komen, die toch als goed aangemerkt kunnen worden. Dit kan voorkomen worden door een verdere specificatie in de vraagformulering op te nemen, of door er meer een aanvulvraag van te maken.

VOORBEELD

Een belangrijk verschil tussen de klassieke en de keynesiaanse economie betreft de relatie tussen vraag en aanbod. Beschrijf dit verschil.

VOORBEELD

De klassieke en de keynesiaanse economie verschillen in opvatting over het verband tussen vraag en aanbod. Bij welke school past de stelling dat de produktie daalt bij onvoldoende vraag?

VARIANT

(zelfde stam)

Bij welke school past de stelling dat de totale vraag zich bij het aanbod aanpast?

VARIANT

De relatie tussen vraag en aanbod wordt wezenlijk anders opgevat door welke twee scholen van economisch denken?

VOORBEELD

Een belangrijk verschil tussen de klassieke en de keynesiaanse economie betreft de veronderstelde oorzaak van werkloosheid. Bij welke school past de gedachte dat werkloosheid het gevolg is van onvoldoende vraag en daardoor gedaalde produktie?

Antwoord: de keynesiaanse.

VARIANT

(zelfde stam)

Bij welke school past de gedachte dat werkloosheid ontstaat doordat van een bepaald goed te veel wordt gemaakt en van een ander te weinig (d.w.z. dat werkloosheid door herscholing van werknemers te verhelpen is?)

Antwoord: de klassieke.

In deze vragen komt nog een ander belangrijk verschil tussen deze twee scholen naar voren. Ook dit verschil moet in de onderwerpenlijst dan wel genoemd zijn.

Merk op dat al deze vragen gesteld zijn op een tamelijk hoog niveau van abstractie. Dat is een nadeel dat hoort bij het alleen maar om globale verwerking van de opgegeven literatuur vragen. Het alternatief is dat de student op een meer actieve wijze met de stof omgaat. Wanneer de student zich oefent op het bedenken van nieuwe voorbeelden bij termen en relaties die in de literatuur voorkomen (of het benoemen en herkennen van nieuwe voorbeelden), dan kunnen toetsvragen geconcretiseerd worden door van nieuwe voorbeelden gebruik te maken. Het nogal abstracte karakter van toetsvragen over globaal bestudeerde literatuur brengt ook met zich dat er maar een beperkt aantal vragen over geschreven kunnen worden. Wie alleen opstelvragen stelt, waarin gevraagd wordt in eigen woorden de kern van een bepaald hoofdpunt weer te geven, beschikt over evenveel vragen als er hoofdpunten in de lijst voorkomen. Dat hoeft natuurlijk niet bezwaarlijk te zijn: de student die al deze hoofdpunten in eigen bewoordingen weer kan geven, beheerst de stof perfect (naar de maatstaf van het gestelde en beperkte onderwijsdoel). Bij aanvulvragen is het mogelijk per hoofdpunt een klein aantal verschillende vragen te formuleren, zoals de gegeven voorbeelden ook laten zien.

6.3 Analyse

Geconcentreerde aandacht voor de hoofdpunten in de literatuur is een eerste stap op de ladder van het meer bewust met teksten omgaan. Een volgende stap is het bewust, expliciet gestructureerd te lijf gaan en verwerken van de opgegeven literatuur. Zolang daarbij alleen gewerkt wordt met de informatie die in de tekst zelf gegeven wordt, kunnen we dat *analyse* noemen. Wanneer informatie uit de tekst verbonden wordt met al aanwezige kennis of wanneer uit een tekst gevolgtrekkingen gemaakt worden die niet door de auteur al werden aangereikt, is dat *inferentie* (zie 6.4). Beide termen moeten worden opgevat als een pragmatische etikettering voor een vat met gevarieerde inhoud. Het analytisch bestuderen van een tekst kan op heel uiteenlopende zaken gericht zijn van de betekenis van gebruikte woorden tot de bedoeling achter een heel boekwerk, van de empirische status van gedane uitspraken tot de logische samenhang in een gegeven argumentatie enz.

De term 'analyse' suggereert een verwantschap met de categorie 'analyse' in de cognitieve taxonomie van onderwijsdoelen van Bloom c.s. (1956). Die verwantschap is zeker wel aanwezig, en daar zal ik ook gebruik van maken. Maar er is een heel wezenlijk verschil in de benadering van Bloom c.s. en die welke in deze paragraaf uitgewerkt zal worden. Bloom c.s. beschouwen het analytisch omgaan met tekst als een vaardigheid op zich, te vergelijken met schriftelijke uitdrukkingsvaardigheid. Voor middelbaar en hoger onderwijs is het een belangrijk doel de student deze vaardigheid bij te brengen. Bij Bloom c.s. zijn analytische toetsvragen dan ook bedoeld om deze vaardigheid zelf te toetsen. Daarbij is de standaard gang van zaken dat de student een nieuw stuk tekst krijgt voorgelegd met een aantal vragen die via analyse van de gegeven tekst te beantwoorden zijn. Een belangrijke Amerikaanse gestandaardiseerde toets als de New High School Equivalency Examination bestaat voor een aanzienlijk deel uit precies dergelijke analytische vragen, ook voor vakken waarvan je zou denken dat het daarbij toch allereerst om kennis en inzicht in inhoudelijke onderwerpen zou gaan. Welnu, voor dergelijke analytische vragen zal ik geen vraagemogelijkheden geven want dat zijn vragen die niet algemeen te gebruiken zijn, maar

slechts voor het studieonderdeel dat in het bijzonder gericht is op het bijbrengen van dergelijke analytische vaardigheden in algemene zin.

Hoe zijn analytische vragen dan wel te gebruiken? Op welke manier is het analytisch omgaan met teksten dan wel zinvol? Vergelijk analyseren weer met die andere algemene vaardigheid: schrijven. Ook zonder dat schrijven een expliciet doel van het onderwijs is, speelt het een belangrijke rol omdat de student zich schriftelijk moet kunnen uitdrukken. Op dezelfde wijze is analyseren van belang: als hulpmiddel of techniek om je een tekst op bepaalde wijze 'eigen' te maken. Iemand die zich een tekst door analyse eigen heeft gemaakt, heeft een ander, dieper inzicht in het in die tekst behandelde onderwerp dan iemand die slechts op globale wijze van diezelfde tekst kennis heeft genomen. Daar ligt dan ook de aansluiting met de te schrijven toetsvragen: het is nu mogelijk om de toetsvragen te richten op de 'resultaten' van het analytisch bestudeerd hebben van een tekst. Merk op dat dit iets anders is dan het schrijven van toetsvragen die tekstanalyse vragen uit te voeren op een gegeven (nieuw) stuk tekst, ook al zullen deze vragen naar de vorm veel gemeenschappelijks kunnen vertonen met de eerstgenoemde.

Dat klinkt veelbelovend, maar hoe voorkom je dat toetsvragen die gericht zijn op het *analytisch verwerkt* hebben van de literatuur niet veeleer dat *analytisch vermogen* van studenten toetsen dan de kennis en het inzicht die via analyse verkregen zijn? Dit is een probleem, inderdaad, maar het is niet ernstiger dan het vergelijkbare probleem dat zich voordoet bij toetsen waarbij het mede aankomt op het vermogen van de student zich schriftelijk uit te drukken. Of meer algemeen: bij de voorbereiding op een tentamen speelt het op intelligente wijze omgaan met de leerstof altijd een rol, maar dat wil nog niet zeggen dat toetsresultaten eerder een weerspiegeling van verschillen in intelligentie dan van verschillen in stofbeheersing zouden zijn. Dat neemt echter niet weg dat het problematisch kan worden wanneer de vragenschrijver zich van de mogelijkheid van verwarring van beide niet bewust is. Wie vragen stelt die alleen maar te beantwoorden zijn door studenten die analytische hoogstandjes kunnen maken, gaat ver over de schreef. De analytische vaardigheden die bij de student veron-

dersteld worden, moeten redelijk zijn, liever nog bescheiden verondersteld worden. Aan de didactische kant kan de docent er ook veel aan doen door glashelder te maken op welke wijze de student analytisch met de literatuur om kan gaan of verwacht wordt om te gaan. Dat kan gebeuren door een aantal modelmatige voorbeelden te geven en door te laten oefenen op toetsvragen die op die analyse geënt zijn (en daarbij uit te leggen op welke wijze de toetsvragen ingaan op analytisch 'verwerkte' informatie). Daarbij kan een lijst van analytische vragen gebruikt worden, samen te stellen uit de lijst die ik verderop in deze paragraaf zal geven.

Analytisch lezen is 'lezen wat er staat', zou je kunnen zeggen. En dat is niet altijd even eenvoudig.

VOORBEELD

'Als Nederland het besluit over het stationeren van kruisraketten uitstelt of afwijst, kan dit het succes van de komende wapenonderhandelingen met de Sovjet-Unie in gevaar brengen.'

Het analyseren van deze uitspraak kan aanleiding vormen tot het schrijven van een uitgebreid opstel. Het thema is de relatie tussen een Nederlands besluit en ontwapeningsonderhandelingen. Er wordt een oorzakelijk verband tussen beide gelegd. In dit geval is dat een kansmatig (stochastisch) verband: het staat niet vast dat er zo'n oorzakelijk verband is, maar de mogelijkheid ertoe bestaat. Het is een verband dat niet empirisch aantoonbaar is, ook niet wanneer Nederland stationering afwijst en de onderhandelingen mislukken. Dat bestempelt de uitspraak tot een uitspraak met een suggestief karakter. (Hij is van Amerikaanse functionarissen, NATO-hoofdkwartier Brussel, augustus 1981). Een wazige term in deze uitspraak is 'succes': de betekenis daarvan is niet duidelijk, terwijl de betekenis van de uitspraak in zijn geheel er toch van afhangt. Enzovoort, enzovoort.

Analytische vragen zijn bijvoorbeeld: Wat is het thema van deze paragraaf?; Welke argumenten voor of tegen worden aangevoerd?; Worden er voorbeelden gegeven?; Welke conclusies worden getrokken?

Analytische vragen zijn op verschillende manieren in te delen. Bij het schrijven van toetsvragen is het gebruik van een indeling in 'bewust' en 'kritisch' handig: er kan gevraagd worden of de student zich ervan bewust is dat een gelegd verband oorzakelijk van karakter is, terwijl kritische vragen bijvoorbeeld de empirische onderbouwing voor een verondersteld oorzakelijk verband betreffen.

Voor het schrijven van toetsvragen kan begonnen worden met analytische vragen te bedenken ervan uitgaande dat studenten bij de toets de beschikking over de betreffende tekst hebben. Is het niet de bedoeling dat studenten bij de toets de bijbehorende tekst als gegeven krijgen, dan kunnen de vragen later redactioneel geherformuleerd worden.

In de volgende lijst met mogelijke analytische vragen, die als niet meer dan suggestief en zeker niet als volledig moet worden gezien, is een ordening aangebracht van vragen betreffende woorden en termen, via vragen over relaties, naar vragen over tekstgedeelten.

Woorden

- Heeft dit woord een gunstige of ongunstige connotatie?
- Is het woordgebruik op meerdere plaatsen suggestief? (geef voorbeelden)

Termen

- Welke (nieuwe) termen worden geïntroduceerd?
- Is deze term gedefinieerd (omschreven, van duidelijke voorbeelden voorzien)?
- Wanneer later in een tekst concrete voorbeelden van een eerder ingevoerde term (of relatie) voorkomen, vraag dan dergelijke voorbeelden ook expliciet te benoemen (wanneer de auteur dat niet gedaan heeft).
- Maken gegeven voorbeelden voldoende duidelijk wat de betekenis van de betreffende term is? Zo nee, op welke punten blijven er onduidelijkheden? Kun je zelf nieuwe voorbeelden geven, waarbij het onduidelijk is of deze term erop van toepassing is of niet?

Relaties

- Welke relatie wordt genoemd (geïntroduceerd, behandeld)?

- Wat is het karakter van deze relatie? (oorzakelijk, volgorde, logisch, overeenkomst, verschil enz.; zie inleiding hoofdstuk 5).

- Is de relatie gedefinieerd (omschreven, van duidelijke voorbeelden voorzien)?

- Maken de gegeven voorbeelden de betekenis van de relatie voldoende duidelijk?

Zo nee, wat zijn de onduidelijkheden?

- Is de relatie empirisch van aard (is de juistheid ervan aan concrete gegevens te toetsen)?

Wordt een geponeerde relatie met empirische gegevens (resp. een theoretische of logische argumentatie) onderbouwd? (zie ook onder *Argumentatie* hieronder). Is de relatie in objectieve bewoordingen beschreven of in suggestieve?

Tekstgedeelte (uitspraak, alinea, paragraaf, hoofdstuk, boek)

- Wat is het thema (hoofdpunt, belangrijkste idee)?

- Hoe wordt het thema uitgewerkt? (voorbeelden, argumenten voor en tegen, historische ontwikkeling, bekende onderzoeksresultaten, verdere detaillering, relatie tot andere thema's, toepassingen enz.). Is de uitwerking van het thema adequaat (wordt er afgedwaald, verliest de uitwerking zich in details, wordt een empirisch verifieerbare stelling alleen logisch of verbaal uitgewerkt)?

- Werk de onderlinge relaties tussen de verschillende thema's uit (formuleer welke die relaties zijn, maak een schema van de onderlinge verbanden). Is de tekst volledig in de behandeling van onderlinge relaties tussen de verschillende thema's (is het geheel een samenhangende uiteenzetting, of zitten er leemten in; zijn er thema's die in het geheel niet aan de overige besproken thema's gerelateerd zijn)?

Argumentatie

- Is de uitspraak (stelling, mededeling) feitelijk van aard, is de juistheid empirisch te toetsen, is de juistheid te demonstreren op theoretische of logische gronden, is de uitspraak een mening of veronderstelling?

- Welke argumenten worden aangevoerd (voor zowel als te-

gen, empirische zowel als theoretische)?

- Is de argumentatie eenzijdig (overwicht van argumenten die voor de stelling pleiten, suggestieve formulering van argumenten)?
- Is de aard van de aangevoerde argumenten feitelijk of zijn het meningen?
- Worden feitelijke argumenten met voldoende gegevens onderbouwd?
- Wat is de bron van de verstrekte gegevens?
- Zijn gegeven argumenten relevant met betrekking tot het gestelde?
- Wat is de structuur van de argumentatie (wat is de onderlinge relatie van de verschillende gegeven argumenten)?
- Is de logica van de argumentatie te onderkennen? Zo ja, geef de logische structuur aan. Is deze logica 'correct'?
- Is de argumentatie inductief van karakter? Maakt de auteur duidelijk wat de beperkte waarde van zijn of haar inductieve logica is? Zijn de conclusies van de auteur in overeenstemming met het inductieve karakter van de argumentatie? Wordt er te sterk gegeneraliseerd?

In iedere discipline zijn er wel een aantal analytische vragen die voor dat vak van bijzonder belang zijn en die de student bij het doorwerken van de literatuur routinematig moet kunnen stellen en beantwoorden. Het is niet gezegd dat de analytische vragen die in uw vak van grote betekenis zijn ook in bovenstaande lijst voorkomen.

Het is aan te raden uit bovenstaande uitgebreide lijst van mogelijkheden een klein maar select aantal vraagmogelijkheden voor gebruik in de toets te kiezen. Zorg ervoor dat studenten hun literatuur dan ook met analytische vragen in het achterhoofd bestuderen. Geef ze het lijstje met de te verwachten vragen van analytische aard. Geef onderwijs in het analytisch omgaan met tekst (door daar zelf in uw college een 'model' van te geven, door in werkgroepen aandacht te besteden aan deze tekstanalyse). Wanneer studenten over een set analytische vragen beschikken zoals die ook op het tentamen te verwachten zijn, bevordert dat het analytisch lezen, met de daaruit resulterende betere verwerking van de stof.

Analytische toetsvragen hebben enigszins het karakter van reproductievragen, zij het dat nu niet gevraagd wordt tekst of informatie letterlijk te reproduceren. De waarde van deze analytische vragen lijkt dan ook beperkt, maar dat is grotendeels schijn. Allereerst heeft het gebruik van analytische vragen een 'feed-forward' -effect op de wijze waarop de student de literatuur verwerkt en dat kan op zich al rechtvaardiging genoeg zijn. Maar er komt bij dat voor een literatuurtentamen soms niet veel meer verwacht mag worden dan het analytisch verwerken van de gegeven tekst, welnu dan kan men daar ook tevreden mee zijn. Tenslotte is er nog de mogelijkheid analytische vragen over de literatuur te combineren met vragen die inferenties betreffen; zie de volgende paragraaf.

6.4 Inferentie

Met de term 'inferentie' worden alle vormen van interpretatie aangeduid waarbij de letterlijke tekst 'verrijkt' wordt met duidingen, verder reikende gevolgen, verbanden met andere beschikbare informatie en dergelijke. Het is duidelijk dat het maken van inferenties veelal een gegeven tekst veronderstelt. En ook dat het analyseren van tekst het maken van inferenties vergemakkelijkt en bevordert. Het verschil tussen analyse en inferentie is overigens van tamelijk triviale aard: je zou kunnen zeggen dat inferentie hetzelfde is als analyse, maar dan op de tekst *samen* met andere gegevens waarover we kunnen beschikken. Dat kan zowel voorkennis van inhoudelijk karakter zijn als de uitgebreide kennis-van-de-wereld die ieder als taalgebruiker tot zijn beschikking heeft. Analyse: Is de uitspraak empirisch onderbouwd; wat is de relatie tussen de gegevens en de conclusie; volgt de conclusie op geldige wijze uit de gegevens enz. Inferentie: Welke gegevens heb je nodig om de gegeven uitspraak waar te maken; zijn er uit de verstrekte gegevens ook ander conclusies te trekken; hoe verhoudt zich de gegeven informatie met andere kennis die ik heb en die relevant is enz.

Dat betekent dat het materiaal dat in 6.3 gepresenteerd is hier uitdrukkelijk bekend wordt verondersteld. Inferentie bouwt in sterke mate verder op analyse. Het betekent ook dat inferentie hier een behandeling krijgt die parallel is aan die van analyse

in 6.3. Het maken van inferenties bij het lezen van tekst wordt als een min of meer vanzelfsprekende vaardigheid beschouwd, zoals het kunnen schrijven en analyseren vanzelfsprekende vaardigheden zijn (waaraan geen hoge eisen gesteld worden). Hoewel aan vaardigheid in het maken van inferenties ook in het onderwijs wel enige (zijdelingse) aandacht besteed zal moeten worden, hebben de onderwijsdoelen toch niet (in de eerste plaats) betrekking op het bijbrengen van deze vaardigheid, en heeft de toetsing *niet* als bedoeling deze *vaardigheid* vast te stellen. Vergelijk wat in de voorgaande paragraaf over het kunnen analyseren is opgemerkt.

Inferenties zijn te verdelen in 'taalkundige' inferenties en 'inhoudelijke' inferenties. De taalkundige inferenties zijn tamelijk vanzelfsprekend, vinden soms zelfs als automatisme bij het lezen al plaats. De inhoudelijke inferenties zijn van bijzonder belang voor het opnemen en onthouden van de informatie die in de tekst wordt aangeboden. Het gaat immers om de integratie van de nieuwe informatie met de al beschikbare kennis. Ook deze inhoudelijke inferenties worden vaak op vanzelfsprekende wijze al tijdens het lezen gemaakt, en dat gebeurt in sterkere mate bij het *herlezen* van de tekst dan bij de eerste kennismaking.

'Het ziet ernaar uit dat bij eerste lezing een tekst op een tamelijk oppervlakkig semantisch niveau wordt verwerkt, terwijl veel studenten bij het herlezen de informatie in verband brengen met een ruimere context. Dit is een ervaring die we allen wel kennen. Een tweede keer een gedicht of roman lezen, of een film zien, leidt vaak tot een nieuw begrip, waaronder ook een doordringen tot de bredere betekenis van het werk' (Friedman en Rickards, 1981).

Het belang van inferentiële vragen, evenals dat van analytische, is tweeledig: het gebruik van dergelijke vragen heeft een 'feed-forward'-werking op het studiegedrag, waardoor de literatuur op een meer geïntegreerde wijze verwerkt zal worden dan anders misschien het geval zou zijn, terwijl er daarnaast in het karakter van de aangeboden stof een aanleiding kan zijn om een bepaalde nadruk op inferentie te leggen. Bij de meeste literatuurtentamens zal toch wel gelden dat er in de opgegeven literatuur bepaald samenhangen te onderkennen zijn, een zekere onderlinge verbondenheid van de informatie, die de stu-

dent er als zodanig ook uit moet kunnen halen. De student moet zich de instelling eigen maken om nieuw aangeboden informatie zelf in verband te brengen met kennis die hij of zij al heeft. De student kan daarmee geholpen worden wanneer hij voorgelicht wordt over en enige oefening krijgt in de aard van de interferenties waar het in deze discipline in het bijzonder om gaat, Tenslotte is het niet overal zo geregeld als in de wiskunde en in de geformaliseerde wetenschappen, waarbij studenten via het maken van talrijke opgaven in de gelegenheid gesteld worden nieuwe informatie heel expliciet te relateren aan de overige informatie waarover ze al beschikken. Het maken van inhoudelijke inferenties zal de student in het begin niet al te gemakkelijk afaan. Het kenmerk van onderwijs is immers dat er voortdurend informatie wordt aangeboden, die in sterke mate nieuw is voor de student en waarbij het misschien niet direct voor de hand ligt hoe deze informatie te relateren is aan wat overigens al tot het repertoire van kennis en vaardigheden van de student behoort. Dat is een geheel andere situatie dan die waarin de docent zich bevindt wanneer deze nieuwe informatie op het eigen vakgebied verwerkt: de relevante dwarsverbindingen worden snel en bijna vanzelfsprekend gelegd. De docent beschikt ook al over een uitgebreid en stevig geconsolideerd netwerk van informatie op het eigen vakgebied en juist bij zo'n netwerk is het gemakkelijk er nieuwe informatie aan vast te knopen. Van de student wordt daarentegen juist verwacht dat hij zo'n goed verbonden netwerk van kennis begint *op te bouwen*. Daar volgt dan ook uit dat de student ondersteund moet worden bij het maken van inferenties, daartoe gestimuleerd moet worden, en dat de docent bescheiden moet zijn in de eisen die hij op dit punt aan de studenten denkt te kunnen stellen, (Hulpmiddelen voor de student zijn: het meerdere keren doornemen van de stof met het doel deze meer geïntegreerd te verwerken, het schematiseren van de stof, het hanteren van een kleine lijst van inferentiële vragen die 'standaard' te stellen zijn.)

Voor het schrijven van toetsvragen kan er om te beginnen van uitgegaan worden dat de studenten de tekst ook *gegeven* krijgen. Is dat in feite niet het geval, dan kan in een later stadium de redactie van de geschreven vragen aan die omstandigheid worden aangepast. Ik geef hier weer een suggestieve lijst met

vraagmogelijkheden, algemeen geformuleerd:

Woorden

- Waarom gebruikt de auteur hier een suggestief woord? (wat wil de auteur ermee bereiken, en waarom; had hij een ander woord kunnen gebruiken; duidt het gebruik van dit woord op een bepaalde opvatting, een verborgen veronderstelling)
- Wekt het gebruik van dit woord bepaalde verwachtingen? Welke? (het gaat hier om meerdere betekenissen die een woord kan hebben, om sterke associatieve relaties die met andere woorden kunnen bestaan, zoals kernenergie-milieu-stralingsrisico-Harrisburg)

Termen

- Is deze term al eens eerder gebruikt? Zijn er voorbeelden van gegeven, en welke waren die voorbeelden?
- Kan men nieuwe voorbeelden bij deze term bedenken?
- Met welke andere termen is deze term gerelateerd? En van welke aard zijn die relaties? Zijn daar voorbeelden van gegeven? Enz.
- (waarneembare term) Waar of wanneer kan ik zo'n ding, dier of gebeurtenis werkelijk tegenkomen, verwachten?
- (indirect waarneembare term) Welk instrument of welke onderzoekopstelling is nodig om dit verschijnsel waar te nemen, waarneembaar te maken? Op welke veronderstellingen berust het concluderen vanuit hetgeen waargenomen is tot de aanwezigheid van het bedoelde verschijnsel?
- (abstracte term) Welke aspecten of instanties die met deze term aangeduid worden, mogen verwacht worden in de context van deze uiteenzetting speciaal van belang te zijn?
- (construct) Is het noodzakelijk dit construct te hanteren? Is een ander construct mogelijk? Zouden een ander construct, een andere assumptie, tot andere resultaten leiden?
- (theoretische term) Waarom zijn de relaties van deze term tot andere termen zoals hier uiteengezet en niet anders? Zijn er andere mogelijkheden? Is deze term ook te relateren aan hier niet genoemde (theoretische, of waarneembare) termen?

Bovenstaande vraagmogelijkheden hebben betrekking op termen die in de tekst genoemd zijn. Het kan natuurlijk ook zijn dat de relevante term niet genoemd wordt, maar wel een voorbeeld, instantie, aanduiding, of beschrijving gegeven wordt:

Welke term wordt voor dit voorbeeld gebruikt?

Is het beschreven verschijnsel zo belangrijk dat er een speciale term voor gebruikt wordt (die mogelijk later in de tekst geïntroduceerd wordt)?

Relaties.

Dezelfde vraagmogelijkheden als hierboven voor 'termen' genoemd. Daar werd ook al gewezen op de vraag naar andere relaties dan in de tekst genoemd, van genoemde termen tot niet in de tekst genoemde termen.

- Kan het genoemde verband doorgetrokken of gegeneraliseerd worden? Voer dat dan ook uit, of geef aan waarom er niet gegeneraliseerd mag worden.

- Stellingen, opvattingen, weergegeven gebeurtenissen: welke gegevens of welke theoretische of logische argumenten zijn nodig om de stelling enz. te onderbouwen, te bewijzen, aannemelijk te maken? Welke veronderstellingen zijn impliciet in de stelling enz.?

Argumentatie

- Wat is de theoretische achtergrond van de gegeven argumentatie (tot welke 'school' behoort de auteur) en valt de gegeven argumentatie vanuit deze veronderstelde theoretische achtergrond verder aan te vullen? (een hele theorie als impliciete veronderstelling). Zijn de gehanteerde (dan wel de impliciete) veronderstellingen dwingend van karakter, of hadden ook andere veronderstellingen gemaakt kunnen worden?

- Hoe zou de argumentatie er onder andere mogelijke veronderstellingen uitzien en tot welke eventueel andere conclusies zou ze leiden?

- Wat is de waarde van argumentatie en conclusies wanneer veronderstelde empirische gegevens niet beschikbaar zouden blijken of anders uit zouden vallen dan verondersteld of voorspeld?

- Hoe zou deze argumentatie eruit kunnen zien wanneer ze gegeven was door een auteur uit school Y in plaats van X?

Wanneer het (routinematig) maken van bepaalde inferenties in uw vak van belang is en in de onderwijsdoelen gespecificeerd is (of zou moeten worden), maak dan ten behoeve van

uw onderwijs en als hulpmiddel voor de studenten een lijstje van deze belangrijke inferenties.

Zeer veel inferenties liggen dermate voor de hand dat de goede lezer, de student die voor de tweede maal zijn tekst bestudeert, er niet aan ontkomt ze spontaan te maken. Dat betreft ook dwarsverbindingen door de leerstof. Toetsvragen die op dergelijke inferenties afgaan, zijn altijd te gebruiken. Daar zit wel enig risico in, omdat nog maar moet blijken of de gevraagde inferentie inderdaad van dit 'spontane' karakter is. Let daar bij de controle op de kwaliteit van de toetsvragen (zie hoofdstuk 8) dan ook terdege op.

6.5 Literatuur

HOFSTEE, W. K. B., 'Participatie controle door "onbenullige" toetsitems', *Nederlands tijdschrift voor de Psychologie*, 1973, 28, 189-198. Ook in: A. G. VROON, & S. E. M. EVERWIJN (red.), *Handboek voor de onderwijspraktijk*, Deventer, Van Loghum Slaterus (losbladig).

NOORDMAN, L. G. M. & W. VONK, 'Lezen: het begrijpen van tekst', *Nederlands tijdschrift voor de Psychologie*, 1981, 36, 385-408.

Een overzicht van onderwijspsychologisch onderzoek dat de laatste jaren op dit gebied tot ontwikkeling is gebracht.

GAGNÉ, E. D., 'Long term retention of information following learning from prose', *Review of Educational Research*, 1978, 48, 629-665.

REDER, L. M., 'The role of elaboration in the comprehension and retention of prose: a critical review', *Review of Educational Research*, 1980, 50, 5-53

ODELL, L., 'Teaching writing by teaching the process of discovery: an interdisciplinary enterprise', in: L. W. GREGG & E. R. STEINBERG, *Cognitive processes in writing*, Hillsdale, New Jersey, Erlbaum, 1980.

Het maken van inferenties als opwarmer voor het schrijven ('tagmemic theory').

7 Problemen stellen

Er is een categorie vragen die niet goed in een van de voorgaande hoofdstukken is in te passen: dat zijn de vragen waarbij de student enige vrijheid wordt gelaten in de wijze van beantwoorden, waar bij voorkeur eerst een oplossingsplan wordt opgesteld voordat er gerekend enz. gaat worden en waar probleemrelevante informatie op eigen initiatief ingebracht moet worden. Het zijn vragen waar je eerst eens over moet nadenken, in tegenstelling tot vragen waarop je het antwoord weet of kunt zoeken. De toetsvragen in de voorgaande hoofdstukken hebben met elkaar gemeen dat ze gericht bepaalde informatie (terug)vragen; ze toetsen niet op het vermogen om diezelfde informatie spontaan, ongevraagd en zelfstandig te kunnen produceren.

Problemen doen dat in meerdere of mindere mate wel: de student moet nu zelf bedenken welke informatie bruikbaar zou kunnen zijn om het gestelde probleem dichterbij een oplossing te brengen.

In het dagelijks taalgebruik is een probleem een moeilijk vraagstuk. In dit hoofdstuk geef ik een meer specifieke betekenis aan deze term (als stipulatieve definitie, zou je kunnen zeggen): Onder een *probleem* wordt hier verstaan een vraagstuk dat voor zijn oplossing een *planmatige* aanpak van de student vraagt, waarbij de student *zelfstandig* probleemrelevante informatie moet kunnen inbrengen. Dit laat uitdrukkelijk de mogelijkheid open dat problemen wel eens makkelijk kunnen zijn of routinematig opgelost kunnen worden.

Het hoofddoel van dit hoofdstuk is het geven van vuistregels voor het bedenken, construeren en stellen van problemen; die vuistregels komen in 7.3 aan de orde. Evenals bij het schrijven van 'gewone' toetsvragen is er voorafgaand aan het stellen van problemen een leerstofinventarisatie (7.2) te maken, die het materiaal oplevert waarop de vuistregels losgelaten kunnen worden. Voor zowel de vuistregels als voor de inventarisatie is het noodzakelijk om te beschikken over een *taal* om op vruchtbare wijze te kunnen spreken over de diverse activi-

teiten waarin het aanpakken en oplossen van problemen uiteengelegd kan worden. Vandaar dat in 7.1 een analyse wordt gegeven van wat er zoal bij het aanpakken en oplossen van problemen komt kijken; tegelijkertijd wordt dan een (zo spaarzaam mogelijke) terminologie ingevoerd.

Het kan buitengewoon moeilijk zijn om in didactisch en ander opzicht 'greep' te krijgen op dat 'oplossen van problemen'. Voor een aantal vakgebieden is daar al een hoopgevend begin mee gemaakt, zie de verwijzingen in de komende paragrafen; dezelfde literatuur, samen met de sterk vereenvoudigde aanwijzingen die in dit hoofdstuk gegeven worden, kunnen de docent in andere vakgebieden helpen daar tenminste een schuchter begin mee te maken.

7.1 Over problemen gesproken

De vuistregels voor het stellen van problemen, zoals die in 7.3 gegeven worden, sluiten aan op de verschillende 'stappen' waarin het aanpakken en oplossen van problemen uiteengelegd kan worden. In fig. 7.1 is schematisch aangegeven welke stappen dat zijn, en hoe deze onderling gerelateerd zijn. Bij de meeste problemen laat deze cyclus van stappen zich in en of andere vorm in de oplossing herkennen. En dat is dan ook de bedoeling bij het presenteren van dit schema: het biedt een aantal vaste aanknopingspunten om over een gegeven probleem en zijn oplossing te kunnen spreken en ook om meer in het algemeen uiteen te kunnen zetten om welke specifieke

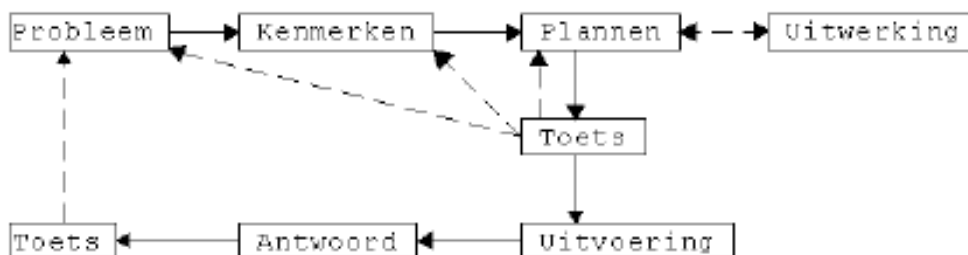


fig. 7.1 De probleemaanpak in schema

vaardigheden het gaat bij het oplossen van problemen in het eigen vakgebied. Op dat laatste sluiten de inventarisatie (in 7.2) en daarop weer de vuistregels (in 7.3) aan. De gekozen terminologie krijgt hieronder een verdere toelichting, waarvan de algemene strekking is dat deze termen een zeer brede betekenis krijgen toebedeeld. Let daarbij vooral op de ongewone invulling van wat onder een 'plan' verstaan kan worden (daar wordt bijvoorbeeld ook een hypothetisch antwoord onder begrepen).

Misschien belangrijker dan de terminologie zijn de relaties zoals in fig. 7.1 met pijlen aangeduid: het gaat er in het onderwijs niet zozeer om dat de student beseft dat hij een bepaalde stap moet maken als wel dat hij in staat is om die stap te maken. Vooral de stap van expliciet gemaakte *kenmerken* van de probleemsituatie naar mogelijke *plannen* voor de oplossing is heel wezenlijk bij het oplossen van problemen. Tenslotte wil ik benadrukken dat dit schema niet dogmatisch bedoeld is: in de praktijk zullen de verschillende onderscheiden stappen min of meer in elkaar kunnen overvloeien en in sterke onderlinge wisselwerking kunnen staan. Voor een bepaalde klasse van problemen in een specifiek vakgebied zal het soms mogelijk zijn om dit globale schema te vervangen door een meer gedetailleerd schema met meerdere expliciet aangegeven terugkoppelingsmomenten.

Problemen. De problemen waarmee in het hoger onderwijs gewerkt wordt zijn van een grote verscheidenheid. Dat kan ook binnen ieder afzonderlijk vakgebied gelden: analyseer dan ieder type of iedere klasse van problemen afzonderlijk.

Natuurlijk zijn er heel algemene typen van problemen te onderscheiden, bijvoorbeeld 'gesloten' problemen waarbij een heel specifiek antwoord verlangd wordt (bewijzen en berekenen) en 'open' problemen waarbij het antwoord grote individuele vrijheid toelaat (essay, constructieproblemen). Dergelijke typering is op talrijke manieren te maken en hoewel het probleemtype zeker van invloed is op de typische aanpak van het probleem, zie ik in dit hoofdstuk ervan af om expliciet in te gaan op dergelijke verschillen en benadruk ik liever wat er in grote lijnen gemeenschappelijk is in de aanpak van heel verschillende probleemtypen.

In het volgende zal ik bij voorkeur de aanpak van schaakproblemen als voorbeeld nemen. Niet omdat er nogal wat denkpsychologisch onderzoek op dit gebied is gedaan (bijv. De Groot, 1978), maar gewoon omdat deze problemen voor de meeste lezers inzichtelijk zijn en omdat de afzonderlijke punten uit fig. 7.1 er goed in uitkomen.

Kenmerken. Een aangeboden probleem wordt allereerst zorgvuldig bestudeerd: de probleemstelling wordt geanalyseerd (zie 6.3), er worden mogelijk al enkele inferenties gemaakt (zie 6.4), het probleem wordt in een andere, meer toegankelijke vorm gebracht (vertaald, getekend, in formule gebracht; zie 5.1). Deze activiteiten zijn erop gericht die kenmerken van de probleemsituatie naar voren te brengen waarop met plannen aangesloten kan worden. Bij het schaken gaat het om het vinden van de belangrijkste kenmerken van de stelling; bij juridische problemen (casus) om het aanwijzen van het centrale probleem. Het identificeren van belangrijke kenmerken in een probleemsituatie is een (eventueel afzonderlijk) te onderwijzen, te oefenen en te toetsen vaardigheid (zie 5.1, 6.3 en 6.4). Of de zo gevonden kenmerken een oplossing mogelijk maken, zal echter nog moeten blijken; het is altijd mogelijk dat in een later stadium deze analyse uitgebreid of overgedaan moet worden.

SCHAKEN

Relevante kenmerken van een schaakstelling kunnen bijvoorbeeld zijn: de open of halfopen verticale lijn, de goede en de slechte loper, kenmerken betreffende de pionnenstelling of de bouw van de koningsvleugel. Het zijn die kenmerken waarvan de ervaring heeft uitgewezen dat ze van strategisch belang zijn; de theorie is de systematische neerslag van deze ervaring. Euwe (1966): 'In het algemeen zal niet één kenmerk de situatie beheersen; er zullen meer kenmerken aanwezig zijn, waarvan nu eens het ene en dan weer het andere overweegt. Wij moeten dus álle belangrijke kenmerken der stelling opsporen en ook hun onderlinge waarde taxeren alvorens een plan te gaan vormen.' En: 'De opsporing van het belangrijkste of beter gezegd het "leidende kenmerk" is derhalve de eerste taak van de speler.'

VOORBEELD (juridisch casus)

'Persoon A leent een kostbaar boek aan persoon B. Door een misverstand verkeert B in de mening dat A hem het boek ten geschenke heeft gegeven. Enige tijd later geeft B het boek op zijn beurt als verjaardagsgeschenk aan C, A bezoekt C en treft daar zijn boek aan. A vraagt C hem het boek terug te geven. C weigert, omdat hij naar zijn zeggen het boek gekregen heeft van een betrouwbare vriend, die een boek dat niet van hem was niet zou weggeven.

A wil nu zijn boek in rechte terugvorderen.' (Crombag c.s., 1972).

Het centrale probleem, dat wat het casus kenmerkt, is dat A zijn boek terug wil hebben van C.

VOORBEELD (wiskundige analyse)

Evalueer de onbepaalde integraal: $y = \int 1 / (0,5 + \sqrt{x}) dx$

Leidend kenmerk van dit probleem is $x^{1/2}$, of $1 + x^{1/2}$, het aangrijpingspunt is om de integraal met behulp van de substitutiemethode op te lossen.

Het laatste voorbeeld illustreert nog eens dat 'problemen' niet altijd hoeven te verschillen van bepaalde soorten opgaven die in voorgaande hoofdstukken behandeld zijn. De kern van dit probleem is immers de keuze van de juiste techniek (hier de substitutiemethode) uit een klein aantal mogelijk bruikbare technieken.

Hoewel het in dit boek niet over didactiek gaat, is deze opmerking van Euwe hier toch wel interessant: 'Hij, die de partij volgens vaste regels behandelt en zich voortdurend rekenschap geeft van de aanwezige kenmerken, zal zijn spel veel gemakkelijker kunnen perfectioneren dan iemand die zuiver op gevoel of berekening afgaat en nooit vraagt naar het waarom.'

De opgespoorde kenmerken in de probleemsituatie dienen als springplank voor het opstellen van plannen, zodat de betekenis van deze kenmerken pas duidelijk wordt in samenhang tot die plannen. Over opsporen gesproken: dat is precies wat de student *expliciet* moet doen en waar de ervaren probleemop-

losser misschien in het geheel niet bewust bij stilstaat. Hetzelfde verschil tussen ervaren en onervaren probleemoplossers zullen we hierna bij het toetsen van opgeworpen plannen tegenkomen.

Plannen. Hoewel de andere stappen uit het aanpakschema niet gemist kunnen worden, zit de crux van het oplossen van een probleem toch wel in het weten op te stellen van plannen. Omdat we niet kunnen veronderstellen dat de probleemoplosser ergens tussen zijn grijze hersencellen een klein mannetje heeft zitten dat via de zintuigen aangereikte problemen van plannen voorziet, zal hier uiteengezet moeten worden hoe plannen dan wel tot stand komen. Twee overwegingen zijn daarbij van belang:

1. De koppeling van kenmerken naar plannen moet duidelijk gemaakt kunnen worden en tevens hoe voorlopige plannen via toetsen uitgewerkt worden tot uitvoerbare plannen.
2. Het gaat hier niet om 'creatief' oplossen van problemen, maar om vaardigheden in aanpakken en oplossen die *onderwezen* zijn; met name moet de student over nogal wat vakkenkennis beschikken, en zal hij geoefend zijn in het toepassen van die kennis in bepaalde klassen van problemen. Denk daarbij aan de beheersing van termen en relaties, zoals in de hoofdstukken 4 en 5 behandeld.

Hiermee is nog niet aangegeven wat hier zoal onder 'plannen' verstaan kan worden. Ik wil beginnen met een voorbeeld:

SCHAKEN

Bij de halfopen verticale lijn is een gevolgsregel of plan: 'Bezetting met torens en/of dame. Bezet het eerste veld van deze lijn, liefst met een paard, en marcheer met de pionnen zodanig naar voren, dat de tegenpartij op de half open lijn of naburige lijnen één of meer zwakke pionnen krijgt (geïsoleerde of achtergebleven pionnen)' (Euwe).

De schaaktheorie voor gevorderde spelers bevat een klein aantal van deze 'gevolgsregels': er zijn niet veel meer dan een tiental van deze stellingskenmerken, met bij ieder kenmerk een of meer bijbehorende spelstrategieën.

Het behoeft geen toelichting dat het voor de schaakspeler niet

voldoende is om deze algemene regels uit het hoofd te leren: hij krijgt er pas enige beheersing over wanneer hij de regels heeft (zien en) leren toepassen. Er is volgens Euwe kennis van zaken, oefening en techniek voor vereist, maar die liggen binnen het bereik van de 'gevorderde' speler. Euwe: 'Is men eenmaal op de hoogte van een sluitende theorie, dan kan men op grond van bepaalde stellingskenmerken vaststellen, welke gedragslijn vereist is. De strategie kan worden uitgestippeld en aldus biedt de theorie voor de praktische partij een flink houvast.' Dat *houvast* op het probleem is belangrijk, en dat zit voornamelijk in de drastische en rationele inperking van het aantal speelmogelijkheden. De Groot (1978): 'Het plan maakt lange termijn ontwikkelingen mogelijk; het geeft de speler greep op de stelling; de veelheid van mogelijkheden wordt er door georganiseerd en ingeperkt; en het verengt het aantal te bestuderen zetten. Alleen dat wat in het plan past, hoeft overwogen te worden.'

Een belangrijk verschil tussen de gevorderde schaker en de schaakmeester zit in het zeer veel grotere aantal stellingskenmerken dat de schaakmeester weet te onderscheiden, mede omdat hij op de hoogte is van wat bij ieder van die spelsituaties een geschikt plan voor spelvoortzetting is. Docenten zijn zich doorgaans heel goed bewust van de analoge situatie ten aanzien van de problemen in het eigen vakgebied: ter behandeling in het onderwijs wordt daar een geschikte maar heel beperkte keuze uit gedaan, zodat de student zich over die klasse van problemen in enkele maanden tijds een redelijke beheersing kan verwerven; voor het verwerven van enig 'meesterschap' zijn immers veeleer jaren van studie en ervaring nodig. Een en ander betekent niet dat de student onmachtig zou zijn om problemen uit een nieuwe klasse aan te pakken: dat zal hem alleen erg veel moeite en vallen en opstaan kosten, omdat hij alles expliciet uit moet zoeken wat de ervaren probleemoplosser bijna onbewust uit zijn mouw kan schudden. Dat is zo bij schaken, en dat zien we in al die vakken waar het oplossen van problemen een prominente plaats in het onderwijs inneemt.

Heeft de analyse op kenmerken ertoe geleid dat het 'type probleem' herkend is en de daarbij vanuit de theorie bekende mogelijke plannen aangegeven zijn, dan zullen deze mogelijke

plannen eerst nog getoetst en dan verder uitgewerkt moeten worden. Zo wordt in het schaakspel een gekozen strategie verder uitgewerkt tot een te volgen tactiek:

‘We onderscheiden in het schaakspel strategie en tactiek. De strategie behelst de doelstelling en het vormen der plannen. De tactiek omvat de uitvoering der plannen. Populair gezegd komt het bij de strategie op denken aan, bij de tactiek op kijken’ (Euwe, 1966).

‘Kijken’ staat natuurlijk toch voor redeneren, en daar zit bij het schaken een sterke visuele component in. De strategie, het globale plan, wordt uitgewerkt tot de te volgen tactiek. Anders dan Euwe, wil ik eraan vasthouden dat een gekozen tactiek een *gedetailleerd* plan is. Dit illustreert dat de vorming van een uitvoerbaar plan in fasen kan gebeuren, waarin een eerste globaal plan, na toetsing, uitgebouwd wordt tot een steeds meer gedetailleerd plan. Zo kan het, en zo hoeft het dus niet altijd te gaan. Een omvangrijk probleem kan bijvoorbeeld eerst opgesplitst worden in deelproblemen die achtereenvolgens op te lossen zijn en in die volgorde elkaars oplossing ook mogelijk maken. Het kan ook zijn dat een gedetailleerd plan ontwikkeld moet worden zonder dat een globaal plan al beschikbaar is. Een voorbeeld van het laatste is de schaakstelling die zich niet voor een strategische benadering leent, maar waarin ‘een nauwkeurige berekening van de mogelijkheden van het ogenblik vereist is (tactische methode).’ Euwe: ‘Men spreekt in deze gevallen van *combinatiespel*, in tegenstelling met het *positiespel* waar de kenmerken van de stelling de hoofdrol spelen.’

Bij het schaken kan een ‘plan’ een strategie zijn (globaal plan) of een tactiek (gedetailleerd plan), waarbij de strategie aansluit op de kenmerken van de stelling en de tactiek een uitwerking van de strategie is, waarbij met die kenmerken (en eventuele andere kenmerken) alleen maar rekening gehouden hoeft te worden. Zo zou het aanpakken van problemen in sommige vakgebieden ook beschreven kunnen worden: er worden een beperkt aantal *probleemtypen* behandeld, waarbij het type *herkend* kan worden aan bepaalde kenmerken of een configuratie van kenmerken en waarbij een of enkele passende strategieën (globale plannen) horen. Een eenvoudig voorbeeld zijn problemen waarvoor bij een gegeven onderzoekopzet en

resultaten daarvan een statistische methode voor het toetsen van de onderzoekshypothese gevonden moet worden (vergelijk 5.2). Plannen kunnen echter nog in vele andere gedaanten voorkomen, waarvan er hier enkele genoemd worden.

Het plan in de gedaante van een *voorlopige of hypothetische oplossing*. Een goed voorbeeld daarvan vinden we bij diagnostische problemen, waarbij aan de hand van de geanalyseerde kenmerken van de probleemsituatie een aantal mogelijke diagnoses geopperd worden of althans één mogelijke diagnose. De verdere aanpak bestaat dan uit het toetsenderwijs komen tot een definitieve diagnose of toetsenderwijs een globaal gestelde diagnose verder verfijnen. De koppeling van geanalyseerde kenmerken in de probleemsituatie naar deze voorlopige diagnoses is een heel directe, als zodanig meestal geleerde en geoefende: bij kenmerk A, of configuratie van kenmerken B en C, moeten/kunnen de (globale) diagnoses X en Y overwogen worden. Het 'plan' is hier simpelweg de geopperde diagnose; voor diagnostische problemen zou je in fig. 7.1 in plaats van 'plan' in datzelfde hokje kunnen invullen 'diagnose' of 'differentiaaldiagnose' en dergelijke.

Het metaplan. Bij veel problemen kan niet in een keer een plan voor de oplossing van het probleem in haar geheel gemaakt worden. We zagen dat al bij diagnostische problemen, waarbij de achterliggende bedoeling (het metaplan) is om via toetsing en extra informatieverzameling een steeds betere diagnose te kunnen stellen (rivaliserende diagnoses uit te kunnen schakelen).

Een metaplan is een plan om een probleem op te lossen door eerst bepaalde deelproblemen op te lossen. Dergelijke 'metaplannen' worden eveneens onderwezen, maar dat zal vaak gebeuren in de vorm van gegeven voorbeelden of demonstratie van de bedoelde aanpak. Een voorbeeld is de 'aanwijzing' bij problemen in de thermodynamica om eerst te proberen een of meer geldige wetten (Mettes en Pilot, 1980: kernbetrekkingen) te vinden waarin de gevraagde grootte voorkomt; dit is een deelprobleem en de oplossing hiervan brengt

veelal een nieuw deelprobleem met zich mee omdat er in de zo gevonden wet of wetten nieuwe onbekenden voorkomen, waarvoor op dezelfde wijze nieuwe en geldige wetten gevonden moeten worden. Het metaplan is om op dezelfde wijze een aantal wetten (kernbetrekkingen, vergelijkingen) te vinden, waarmee het in beginsel mogelijk moet zijn om de gevraagde onbekende uit te drukken in termen van hetgeen gegeven is of aangenomen mag worden. De volgende stap in dit metaplan is het vinden van de wiskundige bewerkingen die tot dat verlangde resultaat kunnen leiden. Ieder van deze deelproblemen laat zich afzonderlijk in het schema van fig. 7.1 afbeelden. In het beschreven thermodynamica-probleem wordt de oplossing geleidelijk opgebouwd uit de oplossing van de achtereenvolgens aangepakte deelproblemen. In andere gevallen kan het zijn dat deelproblemen los van elkaar worden opgelost en dat vervolgens uit deze deeloplossingen nog een oplossing voor het gehele probleem gevonden, geconstrueerd of samengesteld moet worden.

Andere metaplannen zijn: bij een nieuw type probleem eerst nagaan of er analoge of verwante problemen bekend zijn waarvoor een oplossingsroute al bekend is; proberen een complex probleem te versimpelen, bijvoorbeeld eerst een speciale variant of een extreem geval aan te pakken en vandaaruit terug te keren naar het meer algemene of complexe probleem; eerst een model maken en daarmee experimenteren of eerst een simulatie uitvoeren. Het zal duidelijk zijn dat metaplannen gemakkelijk de vorm kunnen aannemen van specifieke methoden voor het aanpakken van problemen, bijvoorbeeld voor bewijsproblemen de techniek 'bewijzen vanuit het ongerijmde'. Ook voor deze metaplannen geldt dat de student daarmee heeft leren werken, ze in zijn repertoire van aanpaktechnieken heeft opgenomen, voor zover de problemen ten minste in een (eind)toets met behulp van dergelijke metaplannen aangepakt moeten kunnen worden.

De koppeling van kenmerken naar plannen, welke bepaalde vorm die 'plannen' ook mogen hebben, heeft als basisvorm: 'als dit, doe of zoek dan dat', 'als je dit symptoom tegenkomt, check dan altijd deze mogelijke diagnose'. In de psychologische (en Kunstmatige-Intelligentie) literatuur over het oplossen van problemen worden deze conditie-actie-paren aange-

duid met de technische term 'produkties'. Dergelijke produkties zijn de bouwstenen van computerprogramma's voor het oplossen van problemen waarbij met een groot kennisbestand wordt gewerkt, bijvoorbeeld voor het stellen van diagnoses (zie Duda en Gaschnig, 1981, voor een recent overzicht van dergelijke computerproduktiesystemen).

Het vervelende van die koppeling van kenmerken naar plannen is nu dat het in het onderwijs nog vaak voorkomt dat de leerstof niet in de daartoe juiste vorm wordt aangeboden of door de studenten wordt verwerkt. Het gaat erom dat bij een gegeven kenmerk de bijpassende informatie geproduceerd moet kunnen worden, maar dat hoeft niet de manier te zijn waarop de student die informatie beheerst! Het maakt veel verschil of men leert bij een bepaald symptoom altijd aan een bepaalde mogelijke diagnose te denken of dat men leert dat die diagnose past bij kenmerken *zus en zo*. Immers, wie leert bij een bepaald kenmerk altijd diagnose X te overwegen, hoeft bij ieder nieuw geval waarin zich dat kenmerk voordoet niet eerst te zoeken onder alle mogelijke ziektebeelden totdat de 'passende' diagnose X gevonden wordt. Wie leert bij ieder ziektebeeld op te sommen in welke gevallen dat als diagnose geopperd kan worden en niet het omgekeerde, zal in ieder nieuw geval alle ziektebeelden langs moeten zoeken om mogelijk passende te 'vinden'.

Hier ligt een van de oorzaken waarom studenten onverwacht veel moeite kunnen hebben met het toepassen van hun kennis bij het oplossen van problemen (zie ook Simon, 1980).

Toetsen. Een geopperd plan hoeft nog geen goed plan te zijn en om daar achter te komen is het doorgaans niet nodig om te proberen het plan uit te voeren. Een voorlopige diagnose wordt meestal niet op haar juistheid getoetst door te kijken of de patiënt goed reageert op de medische behandeling die bij die diagnose behoort. Nee, het algemene stramien is dit: bepaalde kenmerken leiden tot een of meer globale plannen, en of een globaal plan 'past' bij het onderhavige probleem wordt getoetst door na te gaan of in de probleemsituatie bepaalde andere kenmerken voorkomen. Bijvoorbeeld: wil deze voorlopige diagnose mogelijk juist zijn, dan moet een bepaald laboratoriumonderzoek leiden tot dit specifieke resultaat. Een juridisch voorbeeld (Crombag c.s., 1972):

VOORBEELD

Een cruciale passage in een juridische casus is 'Jansen overhandigt een boek aan Pietersen', wat zich in juridische termen zou kunnen laten vertalen als ofwel 'levering', 'schenking', of 'het aangaan van een overeenkomst van bruikleen'. De te kiezen vertaling leidt meteen tot een bepaald programma voor de verdere uitwerking van de casus, en is dus een voorlopig 'plan'. Welnu, de voorwaarden waaronder er gesproken kan worden van 'levering' enz. liggen vast in rechtsregels (o.a. in wetsartikelen, maar ook in jurisprudentie), en getoetst moet worden of het voorliggende probleem die voorwaarden bevat.

Voor een natuurkundig probleem geeft Larkin (1981) de volgende beschrijving:

VOORBEELD

'Zo kan de deskundige bij het zoeken naar een plan voor de oplossing besluiten het eerst eens met "momenten" te proberen. Hij is dan in staat om alle momenten die in het probleem voorkomen te "zien" en weet dan automatisch welke informatie wel en welke niet te verkrijgen is. Hij kan dan eenvoudig deze benadering verwerpen of voortzetten met een vrij grote zekerheid over het eventueel welslagen ervan. Zonder dit vermogen om automatisch informatie te genereren zou het ontwerpen van plannen een heel omslachtige werkwijze zijn die misschien nauwelijks te onderscheiden is van het simpelweg uitwerken van probeersels.' Wat de deskundige hier 'automatisch' doet, is het toetsen van het voorlopige plan. De nog niet zo goed ingevoerde student zal dat niet 'automatisch' kunnen. Larkin: '(...) onervaren oplosers opperen bij een overwogen regel enkele algemene uitspraken over de toepasbaarheidsvoorwaarden, en die vergelijken zij expliciet met de gemaakte afbeelding van het probleem.'

Wat de ervaren oplosser 'automatisch' en pijlsnel doet, moet de onervaren oplosser heel bewust en enigszins moeizaam doen. Dat is vaak ook het verschil tussen de docent en zijn studenten, en dat kan het onderwijzen van het aanpakken en oplossen van problemen bemoeilijken. In stappenschema's voor het aanpakken van juridische casus (Crombag c.s., 1972), problemen in de thermodynamica (Mettes en Pilot,

1980) of bestuursproblemen (Terlouw e.a., 1981) wordt dan ook veel aandacht aan dergelijke expliciet door de student uit te voeren toetsen gegeven.

Een bijzondere vorm van toetsen is die waarbij blijkt dat de probleemstelling onvoldoende informatie bevat om de toets uit te kunnen voeren. Dat kan ertoe leiden dat voor het ontbrekende gegeven een bepaalde *aanname* gedaan wordt.

THERMODYNAMICA

‘Omdat betrekkingen slechts onder bepaalde voorwaarden geldig zijn, en in de *praktisch* voorkomende gevallen daar zelden aan is voldaan, moet men veelvuldig aannamen doen (dat iets verwaarloosbaar is) c.q. vereenvoudigingen uitvoeren, (...) Het is daarbij bovendien van belang dat de oplosser in staat is vast te stellen welke aannamen wèl en niet zijn toegestaan. Daarvoor moet hij een indruk zien te krijgen van de fout die mogelijk wordt gemaakt bij het doen van de aanname. Bovendien zal hij eerst na moeten gaan of de aanname noodzakelijk is, en de redenen voor de aanname expliciet moeten maken.’

(Mettes en Pilot, 1980).

Het is niet altijd mogelijk om ieder plan meteen volledig te toetsen. Het opwerpen van plannen en bijstellen van plannen gaat in voortdurende wisselwerking met die toetsen die wel uitgevoerd kunnen worden. Dat kan er dus toe leiden dat meerdere plannen tegelijkertijd gehandhaafd worden en steeds verder uitgewerkt, totdat er voldoende informatie uit verdere toetsing is verkregen dat het mogelijk wordt om bepaalde plannen definitief te laten varen.

Het expliciet uitvoeren van de stappen: analyse op kenmerken - plan opperen - toetsen, is van buitengewoon belang bij het aanpakken van problemen. Het is een veel voorkomende beginnersfout om na het lezen van de probleemstelling meteen een oplossing te gaan uitwerken. In de thermodynamica uit zich dat in het te snel overgaan tot het manipuleren van formules en reduceren van het aantal onbekende grootheden (en zonder een geschikt plan mislukt dat meestal).

Uitvoeren - antwoord - toetsen. Is er een voldoende gedetailleerd plan opgesteld, dan is het uitvoeren van het plan, eventu-

eel het formuleren van het antwoord, en het toetsen of het antwoord voldoet als antwoord op het gegeven probleem niet 'problematisch' meer. Mettes en Pilot (over problemen in de thermodynamica) spreken over het 'uitvoeren van standaardbewerkingen'. Het kan blijken dat het plan ondeugdelijk is, dat het geen antwoord oplevert, of een ondeugdelijk of onaanvaardbaar antwoord; dan wordt een beter plan gezocht, eventueel na een nieuw onderzoek naar relevante kenmerken in de probleemsituatie. Meestal wordt voor het uitvoeren, formuleren van het antwoord en toetsen van het antwoord, aanspraak gedaan op bekende en geoefende technieken en vaardigheden (wiskundige technieken bijvoorbeeld).

7.2 Inventarisatie

De vuistregels voor het stellen van problemen, die in de volgende paragraaf worden gegeven, worden toegepast op 'de leerstof'. Het zal doorgaans nodig zijn om 'de leerstof' op een geschikte manier te beschrijven of te inventariseren om efficiënt met die vuistregels te kunnen werken. Het ligt voor de hand om te inventariseren naar de verschillende onderdelen in het schema van fig. 7.1 aangegeven. Inventariseren hoeft niet hetzelfde te zijn als het verzamelen en ordenen van 'oude' problemen, maar dat laatste zal wel een heel zinnige aanvulling van en controle op de inventarisatie kunnen vormen.

Verskillende werkwijzen zijn mogelijk. Zo kunnen eerst 'oude' problemen geanalyseerd worden op relevante kenmerken, plannen, toetsen enz. Beter lijkt het om die 'oude' problemen te gebruiken als controle achteraf op de eerst alleen op grond van analyse van de leerstof te verrichten inventarisatie. Zo kan een logische of *a-priori*-inventarisatie aangevuld worden met al hetgeen uit de 'oude' problemen blijkt ook nog gevraagd te worden.

Kenmerken. Om de relevante kenmerken van een probleemsituatie te identificeren, moet het probleem behoorlijk geanalyseerd worden, moeten er misschien tekeningen, schema's of vertalingen gemaakt worden. Zie voor deze zaken wat in voorgaande hoofdstukken behandeld is (vooral 5.1, 6.3 en 6.4). Probeer nu een opsomming te maken van alle mogelijk

voorkomende kenmerken die van belang kunnen zijn bij het kiezen van plannen. Zo geeft bijvoorbeeld Euwe (1966) een opsomming van stellingskenmerken, waar de gevorderde schaakspeler op aan moet sluiten bij de keuze van een strategie, dat is een plan voor de voortzetting van de partij. In de medische diagnostiek kan, ten minste voor deelgebieden, een opsomming gemaakt worden van de belangrijke kenmerken, symptomen of syndromen, die goede aanknopingspunten voor voorlopige diagnoses bieden (zie De Graaff en Galesloot, 1981, voor een ingang tot de literatuur over het oplossen van medische problemen).

Maak zonodig onderscheid tussen kenmerken waarop direct met mogelijke plannen kan worden aangesloten en kenmerken die het mogelijk maken om allereerst het *type* of de *klasse* van het voorgelegde probleem te bepalen (waarna dan mogelijk andere kenmerken van belang zijn om binnen dat type probleemstelling tot de keuze van mogelijk bruikbare plannen te komen). Dit onderscheid is bijvoorbeeld te vinden bij Mettes en Pilot, 1980: 'mogelijk bruikbare kernbetrekkingen (= 'deelplannen') kunnen gekozen worden vanuit 'één of enkele thermodynamische kenmerken (type systeem, proces, enz,)' of 'vanuit een typering van het vraagstuk *als geheel* (bijvoorbeeld een "typisch Eerste Hoofdwet probleem", een evenwichtsvraagstuk)'.

Sluit bij het inventariseren aan op de aard van de betreffende klasse van problemen. Bijvoorbeeld: thermodynamische problemen, althans in de analyse van Mettes en Pilot, zijn beter aan te pakken door te werken vanuit het gevraagde dan door te werken vanuit de gegevens, zodat vaak het gevraagde het eerste kenmerk is dat als ingang voor het zoeken naar een bruikbaar plan (kernbetrekking) dient. De kenmerken van de gegevens dienen dan om overwogen plannen (kernbetrekkingen) te kunnen toetsen op hun geldigheid en bruikbaarheid.

Een enigszins volledige lijst van mogelijk leidende kenmerken en van tenminste zijdelings relevante kenmerken kan opgesteld worden door er de theorie van het desbetreffende vak op door te nemen. Dat hoeft niet meteen ook een volledige lijst op te leveren: het kan zijn dat andere onderdelen in de inventarisatie wijzen op andere kenmerken die nog aan deze lijst toegevoegd moeten worden. Ook kan bij het analyseren van 'ou-

de' problemen blijken dat de gemaakte lijst nog verder aan te vullen is.

Plannen. Een eerste lijst van mogelijke plannen kan opgesteld worden door er de theorie van het betreffende vak op door te nemen. Dat levert bijvoorbeeld alle mogelijke globale diagnoses op of alle belangrijke wetten, technieken en instrumenten (zoals de kernbetrekkingen in de thermodynamica).

Voordat zo'n lijst van plannen afgezet wordt tegen de lijst kenmerken kan het wenselijk zijn eerst enige ordening in die misschien wel erg heterogene verzameling van plannen te brengen. Misschien zit er een bepaalde logische structuur in, die zeker bij het oplossen van problemen uitgebuit zou kunnen en moeten worden. Wanneer een vak erg rijk is aan wetten of kernbetrekkingen kan het zijn dat een aanzienlijk deel van die wetten zich af laat leiden van wetten uit een veel kleinere groep. Zie voor dit soort overwegingen ook Mettes en Pilot (1980, blz. 100 e.v.). Om nu een goed overzicht te krijgen van alle relevante koppelingen van kenmerken naar mogelijke plannen kan daarvoor een kruistabel worden gemaakt, met bijvoorbeeld horizontaal de rij kenmerken of clusters van kenmerken en verticaal de rij plannen (wetten, technieken, ziektebeelden enz.).

Bij een andere wijze van inventariseren worden bij ieder gegeven kenmerk alle mogelijke bruikbare plannen gezocht. Deze aanpak ligt vooral dan voor de hand wanneer de leerstof ook al op deze wijze is georganiseerd. Welke werkwijze ook gevolgd wordt, verbindt altijd de kenmerken aan de plannen, want dat is waar het in deze fase van de probleemaanpak om gaat. Het overzicht dat hieruit ontstaat, bij voorkeur in de vorm van een kruistabel, biedt de mogelijkheid om op een andere wijze nog eens systematisch na te gaan of er nog leemten zijn, of er bepaalde eigenaardigheden in dit overzicht zitten, of een en ander zich goed verhoudt tot wat de doelen van het onderwijs zouden behoren te zijn, en dergelijke.

In het bijzonder zou zo wel eens ontdekt kunnen worden dat het onderwijsprogramma een te rijke inhoud heeft om voor studenten nog overzichtelijk te kunnen zijn.

Toetsen. Opgeworpen plannen moeten getoetst worden op hun geldigheid of mogelijke bruikbaarheid.

VOORBEELD

'Vele thermodynamische betrekkingen zijn slechts geldig als het proces aan bepaalde voorwaarden voldoet. We onderscheiden vier kenmerken:

1. reversibiliteit,
2. wel of geen chemische reactie,
3. procesgrootheden die nul zijn, en
4. toestandsgrootheden die constant zijn.'

(Mettes en Pilot, 1980)

Het toetsen van plannen gebeurt altijd tegen bepaalde kenmerken van de probleemsituatie; dat kunnen dus heel andere kenmerken zijn dan de kenmerken die allereerst van belang zijn bij het *kiezen* van plannen. Wanneer deze kenmerken op een of andere wijze al in de gegevens van de probleemstelling voorkomen, is de controle meestal eenvoudig te verrichten (en is het allerbelangrijkste *dat* die toets wordt uitgevoerd). Het kan zijn dat voor het produceren van de gevraagde kenmerken een analyse van de probleemstelling nodig is.

VOORBEELD

'Het vaststellen van de kenmerken van de toestanden waarin het systeem voorkomt. (...) Een systematische analyse omvat minstens een nagaan van volume, druk, temperatuur, massa (aantal molen) en energetische gegevens als de thermodynamische potentiaal, enthalpie, entropie, e.d.'

(Mettes en Pilot, 1980)

Lastiger is het toetsen van de geldigheid van plannen wanneer de daartoe benodigde informatie over kenmerken niet (na analyse) waarneembaar of gegeven blijkt te zijn. Soms moet die informatie dan ingewonnen worden, door navragen, onderzoek, raadplegen van tabellen en dergelijke. Denk bijvoorbeeld aan het extra (laboratorium)onderzoek dat nodig kan zijn om een opgeworpen diagnose op mogelijke juistheid te toetsen. Bij de inventarisatie kan het er niet alleen om gaan voor ieder mogelijk plan op te sommen aan welke voorwaarden de probleemsituatie moet voldoen, wil het plan mogelijk bruikbaar zijn, maar daar ook deze complicaties aan toe te voegen: op welke manieren informatie die niet gegeven is toch door verder onderzoek en dergelijke geproduceerd moet kun-

nen worden.

Sommige voorwaarden zijn binnen een gegeven probleemstelling niet te toetsen, eenvoudig omdat de nodige gegevens ervoor definitief ontbreken of omdat een onderzoek veel te kostbaar, tijdrovend, pijnlijk of iets dergelijks zou zijn. Geef dan aan hoe met die voorwaarden dan wel omgegaan moet worden. Meestal zal de student er dan een bepaalde *aanname* over moeten doen, eventueel ondersteund door argumentatie of door een onderzoekje naar tenminste de globale aanvaardbaarheid van die aanname in die situatie. Aannames die regelmatig voorkomen bij problemen in het betreffende vakgebied horen in de inventarisatie opgenomen te worden. Wanneer er geen aannames gemaakt kunnen of mogen worden, zal de betreffende voorwaarde op een andere manier in de aanpak en de oplossing van het probleem meespelen. Misschien moet deze voorwaarde *beredeneerd* kunnen worden uit andere beschikbare gegevens: als noodzakelijkerwijs een van drie mogelijkheden het geval is en je kunt er twee uitsluiten, dan is de derde mogelijkheid het geval. Of het probleem moet verder aangepakt kunnen worden zonder aanvankelijk deze voorwaarde te kunnen toetsen, omdat de nodige informatie pas later beschikbaar komt of omdat de uitkomst zelf als toets dient; een speciale en mogelijk veel voorkomende vorm is het louter algebraïsch in de verdere berekeningen meenemen van een grootte die niet bepaald kan worden.

Uitvoeren, antwoord en toetsen. Het gaat hier om het opsommen van de technieken of vaardigheden die nodig zijn bij het uitvoeren van uiteindelijk gekozen en uitgewerkte plannen. Soms is zo'n plan tevens het 'antwoord', zoals een definitieve diagnose dat is (de behandeling die aansluit op de diagnose wordt meestal als een afzonderlijk 'probleem' gezien). Maar meestal vraagt juist de uitvoering om een vrij grote vaardigheid in het hanteren van speciale technieken, zoals het oplossen van problemen in de thermodynamica nogal veel wiskundige routine vraagt. Wanneer er sterke relaties zijn tussen bepaalde plannen en bepaalde technieken neemt de inventarisatie deze bij elkaar op; in het andere geval kan volstaan worden met een algemeen geldende lijst van te beheersen technieken enz.

Wanneer er voor de controle van de redenering, de tussenre-

sultaten of het eindresultaat bepaalde welomschreven handelingen verricht of technieken toegepast moeten worden, dan horen die in de inventarisatie thuis. Lastiger is het wanneer antwoorden niet eenduidig goed zijn, maar op aanvaardbaarheid beoordeeld moeten worden. Zijn daar bepaalde criteria voor gegeven, dan horen die in de inventarisatie, eventueel gekoppeld aan die bepaalde problemen of typen antwoorden waar ze alleen voor gelden. Zijn dergelijke criteria niet expliciet gegeven en moet de student zijn antwoord toch waarderen naar eisen die gewoonlijk in dat vakgebied gesteld worden, eisen die docenten plagen te hanteren of eisen die de student zelf moet kunnen formuleren, dan moet iets daarvan in de inventarisatie toch nader omschreven worden.

'Oude' problemen. Verzamel 'oude' problemen, orden deze op onderwerp, type of iets dergelijks. Analyseer problemen en oplossingen in de termen van fig. 7.1 en leg de resultaten van deze analyses naast de al gemaakte inventarisatie.

De vergelijking zal erin resulteren dat de inventarisatie hier en daar uitgebreid zal worden met punten die aanvankelijk over het hoofd zijn gezien. Het kan echter ook zijn dat de inventarisatie veel 'breder' is geweest dan blijktens de vragen in de praktijk nodig was geweest, en dan kan er gesnoeid worden. Een verschil tussen inventarisatie en wat er in de praktijk gevraagd blijkt te worden, kan natuurlijk ook wijzen op bepaalde leemten in het scala van problemen dat aan studenten wordt voorgelegd, dan moeten er problemen gevonden worden die in deze leemten passen, en daartoe kunnen de vuistregels in de volgende paragraaf gebruikt worden. Bij het stellen van nieuwe problemen kan een overzichtelijk geordende verzameling van de oude problemen overigens heel goede diensten bewijzen.

7.3 Vuistregels

De uiteenzetting over problemen, hun aanpak en oplossing in 7.1 en de daarop aansluitende inventarisatie, bieden veel nieuwe handvatten bij het bedenken en stellen (formuleren) van nieuwe problemen. In deze paragraaf worden een aantal vuistregels gegeven die het stellen van problemen vergemakkelijk-

ken en richting geven. Dat gaat zowel over 'ideeën' voor nieuwe problemen, als over een techniek om nieuwe problemen te 'construeren'.

Een 'idee' voor een probleem. Een 'idee' voor een probleemstelling is bijvoorbeeld een concrete probleemsituatie of een bepaalde combinatie van een situatie met een bepaald kenmerk en een bruikbaar plan. Het zoeken naar een idee voor een nieuw te stellen probleem wordt dan het zoeken naar een dergelijke nieuwe situatie of een aardige combinatie. Probeer in ieder geval te omschrijven wat in termen van fig. 7.1 een 'idee' voor een nieuw probleem is, omdat daardoor het zoekproces in de literatuur enz. richting krijgt.

Een altijd bruikbare bron voor ideeën zijn andere studieboeken dan die welke door studenten gebruikt worden. Niet omdat de theorie daarin op iets andere wijze gepresenteerd wordt, maar omdat de daarin opgegeven problemen bruikbare ideeën voor nieuw te stellen eigen problemen kunnen bevatten.

Niet alleen eigentijdse, maar ook verouderde studieboeken kunnen daarvoor gebruikt worden. Problemen zijn niet altijd direct herkenbaar als mogelijk bruikbaar, maar moeten eerst uitgewerkt, opgelost en geanalyseerd worden op relevante kenmerken, bruikbare en onbruikbare plannen en dergelijke; deze inspanningen vormen geen verloren tijd en energie, omdat ze meteen het materiaal voor het modelantwoord opleveren.

Artikelen in de vakliteratuur vormen een onuitputtelijke bron voor ideeën. Voor sommige vakken, zoals rechten, levert zelfs de dagbladpers talrijke bruikbare thema's. Het zoeken in deze bronnen gebeurt met enig overleg, bijvoorbeeld met het oog speciaal gericht op de probleemsituaties die erin beschreven worden, de plaats van de probleemkenmerken, de gebruikte en eventueel de verworpen plannen. Problemen uit de vakliteratuur zijn natuurlijk niet zonder meer geschikt om studenten in een toetsituatie voor te leggen, maar het aanbrengen van vereenvoudigingen, het aanreiken van extra gegevens en het geven van aanwijzingen kunnen daar veel aan veranderen. De beroepspraktijk, indien van toepassing, is een voor de hand liggende bron van ideeën. Het zal vaak mogelijk zijn om enigszins stelselmatig te inventariseren welke verschillende

typen problemen zich voordoen, welke varianten er binnen ieder type te verwachten zijn en in welk scala van concrete situaties die problemen zich kunnen voordoen. Vakken waarin zo'n stelselmatige benadering mogelijk is (denk aan geneeskunde), zullen de leerstof in niet geringe mate al op deze wijze geordend hebben. Een aardige bron voor nieuwe ideeën is te vinden in de geschiedenis van het betreffende vak. En dan niet alleen de experimenten en theorieën die deze 'geschiedenis gemaakt hebben', maar ook de minder bekende, meer alledaagse worsteling met de problemen van het vak. Wanneer enige historische verdieping tot de onderwijsdoelen behoort, kan dat fraai met het aanpakken en oplossen van problemen gecombineerd worden. Overigens zal verrassend vaak een sterke overeenkomst blijken te bestaan tussen de problemen waarmee vakgenoten in een grijs verleden dagelijks bezig waren en de problemen waar studenten zich nu vertrouwd mee moeten maken; ook dat kan uitgebuit worden bij het zoeken naar ideeën voor nieuw te stellen problemen.

Studenten kunnen zelf nieuwe problemen, of ten minste nieuwe probleemsituaties aandragen. Die zijn weliswaar niet onmiddellijk bruikbaar, maar kunnen wel bewaard worden om er voor studenten in latere jaren nieuwe problemen uit te construeren.

Een heel andere ingang voor het vinden van ideeën is niet in de *inhoud* van problemen gelegen, maar in de *aard* ervan. Dit is verwant aan de nog te behandelen mogelijkheid tot het stellen van deelproblemen. Een illustratieve opsomming:

- niet zelf een probleem laten aanpakken, maar een gegeven oplossing laten controleren, bekritisieren, verbeteren of er alternatieve mogelijkheden voor aan laten geven;
- laten verklaren waarom een gegeven resultaat kan gelden voor een eveneens gegeven probleemstelling (analoog aan bewijsopgaven in de wiskunde); het laten verklaren van verschijnselen;
- een opstel laten maken over een opgegeven thema (dit hoeft niet alleen voor het taalonderwijs een relevant type probleemstelling te zijn);
- een onderzoekopzet laten ontwerpen bij een gegeven theoretische veronderstelling, te verklaren verschijnsel enz,;
- de resultaten van een gegeven onderzoek laten bespreken,

afwijkingen van de verwachte resultaten laten verklaren, alternatieve veronderstellingen laten produceren, een goede voortzetting voor het onderzoek laten bedenken.

Nieuwe problemen construeren. De wetenschap verlegt grenzen: voor ieder opgelost probleem komen er wel enkele nieuwe in de plaats. Maar dit zijn niet de problemen die studenten in een afsluitende toets moeten kunnen oplossen. Integendeel, in het onderwijs gaat het erom studenten door oefening vertrouwd te maken met bepaalde typen problemen. Een nieuw probleem kan dan nog wel moeilijk blijken voor de goed geoefende student, maar het mag hem niet meer voor verrassingen plaatsen. Welnu, het valt te verwachten dat dergelijke problemen op enigszins stelselmatige wijze geconstrueerd kunnen worden op grond van het materiaal dat in de inventarisatie verzameld is. Neem bijvoorbeeld diagnostische problemen. In beginsel kunnen alle mogelijke 'plannen' ofwel diagnoses opgespoord worden, in ieder geval die welke in het studieboek en in het onderwijs behandeld zijn. Bij iedere diagnose kan aangegeven worden bij welke symptomen enz. deze diagnose in ieder geval geopperd en overwogen moet kunnen worden. Dan kan bij ieder koppel van *kenmerk* (symptoom) met *plan* (diagnose) een groot aantal bijpassende problemen gesteld worden door *concrete* situaties te inventariseren of te bedenken waarin dat bepaalde kenmerk op een relevante manier voorkomt. Ik leg de nadruk op concreet, en verwijst naar wat er in voorgaande hoofdstukken al is opgemerkt over het belang van het werken met concrete situaties in tegenstelling tot in abstracte termen beschreven situaties. Omdat er bij een gegeven plan (diagnose) meerdere kenmerken als aanknopingspunt zouden kunnen fungeren, ontstaat bij het construeren van concrete probleemsituaties de uitwaaiering die in fig. 7.2 in beeld is gebracht.

Volgens het schema van fig. 7.2 geconstrueerde problemen zijn niet zonder meer bruikbaar. De richting van de constructie is precies omgekeerd aan de richting waarin de student het probleem moet aanpakken. Daarom moet het geconstrueerde probleem nog geanalyseerd worden op het voorkomen van andere kenmerken die voor het zoeken naar plannen van belang kunnen zijn, plannen die daar bij passen en de wijze waarop de

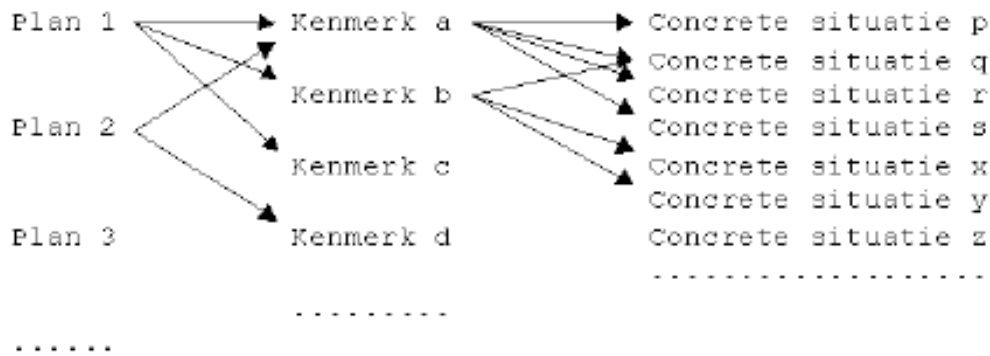


Fig. 7.2 Constructie van een probleemstelling.

onbruikbaarheid van die plannen getoetst kan worden, en dergelijke. Ook wanneer de constructie van nieuwe problemen vertrekt vanuit een inventarisatie van concrete situaties, moeten de voor de hand liggende, maar onbruikbare plannen in de analyse betrokken worden. Constructie en analyse verlopen zodoende tegengesteld, ze geven aanleiding tot verschillende waaiers van mogelijkheden zoals vergelijking van fig. 7.2 met fig. 7.3 laat zien.

Het belang van een analyse zoals die welke in fig. 7.3 abstract is aangegeven, is duidelijk. Zo wordt een beeld verkregen van hoe complex het probleem voor de student is, welke plannen de student door middel van de bijbehorende toetsen moet weten uit te sluiten voor het geval hij niet meteen het 'juiste' kenmerk met het geschikte plan kiest.

De analyse levert belangrijke aanwijzingen voor de definitieve inkleding van een probleemstelling. Het is al gauw noodzakelijk om een eerste probleemformulering uit te breiden met informatie of bepaalde aanwijzingen die in ieder ge-

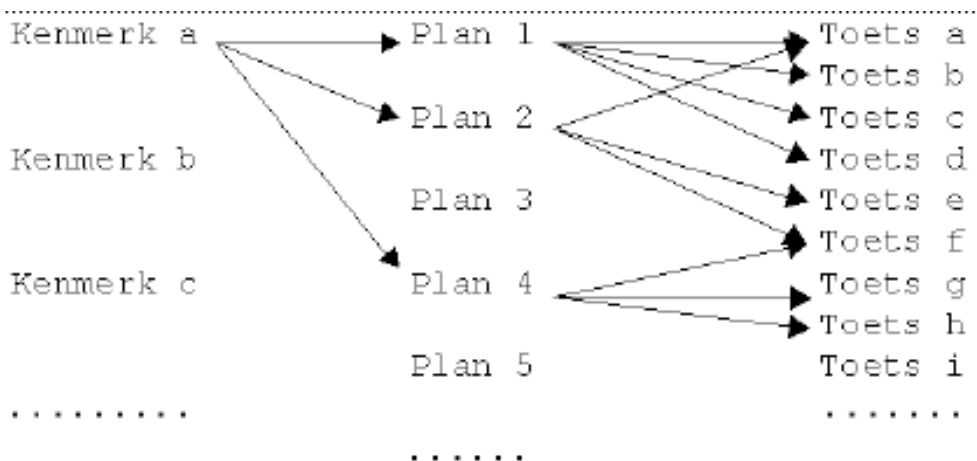


Fig. 7.3 Analyse van een ontworpen probleemstelling

val een aantal plausible maar toch onbruikbare plannen meteen uitsluiten. De analyse moet duidelijk maken dat noodzakelijke informatie voor het toetsen van alle mogelijke plausible plannen in de probleemstelling of anderszins beschikbaar is, ook voor plausible plannen die onbruikbaar zijn en juist door middel van toetsing uitgesloten moeten kunnen worden. Informatie die blijkt te ontbreken kan dan alsnog worden toegevoegd. Ondanks zo'n zorgvuldige analyse moet er nog een onafhankelijke controle op de bruikbaarheid van de ontworpen probleemstelling worden uitgevoerd, bijvoorbeeld door het probleem door een of meer collega's te laten oplossen (zie ook hoofdstuk 8). Vraag die collega's dan ook om aan te geven welke plausible plannen zij nog meer 'gezien' hebben naast hun eigen bruikbare plan, ook in die gevallen waarin zij zelf onmiddellijk een bruikbaar plan 'zagen'.

Deze methode voor het construeren van problemen is voor het ene vakgebied meer zinvol dan voor het andere. Maar ook wanneer deze wijze van construeren minder handig is, kan de analyse uit fig. 7.3 niet gemist worden. Het contrast tussen fig. 7.2 en fig. 7.3 maakt duidelijk dat het onderschatten van de moeilijkheid van een probleem anders onontkoombaar is. Hier doet zich hetzelfde gezichtsbedrog voor als bij het kennisnemen van de oplossing van een gegeven probleem: wie eenmaal de oplossing heeft gezien, is vaak op stel en sprong vergeten welke doodlopende wegen de student te gaan heeft voordat hij het juiste plan geïdentificeerd heeft.

In het bovenstaande is aan het *uitvoeren* geen afzonderlijke aandacht besteed. Wanneer het uitvoeren van plannen zijn eigen problematische aspecten heeft, kunnen 'uitvoeringsproblemen' op dezelfde wijze geconstrueerd worden als hiervoor beschreven is voor problemen in het algemeen. Met uitzondering van moeilijkheden die liggen in het vinden van de juiste combinatie (tactieken!) of in het opzetten van een juiste redenering.

Deze constructieve benadering is een variant op de methode van de rompvragen (zie 2.5). Een koppel van een kenmerk met een passend plan kan men opvatten als een rompprobleem, een abstract probleem dat vraagt om verdere invulling of concretisering. De constructieve methode doet dan ook enige aanspraak op het vermogen van de docent-deskundige

om concrete probleemsituaties te kunnen bedenken of verzamelen bij zo'n abstract 'rompprobleem'.

Het *concrete karakter* van probleemsituaties is van groot belang, maar is ook een bron van moeilijkheden voor de student. De wetten, regels, mogelijke diagnoses en behandelingen uit het tekstboek zien er overzichtelijk uit, zijn gemakkelijk te leren, maar zeer moeilijk onder de knie te krijgen in die zin dat ze ook op heel uiteenlopende concrete situaties toegepast kunnen worden. Larkin (1981) zegt hierover: 'De principes die fundamenteel zijn omdat ze zo algemeen toepasselijk zijn in heel uiteenlopende situaties, zijn in het algemeen ook precies de principes die voor studenten moeilijk te leren zijn omdat ze daartoe moeten leren de betreffende symbolen in zoveel verschillende situaties te interpreteren. Die moeilijkheid wordt nog eens extra aangezet doordat deze wetten met hun algemene symbolen eruit zien als iedere andere gewone vergelijking met symbolen die je mag interpreteren louter met behulp van je kennis van wat dat symbool betekent.' Het feit dat bij het aanpakken van problemen de student 'spontaan' een concrete situatie moet weten te vertalen naar de relevante symbolen, maakt het er voor hem niet gemakkelijker op.

Modelantwoord. Voor het nakijken van antwoorden moet een modelantwoord beschikbaar zijn. Dat modelantwoord kan in grote lijnen opgezet worden volgens het schema in fig. 7.1. Het modelantwoord geeft uiteraard weer hoe de juiste aanpak en oplossing eruitziet, uitgesplitst naar kenmerken, plan, toets op het plan, uitvoering en controle op het antwoord. Eventueel worden alternatieve, ook aanvaardbare oplossingen eveneens expliciet uitgewerkt. Daarnaast zal er soms reden zijn om in het modelantwoord ook aan te geven welke onbruikbare plannen de student moet weten te toetsen en af te wijzen. Het modelantwoord moet voldoende gedetailleerd zijn om er het beoordelingsvoorschrift (zie onder) op te kunnen baseren. Het modelantwoord maakt integraal onderdeel van de probleemstelling uit en de controle op die probleemstelling, bestaande uit antwoorden die door collega's worden gegeven, zal vaak aanleiding geven om het modelantwoord bij te stellen of uit te breiden met aanvaardbare varianten.

Na afloop van de toets kunnen de modelantwoorden (en de

beoordelingsvoorschriften die erbij horen) aan de studenten gegeven worden.

Beoordelingsvoorschrift. Het beoordelingsvoorschrift legt vast hoe het nagekeken antwoord gescoord wordt. De puntentoekenning voor de verschillende onderdelen van aanpak en oplossing wordt bij voorkeur tot in enig detail vastgelegd, waarbij u op het eigen oordeel over het belang van de onderscheiden stappen moet afgaan. Deze puntentoekenning heeft een enigszins subjectief karakter en dat is ook een belangrijke reden om de puntentoekenning tevoren vast te leggen, zodat verschillende beoordelaars wat dit betreft ieder antwoord even 'streng' zullen beoordelen. Wanneer er nogal wat ruimte is voor verschillen van inzicht in de zwaarte van bepaalde onderdelen van het modelantwoord, kan gebruik worden gemaakt van dezelfde methode als die in hoofdstuk 8 voor de controle op toetsvragen wordt gegeven: laat collega's niet alleen de problemen beantwoorden, maar daarbij ook de volgens hen meest gewenste puntentoekenning voor onderdelen van het antwoord aangeven. De resultaten van dit 'experiment' zullen in ieder geval stof tot overleg geven en soms reden tot bezinning over de leerstof. Wanneer er een goede overeenstemming tussen onafhankelijk werkende collega's is, kan het beoordelingsvoorschrift conform het gezamenlijk oordeel worden opgesteld.

Wanneer er op bepaalde punten nogal verschil van inzicht bestaat, ligt het in de rede de achtergronden van die meningsverschillen te onderzoeken en al naar gelang de uitslag van dat overleg het beoordelingsvoorschrift vast te stellen. Stelregel hierbij moet zijn dat de student het voordeel van enig meningsverschil tussen deskundige beoordelaars toebedeeld moet krijgen, omdat er nu eenmaal niet getoetst wordt op het vermogen in de ziel van beoordelaars te kijken.

Een speciaal aandachtspunt vormt de doodgelopen probleemaanpak: de student heeft bijvoorbeeld geen bruikbaar plan kunnen opstellen en daardoor ook geen oplossing kunnen geven. Enige waardering voor het weten te genereren en toetsen van plausibele maar weliswaar onbruikbare plannen kan wel gegeven worden, maar daar houdt het dan wel mee op. Komt het vaak voor dat studenten al in het begin van de probleemaanpak blokkeren, dan is te overwegen of het niet beter is van

deelproblemen (zie hieronder) gebruik te maken, want dan zijn de gevolgen voor de student minder rampzalig. Minder ernstige varianten zijn die waarbij de student al vroeg een fout maakt, maar daarna toch op correcte wijze verder naar een oplossing heeft toegewerkt (een oplossing die op zich natuurlijk niet goed is); of waarbij de student de waarde van een bepaalde onbekende niet heeft kunnen vinden en deze onbekende verder als algebraïsche variabele in zijn berekeningen heeft meegenomen.

Deelproblemen. Het is altijd aardiger om te werken met problemen die van de eerste analyse af tot en met de controle op het resultaat beantwoord moeten worden. Maar het is vaak efficiënter om in plaats daarvan te werken met deelproblemen waarin maar een of twee kritische stappen gevraagd worden. Is het een belangrijk onderwijsdoel dat in concrete probleemsituaties de kenmerken goed worden onderscheiden, dan is het redelijk om in de toets die problemen op te nemen met alleen de vraag naar die bepaalde kenmerken, met weglaten van plannen, toetsen, uitvoeren en controleren. Het weglaten van veel tijd vragende berekeningen is al heel gebruikelijk in de vorm waarbij de inkleding van het probleem zo wordt gekozen dat de berekening gedaan kan worden met 'gemakkelijke' getallen, maar waarom zou men dan niet volstaan met alleen een goed plan voor de oplossing te vragen? Talrijke varianten voor deelproblemen zijn mogelijk, waarbij de zwaartepunten in de onderwijsdoelen aangeven op welke punten die deelproblemen toegespitst kunnen worden. De toets kan zodoende veel scherper toegesneden worden op die onderwijsdoelen.

Moeilijkheidsvarianten. Naast het opsplitsen tot deelproblemen zijn er talrijke andere mogelijkheden om de moeilijkheid van (deel)problemen te manipuleren: geven van extra aanwijzingen, verwerken van irrelevante informatie in de probleemstelling, beschikbaar stellen van naslagwerken, concrete probleemsituaties die meer dan wel minder gewoon of kenmerkend zijn, en dergelijke. Deze mogelijkheden zijn ook in voorgaande hoofdstukken al uitgebreid gegeven. Wees altijd bedacht op het gemak waarmee de moeilijkheid van een bepaald probleem onderschat kan worden en houd daarbij fig. 7.2 en fig. 7.3 voor ogen.

7.4 Literatuur

NICKLES, TH. (ed.), *Scientific discovery, logic, and rationality*, Dordrecht, Reidel, 1980. Ontdekken, het 'zien' van problemen, is een weer opkomend thema in de wetenschapsfilosofie (logic of discovery, the method of analysis).

TUMA, D. T., & F. REIF (eds.), *Problem solving and education: issues in teaching and research*, Hillsdale, New Jersey, Erlbaum, 1980.

NEWELL, A., & H.A. SIMON, *Human problem solving*, Englewood Cliffs, New Jersey, Prentice Hall, 1972.

Een standaardwerk over het oplossen van problemen. Staat met beide voeten in de Kunstmatige-Intelligentietraditie, probeert een beschrijvende (psychologische) theorie van het oplossen van problemen te geven.

BARR, A., & E.A. FEIGENBAUM, *The handbook of artificial intelligence*, Volume I, Los Altos, California, William Kaufman, 1981.

Geeft onder andere een goed overzicht van de algemene technieken voor het aanpakken en oplossen van problemen, zoals die onderzocht of ontwikkeld zijn in Kunstmatige-Intelligentie-onderzoek.

POLYA, G., *How to solve it. A new aspect of mathematical method*, Princeton, New Jersey, Princeton University Press, 1971 (1957).

Een 'klassiek' boek over het aanpakken van problemen, vooral wiskundige problemen.

WICKELGREN, W.A., *How to solve problems, Elements of a theory of problems and problem solving*, San Francisco, Freeman, 1974.

Benadert de aanpak van problemen vooral vanuit de verworvenheden op dit gebied van Kunstmatige Intelligentie.

RUBINSTEIN, M. F., & K. PFEIFFER, *Concepts in problem solving*, Englewood Cliffs, New Jersey, Prentice Hall, 1980.

De aanpak van problemen met behulp van technieken uit de besliskunde en de Operations Research.

VAAGS, D. W., *Over het oplossen van technische problemen*, (proefschrift), T. H. Eindhoven, 1975.

Vooraf leertheoretische (en didactische) aspecten, evenals:

CROMBAG, H. F. M., J. L. DE WIJKERSLOOTH, & E. H. VAN TUYLL VAN SEROOSKERKEN, *Over het oplossen van casusposities*, Groningen, Tjeenk Willink, 1972.

METTES, C. T. C. W., & A. PILOT, *Over het leren oplossen van natuurwetenschappelijke problemen; een methode voor ontwikkeling en evaluatie van onderwijs, toegepast op een cursus thermodynamica*, (proefschrift), Enschede, Onderwijskundig Centrum T.H. Twente, 1980.

DUDA, R. O., & J. G. GASCHNIG, 'Knowledge-based expert systems come of age', *BYTE*, 1981, 238-284.

De computer geprogrammeerd als expert probleemoplosser:

Stand van zaken; vooral medische diagnostiek: de praktische resultaten die Kunstmatige-Intelligentie-onderzoek ook oplevert.

8 Kwaliteit van toetsvragen

Zijn de toetsvragen in concept geschreven, dan is het tijd voor redactioneel bijpolijsten en controle op wenselijke eigenschappen als ondubbelzinnige formulering, dekking van de onderwijsdoelen, mate waarin antwoorden objectief te beoordelen zijn, en dergelijke. Dat hele complex van wenselijke eigenschappen duid ik aan met de term 'kwaliteit'. Wat 'kwaliteit' is en hoe er betekenis en inhoud aan gegeven wordt, is het thema van dit hoofdstuk.

Allereerst bevat de term 'kwaliteit' een normatieve suggestie: 'meer kwaliteit' is 'beter', en wellicht zijn er ondergrenzen voor de verlangde kwaliteit van toetsvragen aan te geven. Daarbij kan men denken aan zoiets als de efficiency van het onderwijs, waar kwalitatief goede toetsvragen het hunne aan kunnen bijdragen. Maar ook aan de rechten en plichten van docenten en studenten tegenover elkaar en wat er vanuit dat meer juridische kader te zeggen valt over wenselijke en vereiste kwaliteit van toetsvragen. Omdat rechtmatigheid dient te prevaleren boven efficiëntie, voor zover redelijk, zal ik beginnen met het bespreken van de eerste (8.1). Daarbij spelen de 'studierechten in het wetenschappelijk onderwijs', zoals Cohen (1981) ze noemt, een bepalende rol. De 'rechten van studenten ten aanzien van het onderwijs dat zij volgen' geven tegelijk de grenzen aan die aan het handelen van de docent gesteld zijn, ofwel aan de verplichtingen die hij heeft tegenover zijn studenten. Dat leidt tot regels voor de procedure, de inhoud en de vorm van examens.

In de volgende paragrafen 8.2 en 8.3 worden aanbevelingen gedaan voor het controleren van de kwaliteit van de ontworpen toetsvragen. Een natuurlijke tweedeling in de kwaliteitscontrole wordt gegeven door de toetsafname: controle *vooraf* moet een redelijke garantie voor de kwaliteit van de toetsvragen opleveren, terwijl controle *achteraf* nodig is om toch in de toets terechtgekomen ondeugdelijke vragen alsnog te kunnen corrigeren of buiten de eindbeoordeling te houden. Veel van de te noemen procedures worden al routinematig gehanteerd. Toch zijn er vaak verbeteringen op mogelijk, bijvoorbeeld door collega's *onafhankelijk* van elkaar de ontwerp vragen te

laten *beantwoorden*, in plaats van ze *plenair* in de vakgroepvergadering te *bespreken*. De algemene strekking van de aanbevelingen is deze: verzamel liever 'harde' empirische gegevens dan af te gaan op opiniërende onderlinge discussie.

8.1 Regels bij het examineren

Wat onder de kwaliteit van toetsvragen kan worden verstaan, volgt in niet geringe mate uit de eisen die studenten daar rechtens aan kunnen stellen. De rechtsregels waar het hier vooral om gaat zijn de in het algemeen rechtsbewustzijn levende *algemene beginselen van behoorlijk bestuur*. Deze, vooral in de administratieve rechtspraak gearticuleerde rechtsregels zijn voor de onderwijsgevende van niet gering belang: ze geven aan tot hoever zijn bevoegdheden tegenover de student reiken. Wilbrink (1977) deed een eerste poging deze algemene beginselen van behoorlijk bestuur te koppelen aan wat een behoorlijke verzorging van het onderwijs en beoordeling van studenten zou moeten zijn. Cohen (1981) heeft dit thema uitgewerkt voor de eisen die aan de verzorging van het onderwijs in al zijn aspecten gesteld kunnen worden. Hij presenteert regels (voorschriften) voor procedure, vorm en inhoud van examens. Deze regels zijn gebaseerd op wettelijke regelingen of afgeleid van algemene beginselen van behoorlijk bestuur. Daarnaast laat Cohen aan de hand van een groot aantal beroepszaken (art. 40 WUB) zien dat deze regels in de praktijk al wel gehanteerd worden door de geschillencommissies. De concrete betekenis van deze regels is dan ook dat schending ervan zal leiden tot beroepsprocedures (ex art. 40), die de docent dwingen alsnog voor behoorlijke kwaliteit van zijn toets(vragen) te zorgen. Dat geeft tevens de urgentie van een goede kwaliteitscontrole aan.

Voor instellingen buiten het wetenschappelijk onderwijs is beroep op examenbeslissingen mogelijk bij de burgerlijke rechter op grond van artikel 1401 van het Burgerlijk Wetboek. Voor openbare onderwijsinstellingen is geen AROB-beroep mogelijk, omdat de wet AROB (Administratieve Rechtspraak Overheids Beschikkingen) op onderwijswetten nog geen beroep openstelt (zie ook Cohen, 1981, blz. 37).

Kan de student de kwaliteit van toetsvragen (van een bepaalde toetsvraag) als beroepsgrond aanvoeren? Cohen (1981, blz. 114): 'Beroep tegen examenvragen zelf en de beoordeling daarvan in het algemeen staat eveneens open; de examenvragen en de beoordelingsnormen vormen de motiveringsgrond. slag van de aangevochten beschikking (d.i.: het gegeven oordeel of cijfer, b.w.)'. Hij wijst op een opmerking van Feldbrugge (in *Trias*, 1973, I, nr. 3, blz. 6) over een beroepszaak in de Leidse Faculteit der Rechtsgeleerdheid:

'... schijnt de commissie impliciet erkend te hebben dat de te toetsen "beoordeling van de examenresultaten" uit art. 40 niet alleen betrekking heeft op de vaststelling van het peil van de antwoorden die een individuele student op een examenvraag heeft gegeven, maar op het gehele beoordelingsproces d.w.z. inclusief de kwaliteit van de examenvragen in het algemeen.'

Tot de kwaliteit van een toetsvraag behoren ook de aanwezigheid van een *modelantwoord* of standaardantwoord en een *beoordelingsvoorschrift*. Bij keuzevragen bestaat het modelantwoord uit het aangegeven juiste (of beste) alternatief en is de scoring gelijk aan de beoordeling: een punt voor een juist aangestreept alternatief. Bij opstelvragen zal een en ander iets minder simpel zijn, maar juist daarbij is het van belang dat er tevoren een modelantwoord wordt opgesteld, aan de hand waarvan antwoorden gescoord kunnen worden op juistheid. Naast deze scoring is het soms noodzakelijk ook de beoordeling in een voorschrift vast te leggen, bijvoorbeeld bij vragen met een ongelijke omvang (gewicht). Cohen gebruikt de term 'beoordelingsvoorschrift' voor scoring en beoordeling tezamen. In het beoordelingsvoorschrift (dat de aanwezigheid van een modelantwoord vooronderstelt) ligt een waarborg voor de gelijke behandeling van studenten, wenselijk op grond van het algemene beginsel van behoorlijk bestuur: 'gelijke gevallen gelijk behandelen, en verschillende naar de maat van hun verschil verschillend', het *gelijkheidsbeginsel*. Ook ligt in modelantwoord en beoordelingsvoorschrift de basis voor een eventuele motivering van de toetsuitslag, en dat is van belang vanuit het *motiveringsbeginsel*. Cohen (1981, blz. 88-89): de examinator moet desgewenst kunnen uitleggen waarom hij juist deze beslissing genomen heeft. 'Het bestaan van een be-

roepsprocedure veronderstelt een motiveringsplicht: als er voor een examenbeslissing geen motivering gegeven hoeft te worden, kan een beroepscommissie de genomen beslissing immers niet beoordelen.' Het motiveringsbeginsel verlangt tevens dat de genomen beslissing gedragen kan worden door de gegeven motivering. Cohen: 'Het examen als zodanig bevat de motivering voor de beslissing; als dat examen verkeerd, want te beperkt van opzet is, betekent dat derhalve dat de motivering voor de beslissing gebrekkig is.'

Hier speelt ook de kwaliteit van de toetsvragen (en de beoordeling van de antwoorden) een rol. De motivering dient feitelijk correct te zijn; zo mogen er geen fouten bij de scoring gemaakt zijn, en dat is alleen na te gaan wanneer er aan de hand van een beoordelingsvoorschrift wordt gewerkt.

Inzagerecht. Cohen (blz. 126): , Na afloop van het examen dient de student de gelegenheid te worden geboden tot (a) inzage in de gestelde vragen, (b) inzage in zijn beoordeelde werk, (c) inzage in de standaardantwoorden en het beoordelingsvoorschrift, (d) bespreking met de examinerator.' Een kwestie van *fair play*, een beginsel van behoorlijk bestuur: de overheid mag haar bevoegdheden niet op dusdanige wijze gebruiken dat de burger daardoor belemmerd wordt in het behartigen van zijn belangen. Het belang van de student is dat hij over de nodige informatie kan beschikken waarop eventueel een beroep gegrond zou moeten worden. Tot de kwaliteit van een toetsvraag hoort dan ook een overtuigend modelantwoord en beoordelingsvoorschrift.

De vragen en opgaven van een examen blijven binnen de duidelijk omschreven omvang van de stof. Cohen (blz. 139): 'De omvang van de stof is een te overzien gebied; alles wat daarbuiten valt, vanzelfsprekend niet. Zou het mogelijk zijn buiten de stof te vragen, dan kan het examen niet aan zijn doel beantwoorden. Het examen is immers de afgeleide van het onderwijs; wil men toetsen in hoeverre dat onderwijs voor een student succes gehad heeft, dan moet men daarover, en niet daarbuiten vragen. De neutrale term "omvang van de stof" heb ik gekozen om duidelijk te maken dat het hierbij niet alleen kan gaan om kenniselementen, maar ook om "inzicht" of "vaardigheden". Tot de 'omvang van de stof' kan alleen

de opgegeven literatuur behoren en bijvoorbeeld niet de inhoud van een niet-verplicht college.

Cohen baseert deze regel op het *kenbaarheidsbeginsel*, onderwijskundigen spreken bij voorkeur over de wenselijkheid van *doorzichtigheid*, terwijl het ook verwant is aan het streven naar *openbaarheid*. Het is geen algemeen beginsel van behoorlijk bestuur, maar is verwant aan of af te leiden van het *fair play* en het *vertrouwensbeginsel*. Het vertrouwensbeginsel: de student moet erop kunnen vertrouwen dat hij tot een voldoende resultaat kan komen wanneer hij afgaat op de informatie die hem door de docent wordt verschaft.

Het examen behoort de onderwijsdoelen naar inhoud en niveau te representeren. In het stadium van het schrijven van toetsvragen betekent dit dat erop gelet moet worden dat alle relevante onderwerpen voorkomen in de leerstofinventarisatie die het uitgangspunt vormt bij het schrijven. Of, wanneer de toets samengesteld wordt door een steekproef te trekken uit een verzameling toetsvragen, dat die verzameling dan ook vragen bevat over alle relevante onderwerpen, eventueel over belangrijker onderwerpen meer vragen dan over minder belangrijke (in proportie zo veel meer als het verschil in belang rechtvaardigt).

De moeilijkheid van de vragen moet overeenstemmen met wat de studenten daarover mogen verwachten. Oefenvragen, een proeftoets, oude tentamenvragen, gegeven voorbeelden en andere informatie: de student moet erop mogen vertrouwen dat de aard van de door hem te beantwoorden toetsvragen daarmee in overeenstemming is (tenzij wijzigingen expliciet zijn meegedeeld).

‘Daarnaast komt het nogal eens voor dat men de examenopgaven net een slag moeilijker maakt dan de opgaven tijdens het onderwijs. De - impliciete! - motivering hiervoor is, dat datgene, wat tijdens het onderwijs geoefend is, triviaal geworden is: iedereen gaat dat goed doen. Wil men dan toch onderscheid maken tussen studenten en examinatoren willen dat ook wanneer daarvoor geen enkele reden is (...), dan moet men het examen moeilijker maken. Het zal duidelijk zijn dat een dergelijke handelwijze volstrekt onjuist is.’ (Cohen, blz. 146).

Toetsvragen zijn duidelijk en ondubbelzinnig. Bij meerkeuzevragen strekt deze regel zich ook tot de alternatieven uit. Bij open vragen vallen modelantwoord en beoordelingsvoorschrift er eveneens onder. Wanneer onduidelijkheden bij de controle vooraf weggewerkt kunnen worden, is dat wel zo prettig. Maar wat te doen wanneer *na afname* van de toets onduidelijkheden blijken? Wanneer voor een meerkeuzevraag ook afleiders als goed alternatief verdedigbaar blijken, dan worden deze afleiders ook als goed antwoord aangemerkt. Wanneer onduidelijkheden of dubbelzinnigheden de studenten in verwarring kunnen hebben gebracht, zal er meestal weinig anders opzitten dan de vraag niet bij het eindoordeel te betrekken. Bijvoorbeeld wanneer geen van de alternatieven een verdedigbaar juist antwoord oplevert. Cohen (blz. 148) maakt een opmerking over meerkeuzevragen met vèrstrekkende betekenis:

‘Wanneer bij een meerkeuzevraag de betrokken examinatoren het onderling niet eens zijn over de vraag welk alternatief juist is, zal men vrijwel altijd moeten aannemen dat de vraag niet aan de geformuleerde regel voldoet.’ Dat betekent dat een dergelijke vraag niet in de toets mag worden opgenomen. Maar bedenk ook dat aannemelijk gemaakt moet kunnen worden dat examinatoren (of andere deskundigen) het onderling eens zijn/waren over het juiste (of beste) alternatief, *onafhankelijk van elkaar*. In 8.2 en 8.3 kom ik hier nog op terug. Voor de beoordeling van antwoorden op aanvulvragen of opstelvragen doet zich veelvuldig de situatie voor dat beoordelaars van elkaar verschillen in de beoordeling van de juistheid van antwoorden en in de waardering van gegeven antwoorden in punten of cijfers. Omdat zich dan de situatie voordoet dat de onderhavige regel geschonden is, terwijl het niet meer mogelijk is de zaak terug te draaien of zo’n vraag helemaal niet mee te rekenen, is er slechts een uitweg: de student het voordeel van de tussen examinatoren bestaande onenigheid geven. Bij verschillende oordelen van meerdere beoordelaars krijgt de student het meest gunstige (voor hem) toegekend. Hoe nu te handelen wanneer achteraf bepaalde vragen niet bij het eindoordeel betrokken worden omdat ze onduidelijk waren? De beste weg lijkt deze te zijn: allereerst de beoordeling uitvoeren *inclusief* de onduidelijke vragen en vervolgens *zonder* deze vragen: de student krijgt het hieruit resulterende

hoogste resultaat als formele uitslag. Wanneer de onduidelijke vragen gesignaleerd worden *voordat* de formele uitslag is vastgesteld en meegedeeld, *kan* deze werkwijze gevolgd worden (bij voorkeur). Worden de onduidelijke vragen gesignaleerd nadat de formele uitslag is meegedeeld, dan *moet* de geschetste werkwijze wel gevolgd worden. De formele uitslag is de geldige uitslag, en daaraan kan alleen iets veranderd worden ten gunste van de student. Wanneer er een informele uitslag voorafgaat aan de formele, bijvoorbeeld door het bekendmaken van de scoringsleutel bij meerkeuzetoetsen, is het wel zo elegant om dan de werkwijze te volgen als was er al een formele uitslag bekendgemaakt. Al met al blijven dit soort wijzigingen achteraf vervelend, ook al omdat de *gelijke behandeling* er niet mee gediend is. Cohen (blz. 133) schrijft naar aanleiding van een beroepszaak over het achteraf verwijderen van toetsvragen uit het eindoordeel:

‘Laten wij eerst vaststellen dat de problemen pas zijn ontstaan nadat er van de zijde van de examinatoren een fout was gemaakt die niet meer te redresseren viel. Is dat het geval, dan is ongelijkheid niet meer te vermijden, maar het voordeel dat sommigen hierdoor behalen, kan niet tevens aan iedereen gegeven worden. Voldoende is dat degenen die het voordeel niet hebben, geen nadeel ondervinden. De polsstok van de beroepscommissie ex art. 40 WUB is geen andere dan die van andere rechters: het gelijkheidsbeginsel in handen van de rechter is in zoverre een beperkt wapen dat het alleen negatief kan werken door een ongelijke behandeling ongedaan te maken, en niet *positief* door iedereen alsnog dezelfde behandeling te geven.’

Gelijke behandeling. Het voorgaande leidt tot de stelregel dat onduidelijke en dubbelzinnige vragen zo veel mogelijk uit de toets geweerd moeten worden door *controle vooraf*. In ieder geval mag niet volstaan worden met alleen een controle achteraf (nabespreking, itemanalyse), omdat maatregelen achteraf de uitwerking van het gelijkheidsbeginsel frustreren.

Samenvattend levert dit voor de kwaliteit van toetsvragen (inclusief modelantwoorden en beoordelingsvoorschriften) het volgende kader op voor controle en bewaking ervan:

- de kwaliteit van de toetsvraag behelst ook de kwaliteit van het bijbehorende modelantwoord (standaardantwoord) en van

het beoordelingsvoorschrift;

- vragen, modelantwoorden en beoordelingsvoorschriften moeten ingezien kunnen worden door studenten die de toets hebben afgelegd;
- vragen gaan niet buiten de voorgeschreven stof; over alle relevante onderdelen van de voorgeschreven stof moet een vraag gesteld *kunnen* worden;
- omvangrijke vragen die opgesplitst kunnen worden in zelfstandige deelvragen, worden opgesplitst wanneer daardoor een betere representatie van de stof in de toets verkregen wordt;
- aard, soort en moeilijkheid van de vragen stemt overeen met wat de student mag verwachten en waarop hij zich heeft kunnen voorbereiden;
- toetsvragen zijn duidelijk en ondubbelzinnig, evenals modelantwoorden en beoordelingsvoorschriften;
- controle voorafgaand aan de toetsing heeft prioriteit boven controle achteraf;
- zowel de controle van de vragen, als het beoordelen van de antwoorden van de studenten, gebeurt zo objectief mogelijk (intersubjectieve overeenstemming is het ideaal) .

8.2 Punten om op te controleren

Gebrekkige kwaliteit van een ontworpen toetsvraag is vaak ook het gevolg van een 'verborgen gebrek'. Er is een instrument nodig om de toetsvragen mee te lijf te gaan (en een methode, zie 8.3). Een afstreeplijst is in ieder geval een goed hulpmiddel; zij kan de zekerheid verschaffen dat geen punten over het hoofd worden gezien. De te presenteren lijsten kunnen natuurlijk worden aangevuld met specifieke vragen die in uw speciale geval van belang zijn. De lijsten kunnen bij de controle gebruikt worden door zowel de toetsvragenschrijver als door andere examinatoren en mogelijk ook door studenten (na afloop van de toets of zelfs tijdens de toetsafname).

Lijst van aanwijzingen voor het formuleren

Tenzij anders aangegeven wordt met 'meerkeuzevraag' alleen de stam van de vraag bedoeld, dus *zonder* de bijbehorende alternatieven.

Bij de controle van meerkeuzevragen is iedere vraag in twee gedaanten beschikbaar: zowel *zonder* als *met* alternatieven.

1. Is de formulering van de vraag *eenvoudig*?

(Vermijd zoveel mogelijk ingewikkelde zinsconstructies, onbekende termen, moeilijke woorden.)

2. Bevat de vraag geen *dubbele ontkenningen*?

(Dat is niet alleen nodeloos ingewikkeld, maar het leidt ook tot dubbelzinnigheid; voor keuzevragen moet de combinatie van een ontkenning in de stam en ontkenningen in alternatieven vermeden worden; wordt bij keuzevragen een ontkenning gebruikt, zorg er dan voor dat de aandacht van de student er op gevestigd wordt door onderstrepen, andere kleur, of toevoegen van de waarschuwing 'let op de ontkenning'.)

3. Is de vraag op verschillende manieren *uit te leggen*?

(Speel hier de advocaat van de duivel, analyseer verschillende mogelijke betekenissen die aan termen in de vraagstelling toegekend zouden kunnen worden. Anders gezegd:

Is de vraag *ondubbelzinnig* geformuleerd? Voor meerkeuzevragen *beide* vormen bekijken: eerst zonder alternatieven en daarna met alternatieven.)

4. Is de vraag ook *ondubbelzinnig* voor studenten die andere literatuur hebben gelezen dan voor de toets is opgegeven?

5. Zijn *universele quantoren* (nooit, alle, zeker, altijd) terecht gebruikt?

(Ook in alternatieven bij meerkeuzevragen.)

6. Is de formulering *specifiek*?

(Wanneer de algemene instructie voor de toets daar geen duidelijkheid over geeft, is het uit de formulering van de vraag dan duidelijk hoe uitgebreid het antwoord moet zijn? Maar ook: Kan de goed voorbereide student weten welk antwoord als een goed antwoord gerekend zal worden?)

7. Is het uit de vraagformulering duidelijk wat de vraag toetst?

(Het is wenselijk dat ook de student die de vraag niet kan be-

antwoorden toch weet waarop hij het antwoord schuldig moest blijven, welke kennis of inzichten hem ontbreken.)

8. Is de vraag *op meerdere manieren goed te beantwoorden*? (Het gaat erom dat ingeval er meerdere goede antwoorden mogelijk zijn, daar bij de scoring rekening mee gehouden wordt: slechts wanneer er onbedoelde antwoordmogelijkheden zijn die toch niet fout gerekend kunnen worden, zal er naar een betere formulering van de vraag gezocht kunnen worden.)

9. Wanneer het te geven antwoord *gemotiveerd* moet worden, is om die motivering dan ook expliciet gevraagd?
(Vooral van belang bij vragen die anders met een eenvoudig 'ja' of 'nee' te beantwoorden zijn.)

10. Geen twee uitspraken tegelijk op juistheid laten beoordelen.
(Dat schept een dubbelzinnige situatie; wel mogelijk: een hoofduitspraak die altijd waar is en de 'reden' die juist of onjuist kan zijn.)

11. Is de formulering niet te ingewikkeld?
(Het is verleidelijk om een slechte meerkeuzevraag te verbeteren door via een scherpere formulering van de stam of van de alternatieven de afleiders ook ondubbelzinnig 'fout' te maken. Het risico is dan dat dergelijke 'verbeterde' meerkeuzevragen eerder zoiets als verbaal begrip van de student meten dan zijn of haar beheersing van het onderwerp van de vraag.)

12. Voor de meerkeuzevraag: is de *stam alleen* een goede aanvulvraag?
(Zo ja, is er een goede reden om er dan toch een keuzevraag van te maken? Het mag niet voorkomen dat de student eerst een studie van de aangeboden alternatieven moet maken, om er achter te komen wat er eigenlijk wordt gevraagd. Het kan wel geoorloofd zijn de aangeboden alternatieven te gebruiken als een verdere definitie van de vraagstelling in die gevallen waarin de stam van de vraag te uitgebreid zou worden wanneer daarin de aard van het verlangde antwoord precies duidelijk gemaakt moet worden.)

13. Is de meerkeuzevraag *te vervangen* door een aanvulvraag?
(Zo ja, is er een goede reden om dat niet te doen?)

14. Bij de *alternatieven* van een meerkeuzevraag:

- a. Zijn de alternatieven zo kort mogelijk geformuleerd?
- b. Komen er in de alternatieven woorden voor die beter in de stam opgenomen kunnen worden?
- c. Sluit de formulering van ieder alternatief aan op de stam?
- d. Zijn dummy-alternatieven (zoals 'geen van deze') te vermijden?
- e. Is er een afleider die als goed antwoord verdedigbaar is?
- f. Is het juiste antwoord gelijk aan het antwoord dat andere examinatoren op de stam (als aanvulvraag dus) gegeven hebben?
(Is dat niet het geval, dan is de meerkeuzevraag mogelijk verwarrend voor de studenten.)

15. *Onbedoelde aanwijzingen* naar afleiders of juiste alternatief.

Het gaat hier om verborgen hints die de toetsslimme student in staat stellen het juiste alternatief op te sporen of ten minste enkele afleiders te ontmaskeren, zonder daar kennis van het gevraagde onderwerp voor nodig te hebben. Onder andere:

- a. Springt het juiste alternatief er in lengte uit?
(Veel korter of langer dan de afleiders?)
- b. Wordt een belangrijke term uit de stam in een van de alternatieven herhaald?
(Doorgaans is dat het juiste alternatief.)
- c. Bevatten de afleiders terminologie die de student onbekend is?
(De student kan dat als teken opvatten dat het een afleider is, maar kan er ook door in verwarring worden gebracht.)
- d. Is de volgorde van de alternatieven random bepaald?
(Met dobbelsteen bijv.; wanneer het juiste alternatief relatief vaak op bijv. de tweede plaats staat, kan de slimme student dat uitbuiten door bij raden altijd nr. 2 te kiezen.)
- e. Zijn afleiders geconstrueerd door er een logische onjuistheid in te verwerken?
(Logisch redeneren behoort zelden tot de onderwijsdoelen.)
- f. Zijn universele quantoren wel terecht gebruikt?

(Bijv. *alle, altijd, nooit, zeker*. 'Zeg nooit nooit': universele uitspraken zijn meestal niet waar, ten tegen een dat een universele uitspraak een afleider is; verwant met punt e.)

g. Kan het juiste alternatief op grond van een onbedoeld kenmerk gematcht worden met de vraagstelling?

Lans en Mellenbergh, in De Groot en Van Naerssen(1973), geven van deze onbedoelde hint dit fraaie voorbeeld:

AFRADER

Wie van de volgende vier filosofen hanteert het onderscheid tussen 'être-en-soi' en 'être-pour-soi'?

1. Heidegger
2. Jaspers
3. Kierkegaard
4. Sartre

(Streef de Fransman aan, dat zal wel goed zijn,)

16. *Ja/nee*-vragen: komen beide alternatieven ongeveer even vaak voor en is dat bereikt zonder te vervallen in het gebruik van ontkenningen ?

(Er is helaas een algemene neiging om dergelijke vragen vaker met 'ja' dan met 'nee' te beantwoorden, ook wanneer beide alternatieven even vaak juist zijn gesleuteld. Sommige studenten hebben die neiging sterker dan andere. Probeer liever niet met deze neiging rekening te houden. Het is een nadeel dat deze vraagvorm nu eenmaal aankleeft.)

17. Bij de *instructie* zoals die bij de toetsafname verstrekt wordt: is deze eenduidig geformuleerd en dekt ze alle vraagvormen die in de toets gebruikt zullen worden?

Puntenlijst voor inhoudelijke aspecten

1. Over welk onderwerp handelt de vraag?

2. Welke vorm van beheersing wordt gevraagd?

(Reproductie, nieuw voorbeeld benoemen, nieuwe toepassing, inferentie, vertaling enz.)

3. Is het abstractieniveau van de vraag hoger gekozen dan wenselijk is?

(Is daar een goede reden voor?)

4. Behoort het gevraagde tot de opgegeven stof?

5. Kan de student weten dat zo'n vraag in de toets kan voorkomen?
(Is zo'n soort vraag in het onderwijs besproken, is erop geoefend enz.)

6. Behoort het gevraagde tot de onderwijsdoelen?

7. Is het een strikvraag?

(Algemeen: wordt er voor een goede beantwoording aanspraak gedaan op vaardigheden of inzichten die niet als zodanig geoefend zijn, die niet voor te bereiden zijn, en waar studenten dus van mogen verwachten dat ze in de toets geen rol spelen?)

8. Is het een opinievraag?

(Opinie vragen horen in een eindtoets in de regel niet thuis, omdat een opinie niet als 'goed' of 'fout' kan worden aangemerkt en daardoor niet beoordeelbaar is)

9. Wordt in de vraag een letterlijk uit de leerstof overgenomen formulering gebruikt?

(Is daar een goede reden voor?)

10. Past de vraagvorm op wat bedoeld wordt te vragen?

(Stelregel: gebruik de meerkeuze-, aanvul- of opstelvraagvorm die op ongeforceerde wijze past bij wat gevraagd wordt.)

11. Is de vraag op te splitsen in zelfstandige deelvragen?

(Zo ja, is er een goede reden om de vraag *niet* op te splitsen?).

12. *Geef een modelantwoord* zonder daar bronnen voor te raadplegen waar de student bij de toetsing geen toegang toe heeft; neem de tijd op die daarvoor nodig is.

13. Welke *karakteristieke fouten* zijn te verwachten?

14. Hoe vaak zal ieder van deze te verwachten fouten voorkomen in de groep studenten die aan de toets deelneemt?

15. Geef aan hoe antwoorden aan de hand van het gegeven modelantwoord (12. hierboven) gescoord zouden moeten worden, d.w.z. op *juistheid* beoordeeld; betrek daar desgewenst ook de te verwachten fouten (13. hierboven) bij.

16. Geef bij de scoring van 15. aan hoe deze *gewaardeerd* zou moeten worden in punten of gewichten op ieder van de afzonderlijk gescoorde onderdelen.

17. *Specifiek voor meerkeuzevragen* komen hier bij:

a. Behoort het kunnen *onderscheiden* tussen het juiste antwoord en de afleiders tot de doelstellingen?

b. Wat is het juiste (of beste) antwoord?

c. Stemt het juiste antwoord overeen met het eerder gegeven antwoord (12. hierboven) bij alleen de stam van de meerkeuzevraag?

d. Orden de afleiders in volgorde van 'moeilijkheid'

e. Zijn er een of meer afleiders die waarschijnlijk iedereen wel weet te vermijden?

f. Hoe moeilijk is het om het juiste antwoord te onderscheiden van de afleider die er het minst van verschilt? g. Hoe vaak zal het juiste antwoord gegeven worden door de groep studenten die aan de toets deelneemt?

(Ga er vanuit dat er niet geraden wordt; de vraag is dus:

hoeveel studenten *weten* het juiste antwoord?)

Algemene vragen

1. Is ervoor gezorgd dat over ieder relevant onderwerp (uit de onderwijsdoelen of zoals genoemd in de inventarisatie) vragen in de toets opgenomen *kunnen* worden?

2. Is ervoor gezorgd dat over belangrijker onderwerpen ook vaker een toetsvraag te verwachten is (in verhouding tot de mate van belangrijkheid)?

(Bijvoorbeeld door bij de lijst onderwerpen een kwantificering te maken van ofwel de *aantallen* vragen die erover in de toets gesteld zullen worden, ofwel de relatieve *waarschijnlijk-*

heid dat er vragen over in de toets gesteld zullen worden.)

3. Is ervoor gezorgd dat er ook vragen gesteld zullen worden over onderwerpen die door de studenten zeer goed beheerst worden? (Ofwel: is ervoor gezorgd dat in deze zin 'gemakkelijke' vragen, die tot de kern van de stof behoren, niet ten onrechte ongeformuleerd zijn gebleven of buiten de toets gehouden?)

4. Is ervoor gezorgd dat de uiteindelijke samenstelling van de toets overeenstemt met wat studenten redelijkerwijs kunnen en mogen verwachten wat betreft de moeilijkheid van de vragen, het soort stofbeheersing dat gevraagd wordt en de tijd die voor volledige beantwoording nodig is?

(Analoog aan de representatieve verdeling in 2. hierboven.)

Al met al lijkt dit een hoop extra werk. Enig extra werk zal altijd wel met controle op de vraagkwaliteit gemoeid zijn, maar zo erg als de eerste aanblik van deze lijsten doet vermoeden, zal het in de praktijk niet zijn. Om te beginnen moet uit deze lijsten een verstandige en werkbare keuze worden gedaan, in overeenstemming met de specifieke eisen die aan die ene toets gesteld worden. Dan zal blijken dat het met enige oefening aan de hand van zo'n kortere lijst al snel mogelijk is om vragen te controleren zonder telkens één oog op de lijst te houden. Bedenk dat niet alle ontworpen vragen een even intensieve controle vooraf nodig zullen hebben. Wanneer 'oude' toetsvragen gebruikt worden, die hun kwaliteit al eerder 'bewezen' hebben, dan is het voldoende om na te gaan of zij nog steeds in overeenstemming zijn met de inmiddels mogelijk gewijzigde leerstof. Voor vragen die verkregen zijn uit 'oude' vragen door bijvoorbeeld een nieuw voorbeeld te gebruiken, is een marginale controle voldoende.

8.3 Onafhankelijke beoordeling van kwaliteit

Hoe zijn de lijsten uit 8.2 te gebruiken voor controle op de kwaliteit van de ontworpen toetsvragen? Het is niet voldoende

wanneer de schrijver van de toetsvragen de lijsten naloopt en afgaande op eigen impressies en ervaringen hier en daar een vraag 'verbetert' of de prullenbak in kiepert. Dat is te vrijblijvend en geeft geen behoorlijke garanties tegenover andere examinatoren en tegenover de studenten dat de vragen nu wel voldoen aan redelijke kwaliteitseisen. Een goede controleprocedure voldoet aan de eis dat eventuele mankementen in (bijna) alle gevallen boven water komen. De sleutel tot zo'n procedure ligt in de aard van de verborgen mankementen: dat zijn vooral die kenmerken van een toetsvraag waarover examinatoren het onderling oneens zijn. Welnu, dan kan men een procedure gebruiken, waarbij die verschillen van inzicht *noodzakelijkerwijs* te voorschijn komen. Het principe is: laat twee of meer examinatoren geheel *onafhankelijk* van elkaar de controle aan de hand van de lijsten uitvoeren. Een van die controlerende examinatoren kan de schrijver van de betreffende toetsvraag zijn; laat deze al bij het schrijven van de vraag een *protocol* opstellen met alle relevante gegevens over die vraag, conform de controlelijsten. Is dat nou echt nodig, zo'n hoop gedoe over toetsvragen waarvan iedere deskundige toch onmiddellijk kan inzien dat ze ondubbelzinnig geformuleerd en al even ondubbelzinnig te beantwoorden zijn? Ja, dat is nodig, want het blijkt eenvoudig niet zo te zijn dat men hierbij op het eigen uiterst stellige, maar toch altijd nog ongewapende oordeel af kan gaan. Ik geef ter illustratie een eenvoudige vraag die zou kunnen voorkomen in een toets aan het eind van een cursus inleiding in de statistiek.

VOORBEELD

Jan Molenaar is de trotse vader van twee kinderen. We komen hem op straat tegen met een jongen die hij voorstelt als zijn zoon. Hoe waarschijnlijk is het dat het andere kind van Molenaar ook een jongen is?

Beantwoord bovenstaande vraag, en motiveer dat antwoord.

Doe een klein onderzoekje door de vraag ook te laten beantwoorden door uw collega's die statistiek doceren. Zo'n onderzoekje werd bij wijze van grap gedaan door een van de auteurs Bar-Hillel en Falk (1982), die de vraag voorlegde aan twee hoogleraren wiskunde. Zij produceerden onmiddellijk en met grote vanzelfsprekendheid het antwoord.

De ene hoogleraar antwoordde dat die waarschijnlijkheid een half was, de ander antwoordde even stellig dat het een derde moest zijn. Goed, dit is een anekdote. Maar we zitten ondanks dat nog wel met het probleem wat dan wel het juiste antwoord is. Bar-Hillel en Falk (1982) hebben daar een boeiende en leerzame uiteenzetting over gegeven, aan de hand van nog enkele andere dubbelzinnige opgaven. Het komt erop neer zoals in de statistiek wel vaker het geval is, dat de opgave in het geheel niet zo eenvoudig is als hij eruitziet en zoals hij door de twee hoogleraren beantwoord werd. De eerste hoogleraar redeneerde dat de waarschijnlijkheid voor de geboorte van een jongen ongeveer gelijk is aan die voor de geboorte van een meisje, beide zijn onafhankelijk van elkaar, dus het antwoord moet een half zijn. De tweede hoogleraar redeneerde dat er drie even waarschijnlijke mogelijkheden zijn: Molenaar heeft ofwel twee jongens, ofwel een jongen als oudste en een meisje als jongste kind, ofwel een meisje als oudste en een jongen als jongste kind, zodat de waarschijnlijkheid dat hij twee jongens heeft een derde moet zijn. Nu is het met wiskundige opgaven als deze noodzakelijkerwijs zo dat er maar een juist antwoord kan zijn, dus een van beiden gaf een onjuist antwoord of beiden gaven een onjuist antwoord. Zorgvuldige analyse, en dat vraagt nogal wat werk zoals tabel 1 in Bar-Hillel en Falk (1982) laat zien, levert als juiste antwoord een half op, maar de motivering van dit antwoord is een andere dan onze eerste hoogleraar had gegeven. De les die hieruit getrokken kan worden is deze: simpel uitziende problemen kunnen in werkelijkheid complex zijn, en een signaal daarvoor kan gevonden worden in verschillende antwoorden van 'deskundigen', zoals die bij een eenvoudig onderzoekje kunnen blijken.

De controle kent een aantal varianten. Allereerst kan de schrijver zichzelf controleren. Een normale procedure is dat een of meer collega's, andere examinatoren, de vragen controleren. Dit zijn de controles vooraf en deze zijn aan te vullen met achteraf te verkrijgen gegevens: de toetsresultaten en eventueel het commentaar van de studenten.

De schrijver controleert zichzelf. Het idee hierachter is: de schrijver moet in staat zijn om na een tussentijd van enkele

weken de te controleren vraag te beantwoorden op dezelfde wijze als op het protocol aangegeven, uiteraard zonder eerst naar het protocol gekeken te hebben. Wanneer antwoorden en protocol op wezenlijke punten verschillen, dan is de vraag ondeugdelijk. Want ofwel blijkt de vraag nu anders uitgelegd te worden dan bij het schrijven de bedoeling was, ofwel er blijkt nu een 'goed' antwoord geproduceerd te kunnen worden dat bij het schrijven van de vraag niet was voorzien. De eerste voorwaarde voor zelfcontrole is dat er zo veel tijd is verlopen tussen het schrijven van de vraag en de controle dat het aannemelijk is dat de schrijver niet uit zijn geheugen hetzelfde antwoord als destijds zal produceren. Andere voorwaarden zijn: er is een volledig protocol beschikbaar, opgemaakt bij het schrijven van de vraag; de vragen worden in toevallige volgorde gecontroleerd; er staan geen onbedoelde aanwijzingen in het te controleren materiaal; meerkeuzevragen worden eerst gecontroleerd zonder de alternatieven, daarna met de alternatieven. De te controleren vragen zijn niet geordend naar onderwerpen uit de literatuur en dragen geen nummer of code die een verwijzing naar de bron in de literatuur kan bevatten. De resultaten van een zelfcontrole zijn bijna even bruikbaar als die van een controle door een collega. Ieder verschil in beantwoording met het oorspronkelijke protocol duidt op een tekortkoming in de vraag of tenminste op een niet eerder onderkende eigenschap, antwoordmogelijkheid, en dergelijke.

De zelfcontrole heeft één overwegend nadeel: de controleur weet dat hij zijn eigen vragen aan het controleren is en kan onbedoeld tot een minder streng oordeel komen dan de objectiviteit verlangt.

Controle door collega's (andere examinatoren). Het is algemeen gebruikelijk ontworpen toetsvragen te bespreken met collega's. Het valt nog te bezien of 'bespreken' wel een voldoende controle is. Bij een gezamenlijke bespreking is een *onafhankelijk* oordeel niet verzekerd. Het kan zijn dat bij een te bespreken *vraag* ook het *antwoord* al gegeven is, en dan vervalt voor een belangrijk deel de mogelijkheid dat collega's met *andere* antwoorden voor de dag komen, daarmee een mogelijke dubbelzinnigheid in de vraagformulering aantonend. Bij een bespreking is het bovendien niet te vermijden dat *opi-*

nies worden uitgesproken, en door anderen ondersteund worden over de juistheid van een bepaalde vraag. Dat is geen goede basis voor een kwaliteitsoordeel.

Beter is het om *empirische gegevens* te verzamelen over verschillen van inzicht tussen onafhankelijk van elkaar werkende examinatoren. Vraag een of meer collega's om de vragen te controleren op dezelfde manier als de schrijver bij de zelfcontrole doet: de lijsten schriftelijk doorwerken zonder ruggespraak met anderen te voeren, en zonder stukken of literatuur te raadplegen waar studenten tijdens het tentamen ook geen toegang toe hebben. Zo worden empirische gegevens verzameld, die te vergelijken zijn met het protocol dat bij het schrijven van de toetsvraag al is opgesteld of die voor meerdere beoordelaars *onderling* vergeleken kunnen worden. Bij deze procedure is er weinig reden meer voor bespreking in de vakgroepvergadering, omdat de gegevens nu sterk voor zichzelf spreken. Immers, waar examinatoren op essentiële punten met elkaar verschillen, andere antwoorden op de vraag geven, een ander alternatief van de meerkeuzevraag als juist aanwijzen, verschillend denken over welk onderwijsdoel de vraag behelst, dan moet de vraag beter geformuleerd of vernietigd worden. Dan valt er niet te onderhandelen. Bijvoorbeeld wat betreft het juiste alternatief bij een meerkeuzevraag: wanneer tenminste één examinerer een als afleider bedoeld alternatief als het juiste heeft aangestreept, dan kan geen onderlinge discussie dat feit verhelen dat ten minste één deskundige deze bepaalde afleider als verdedigbaar het juiste antwoord heeft beschouwd.

Alleen wanneer in dit voorbeeld de 'dissident' ervan overtuigd kan worden dat hij gewoon een fout heeft gemaakt, een fout die ook een student als fout aangerekend dient te worden, dan blijft de betreffende meerkeuzevraag aanvaardbaar. Maar het mag nimmer zo zijn dat in onderlinge discussie een dissidente collega zegt: 'ik zie wel in dat het als goed bedoelde alternatief een goed antwoord is en ik ben bereid om me aan te sluiten bij de keuze van dat alternatief als het juiste alternatief'. Dat levert pseudo-objectiviteit op, in de Amerikaanse literatuur ook wel 'frozen subjectivity' genoemd. Het is evident dat bij zo'n handelwijze al die studenten gedupeerd worden die een antwoord geven dat door ten minste één deskundige ook als antwoord gegeven was, maar door een 'meer-

derheid' van deskundigen werd afgewezen. Het zou de student plaatsen in de onmogelijke positie bij de afweging van zijn antwoorden rekening te houden met wat waarschijnlijk een meerderheidsopvatting onder de examinatoren zal zijn. Om nog maar niet te spreken over de twijfelachtige wetenschapsopvatting die spreekt uit het op deze wijze, bij onderling fiat of meerderheidsopvatting, beslissen wat 'waar' is. Er is geen andere begaanbare weg dan collega's onafhankelijk van elkaar de vragen schriftelijk te laten beoordelen. Desnoods op een sterk gekortwiekte lijst, waarbij in ieder geval gevraagd wordt *om beantwoording* van de te controleren vragen.

Hoe te handelen met de uitkomsten van deze controle? Richtinggevend is daarvoor de rechtmatigheid van de beoordeling van de student (zie 8.1). Dat betekent voor meerkeuzevragen dat het juiste alternatief niet alleen consistent door andere examinatoren ook als het juiste wordt aangemerkt, maar bovendien dat zij als antwoord op de *stam* van de vraag, dus zonder te hebben gezien wat de alternatieven zijn, ook als antwoord geven wat in het juiste alternatief geformuleerd is.

Voor aanvulvragen geldt dat niet in zo sterke mate: afwijkende antwoorden van verschillende examinatoren kunnen wijzen op een laakbare dubbelzinnigheid in de formulering van de vraag; maar het kan ook zijn dat alle antwoorden goed zijn, en dat bij de controle een aantal mogelijk goede antwoorden boven water zijn gekomen waar de schrijver van de vraag nog niet aan had gedacht. Hetzelfde doet zich voor bij opstelvragen en het modelantwoord bij de opstelvraag. Wat de beoordeling betreft: wanneer daarover de opvattingen verschillen, worden de verschillen uitgelegd ten gunste van de student. De student krijgt altijd het hoogste oordeel dat door een examiner gegeven wordt. Zijn er erg grote verschillen in opvatting, dan verdient het aanbeveling de betreffende vraag scherper te formuleren of helemaal weg te laten.

Versillen tussen examinatoren kunnen niet alleen bij controle vooraf blijken, maar ook bij het nakijken van gemaakt werk (wanneer ieder antwoord door ten minste twee examinatoren onafhankelijk van elkaar beoordeeld wordt).

Controle door collega's kost tijd. De methode die de meeste tijd vraagt is de onderlinge bespreking. De hier voorgestelde procedure van onafhankelijke schriftelijke controle kost min-

der tijd en levert veel beter interpreteerbare resultaten op (ook voor verantwoording naar de studenten toe).

Controle door studenten. Studenten worden pas bij de toetsafname met de vragen geconfronteerd. Dat wil nog niet zeggen dat controle door studenten louter controle achteraf zou zijn. Immers, fouten en dubbelzinnigheden in de vragen die tijdens de toetsing worden ontdekt, kunnen op dat moment nog rechtgezet worden. Studenten moeten met zo'n ontdekking (via een surveillant) een verantwoordelijke examinerator kunnen bereiken. Het geven van verduidelijking kan op dat moment alle deelnemende studenten nog ten goede komen. Het belang daarvan is dat voorkomen wordt dat anders door voldongen feiten een ongelijke behandeling van studenten ontstaat. Denk niet dat een uitvoerige controle vooraf alle mogelijke fouten in vraagformuleringen er uithaalt. Voor het overige valt commentaar van studenten onder de rubriek 'controle achteraf'. Allereerst kan gevraagd worden om commentaar op bepaalde vragen op het antwoordformulier bij te schrijven. Studenten worden daar ook expliciet toe uitgenodigd door de belofte dat deze commentaren bestudeerd zullen worden op eventuele consequenties voor de beoordeling. Wie van deze mogelijkheid wat meer werk wil maken, kan ook de studenten een vragenlijstje geven om ze een idee te geven welk soort commentaar op prijs gesteld zou worden (wanneer bepaalde vragen daar aanleiding toe geven). Bijvoorbeeld:- Is een vraag voor tweërlei uitleg vatbaar?

- Zijn meerdere alternatieven als juist te verdedigen?
- Wordt kennis verondersteld die buiten de opgegeven stof valt?
- Had je mogen verwachten dat een bepaald soort vraag niet in de toets zou voorkomen?
- Is onvoldoende duidelijk welk antwoord als correct zal gelden (hoe uitgebreid, al dan geen details, aantal te vermelden tussenstappen)?

Direct na de toetsafname krijgen studenten de beschikking over de juiste antwoorden (of de scoringsleutel voor keuzevragen), en uiteraard mogen ze de vragen behouden (zie 2.4). Op dat moment kunnen studenten tekortkomingen in de vragen, de als juist beschouwde antwoorden of in de scorings-

sleutel ontdekken. Geef dan ook gelegenheid om opmerkingen daarover op het bord van de examinatoren te leggen. Hetzelfde geldt voor de nabespreking van het tentamen (die bij voorkeur zo snel mogelijk te houden is, direct na afloop van de toets is beter dan na enkele weken wanneer al het werk is nagekeken).

Een derde mogelijkheid voor het inwinnen van commentaar van studenten is een intensief doorspreken van de hele toets met een klein groepje studenten (vrijwilligers of een afvaardiging). Bij die bespreking kan dan ook alle materiaal uit de voorbereidingsfase op tafel komen, zoals de protocollen van de vragenschrijvers, modelantwoorden en beoordelingsvoorschriften.

Voor zover commentaar van studenten leidt tot bepaalde maatregelen bij de beoordeling of tot het verwijderen van bepaalde vragen, wordt dat algemeen bekendgemaakt.

Controle achteraf gebeurt door gebruik te maken van de resultaten van de toets, uitgesplitst naar de afzonderlijke vragen. Bij opstelvragen is het van belang de procedure voor het nakijken zo in te richten dat eventuele discrepanties tussen verschillende beoordelaars ook *blijken*. Bijvoorbeeld: ieder antwoord wordt door ten minste twee examinatoren gescoord en gewaardeerd, waarbij geen aantekeningen op het werk van de student gemaakt mogen worden (om het werk van de tweede beoordelaar niet te beïnvloeden). De controle bestaat er dan uit dat gelet wordt op onverwachte antwoorden die 'goed' zijn (die worden alsnog in het modelantwoord en het beoordelingsvoorschrift opgenomen, wanneer meerdere studenten dat antwoord geven). Beoordelaars houden aantekening van alle problemen die ze ervaren in het werken met modelantwoorden en beoordelingsvoorschriften; deze problemen worden besproken en bekeken op mogelijke consequenties voor de beoordeling.

Wanneer alle vragen zijn nagekeken, is ook bekend (of eenvoudig na te gaan) hoe *moeilijk* de afzonderlijke vragen voor deze groep studenten zijn geweest: hoeveel studenten een bepaalde vraag goed hebben weten te beantwoorden. Vergelijk dit nieuwe gegeven met de *schattingen* over de moeilijkheid zoals die bij de controle vooraf zijn gemaakt. Wanneer daar nogal forse verschillen tussen liggen, dan is het zaak om te

analyseren waarin de onvoorziene moeilijkheid van de betreffende vraag schuilt. Bij het bepalen van het verschil tussen verkregen en voorspelde moeilijkheid kan rekening worden gehouden met over-all-verschil in verkregen en voorspelde moeilijkheid. Anders gezegd: het is niet onwaarschijnlijk dat de moeilijkheid van de vragen in het algemeen *onderschat* wordt; dan kijkt men alleen wat nauwkeuriger naar die vragen die in vergelijking tot andere vragen wel erg sterk onderschat zijn. Deze nadere analyse zal doorgaans niet leiden tot het achteraf verwijderen van vragen, tenzij blijkt dat de vraag buiten de opgegeven stof is gegaan.

Bij *meerkeuzetoetsen* gebeurt de analyse van de moeilijkheid van de vragen op dezelfde wijze als hierboven geschetst. De moeilijkheid van een vraag wordt gegeven door de p-waarde op de computeroutput, dat is de proportie studenten die de vraag ofwel goed wisten te beantwoorden, ofwel goed geraden hebben.

De tevoren opgegeven verwachting is uitgedrukt in de proportie studenten die het antwoord zouden *weten*, dus om beide gegevens vergelijkbaar te maken moeten de p-waarden gecorrigeerd worden voor raden. Ga ervan uit dat de raadkans de inverse is van het aantal alternatieven: zijn er vier alternatieven, dan is de raadkans te stellen op $1/4$. Dat is een onderschatting van de werkelijke (maar onbekende) raadkans, omdat studenten doorgaans wel over partiële kennis beschikken op grond waarvan ze een of twee afleiders als afleider onderkend hebben. Voor het doel van de controle achteraf is het werken met een voorzichtige raadkans aanvaardbaar. Gebruik de raadkans dan om de gemaakte schatting op te waarden: geschat dat $6/10$ van de studenten het antwoord weet, levert op dat $6/10 + (1/4) * (4/10) = 7/10$ het juiste alternatief zou aanstrepen. Vergelijk deze laatste waarde met de p-waarde van de computeroutput.

De computeroutput van de meerkeuzetoets levert nog een ander gegeven over de afzonderlijke toetsvragen op: de r-bis, ook wel geschreven als r-it of RIT.

De r-bis duidt aan in welke mate studenten die een bepaalde vraag juist beantwoorden ook tot degenen behoren die op de toets in zijn geheel hoog scoren. Het is een heel lastig of in het geheel niet te interpreteren gegeven. Voor de controle achteraf heeft de r-bis dan ook weinig waarde. De enige functie die

het kan hebben is als signaal voor een tamelijk grove fout in de vraagformulering of in de scoringsleutel: dan kan er wel eens een sterk *negatieve* r-bis gevonden worden. Controleer in zo'n geval of de opgegeven scoringsleutel wel juist is, ook in samenhang met een eventueel onjuiste formulering van de toetsvraag. Maar pas op: het kan best zijn dat er niets bijzonders met de vraag of met de scoringsleutel aan de hand is en dan kan de vraag gewoon gehandhaafd blijven. Er is met andere woorden geen rechtvaardiging voor een automatisme dat toetsvragen verwijdert wanneer er een negatieve r-bis geconstateerd wordt. Er zijn computerprogramma's die zo'n automatische verwijdering als optie kennen; maak nimmer van dergelijke opties gebruik (ook niet wat bepaalde grenzen voor de moeilijkheid of p-waarde van de vragen betreft). Omdat achteraf verwijderen van toetsvragen onbillijkheden met zich mee kan brengen, moet daar voor iedere verwijderde vraag een verantwoording voor gegeven kunnen worden, en het is beslist niet voldoende om te verwijzen naar een computerprogramma-optie die de vragen 'automatisch' verwijderde.

Wellicht ten overvloede: r-bis-waarden in de buurt van 0,0 (zowel iets daarboven als iets daaronder) kunnen perfect aanvaardbaar zijn. Wanneer alle studenten de stof op een niet al te zeer uiteenlopend niveau beheersen, mag je zelfs verwachten dat het merendeel van de r-bis-waarden in de buurt van 0,0 terecht zal komen.

Er is een bijzondere reden om in bepaalde gevallen juist argwanend naar vragen met een *hoge* r-bis te kijken. Anderson (1972): 'Studenten die het op de toets in zijn geheel goed doen zullen een hogere verbale intelligentie ('ability') hebben dan zij die het slechter doen. Vragen die een hoge r-bis hebben, die dus onderscheiden tussen beide groepen studenten, zullen vaker moeilijk vocabulaire bevatten of denkstappen vragen die niet direct te maken hebben met de eigenlijk te toetsen leerstofbeheersing.' Anderson knoopt daar ook de waarschuwing aan vast geen vragen met lage r-bis-waarden uit de toets te verwijderen: dan houd je vragen met relatief hoge r-bis-waarden over, met de door hem genoemde eigenschappen die strijdig kunnen zijn met de doelen van het onderwijs en de toetsing.

8.4 Controlelijsten in beknopte vorm

Lijst 'formulering'

1. Eenvoud
2. Dubbele ontkenning
3. Ondubbelzinnig
4. Te leerstof-specifiek
5. Universele quantoren
6. Formulering specifiek
7. Doorzichtig
8. Meerdere antwoorden mogelijk
9. Om motivering gevraagd
10. Dubbele uitspraak beoordelen

Meerkeuzevragen:

11. Te ingewikkelde formulering
12. Stam alleen goede aanvulvraag
13. Te vervangen door aanvulvraag
14. Bij de alternatieven:
 - a. Kort geformuleerd
 - b. Woorden in de stam op te nemen
 - c. Aansluiting op stam
 - d. Dummy-alternatieven vermijden
 - e. Afleiders echt fout
 - f. Juiste alternatief ook geldig
15. Onbedoelde aanwijzingen:
 - a. Lengte van juiste alternatief
 - b. Herhaling belangrijkste term
 - c. Onbekend jargon in afleider
 - d. Plaats juiste alternatief random
 - e. Logische onjuistheid
 - f. Universele quantoren
 - g. Onbedoeld kenmerk matchen
16. Ja/nee-vragen: frequentie ja/nee
17. Instructie compleet en eenduidig

Lijst 'inhoudelijke aspecten'

1. Onderwerp
2. Beheersingsvorm
3. Abstractieniveau
4. Opgegeven stof
5. Doorzichtig vooraf
6. Doelstelling
7. Strikvraag
8. Opinievraag
9. Letterlijke formulering
10. Vraagvorm passend?
11. Opsplitsbaar
12. Geef modelantwoord (+ tijd)
13. Karakteristieke fouten
14. Frequentie van fouten
15. Beoordelen juistheid
16. Waarderen
17. Meerkeuzevragen:
 - a. Is onderscheiden het doel?
 - b. Streep juiste antwoord aan
 - c. Relatie 17a tot 12
 - d. Orden afleiders op moeilijkheid
 - e. Onwaarschijnlijke afleiders
 - f. Moeilijkheid onderschat
 - g. Proportie die antwoord *weet*

Lijst algemene vragen

1. Onderwerpen representatief
2. Spreiding
3. Makkelijke vragen niet vergeten
4. Over-all doorzichtigheid

8.5 Literatuur

COHEN, M. J., *Studierechten in het wetenschappelijk onderwijs*, (proefschrift in handelseditie), Zwolle, Tjeenk Willink, 1981.

Richtlijnen voor ontwikkeling en gebruik van psychologische tests en studietoetsen. Amsterdam: Nederlands Instituut voor Psychologie, 1978.

Deze richtlijnen hebben een nogal eenzijdige nadruk op wenselijkheden die aangedragen worden vanuit de psychometrie. Dat brengt met zich mee dat ze weinig of niets toevoegen aan de behandeling die COHEN (1981) over hetzelfde onderwerp

geeft. Deze richtlijnen hebben echter wel een zekere bindende werking voor psychologen. Het is te wensen dat het NIP op korte termijn een nieuwe uitgave van de *Richtlijnen* gaat verzorgen waarin het werk van Cohen verdisconteerd is. Zoals de *Richtlijnen* nu geformuleerd zijn steunen ze sterk op de door de American Psychological Association uitgegeven *Standards for educational and psychological tests*, Washington, D.C., 1974 (nieuwe editie in voorbereiding).

Gerefereerde literatuur

- Aitchison, J., & I. R. Dunsmore, *Statistical prediction analysis*. Cambridge: Cambridge University Press, 1975.
- Anderson, J. R., *Cognitive psychology and its implications*. San Francisco: Freeman, 1980.
- Anderson, J. R. (ed.), *Cognitive skills and their acquisition*. Hillsdale, New Jersey, Erlbaum, 1981.
- Anderson, R. C., R. J. Spiro, & W. E. Montague (eds.), *Schooling and the acquisition of knowledge*. Hillsdale, New Jersey, Erlbaum, 1977.
- Anderson, R. C., How to construct achievement tests to assess comprehension. *Review of Educational Research*, 1972, 42, 145-170.
- APA, *Standards for educational and psychological tests*. Washington, DC: American Psychological Association, 1985.
- Bakels, H. L. (red.), *Nederlands recht in kort bestek*. Deventer: Kluwer, 1971.
- Baker, G. P., & P. M. S. Hacker, *Wittgenstein: understanding and meaning; an analytical commentary on the Philosophical Investigations*. Oxford: Basil Blackwell, 1980.
- Bar-Hillel, M., & R. Falk, Some teasers concerning conditional probabilities. *Cognition*, 1982, 11, 109-122.
- Barr, A., & E. A. Feigenbaum, *The handbook of artificial intelligence*, Volume I. Los Altos, California: Kaufman, 1981.
- Belnap, N. D., Jr., & T. B. Steel, Jr., *The logic of questions and answers*. London: Yale University Press, 1976.
- Bertels, K., & D. Nauta, *Inleiding tot het modelbegrip*. Bussum: De Haan, 1969.
- Bhaskar, R., & H.A. Simon, Problem solving in semantically rich domains: an example from engineering thermodynamics. *Cognitive Science*, 1977, 1, 193-215.
- Bloom, B. S., J. T. Hastings, & G. F. Madaus (eds.), *Handbook on formative and summative evaluation of student learning*. London: McGraw-Hill, 1971.
- Borger, R., & F. Cioffy (eds.), *Explanations in the behavioral sciences*. London: Cambridge University Press, 1970.
- Bormuth, J. R., *On the theory of achievement test items*. London: The University of Chicago Press, 1970.
- Breuker, J. *In kaart brengen van leerstof*. Aula 801, Het Spectrum, Utrecht.
- Brink, W. P. van den, Het optimale aantal alternatieven per item. *Tijdschrift voor Onderwijsresearch*, 1979, 4, 151-158.

- Chi, M. T. H., J. P. Feltovich, & R. Glaser, Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 1981, 5, 121-152.
- Cohen, M. J., *Studierechten in het wetenschappelijk onderwijs*. Zwolle: Tjeenk Willink, 1981.
- Crombag, H. F. M., J. L. de Wijkerslooth, & E. H. van Tuyll van Serooskerken, *Over het oplossen van casusposities*. Groningen: Tjeenk Willink, 1972.
- Crombag, H. F. M., Over het inrichten van leersituaties. In A. I. Vroeijenstijn (1981).
- Dewey, J., *Democracy and education*. New York: Macmillan, 1961 (1916).
- Dousma, T., & A. Horsten, *Tentamineren*. Groningen: Wolters Noordhof, 1980.
- Duda, R. O., & J. G. Gaschnig, Knowledge-based expert systems come of age. *BYTE*, 1981, 238-284.
- Elstein, A. S., L. S. Shulman, & S. A. Sprafka, *Medical problem solving; an analysis of clinical reasoning*. London: Harvard University Press, 1978.
- Euwe, M., *Handboek voor de gevorderde schaker*. Den Haag: Van Stockum, 1966.
- Farr, R., *Reading: what can be measured?* Newark, Del.: International Reading Association, 1969.
- Freudenthal, H. *Weeding and sowing. Preface to a science of mathematical education*. Dordrecht: Reidel, 1978.
- Furst, J. E., Bloom's taxonomy of educational objectives for the cognitive domain: philosophical and educational issues. *Review of Educational Research*, 1981, 51, 441-453.
- Gagné, E. D., Long term retention of information following learning from prose. *Review of Educational Research*, 1978, 48, 629-665.
- Graf, E. de, & J. A. M. Galesloot, De ontwikkeling van een toetsmethode voor 'medisch probleemoplossen.' In H. G. Schmidt (red.), *Probleemgestuurd onderwijs; bijdragen tot de Onderwijsresearchdagen 1981*. Harlingen: Flevodruk, 1982. SVO reeks 57.
- Gregson, R. A. M., *Psychometrics of similarity*. London: Academic Press, 1975.
- Groot, A. D. de, *Methodologie*. Den Haag: Mouton, 1962.
- Groot, A. D. de, Some badly needed non-statistical concepts in applied psychometrics. *Nederlands Tijdschrift voor de Psychologie*, 1970, 25, 360-376.
- Groot, A. D. de, *Thought and choice in chess*. Den Haag: Mouton, 1978.
- Groot, A. D. de, & R. F. van Naerssen (red.), *Studietoetsen, construeren, afnemen, analyseren*. 2 delen. Den Haag: Mouton, 1973, 1974.

- Hempel, C. G., *Fundamentals of concept formation in empirical science*. London: The University Of Chicago Press, 1972 (1952).
- Hofstee, W. K. B., Bureaucratisering van de beoordeling? Een beschouwing naar aanleiding van het voor en tegen van multiple-choice-toetsen. *Universiteit en Hogeschool*, 1981, 27, 286-292.
- Hofstee, W. K. B., Participatie controle door 'onbenullige' toetsitems. *Nederlands Tijdschrift voor de Psychologie*, 1973, 28, 189-198. Ook in A. G. Vroon & S. E. M. Everwijn (red.), *Handboek voor de onderwijspraktijk*. Deventer: Van Loghum Slaterus (losbladig).
- Imbos, Tj., & M. Verwijnen, Evaluatie aan de medische faculteit Maastricht. *Metamedica*, 1978, 57, 21-32.
- Jardine, N., & R. Sibson, *Mathematical taxonomy*. London: Academic Press, 1980.
- Kaplan, A., *The conduct of inquiry; methodology for behavioral science*. San Francisco: Chandler, 1964.
- Klausmeier, H. J., *Learning and teaching concepts*. London: Academic Press, 1980.
- Klausmeier, H. J., & P. S. Allen, *Cognitive development of children and youth; a longitudinal study*. London: Academic Press, 1978.
- Kneebone, G. T., *Mathematical logic and the foundations of mathematics*. London: Van Nostrand, 1963.
- Landa, L. N., *Algorithmization in learning and instruction*. Englewood Cliffs, N.J.: Educational Technology Publications, 1976.
- Langerak, W. F., Herkennen versus herinneren, oftewel meerkeuze versus opstelvragen. *Tijdschrift voor Onderwijsresearch*, 1979, 4, 140-143.
- Larkin, J. H., Enriching formal knowledge: a model for learning to solve textbook physics problems. In J. R. Anderson (ed.), *Cognitive skills and their acquisition*. Hillsdale, New Jersey, Erlbaum, 1981.
- Macdonald-Ross, M., Scientific diagrams and the generation of plausible hypotheses: an essay on the history of ideas. *Instructional Science*, 1979, 8, 223-234.
- Markle, S. M., & P. W. Tiemann, *Really understanding concepts: or in frumious pursuit of the jabberwock*. Champaign, Illinois: Stipes, 1970.
- Mellenbergh, G. J., Een onderzoek naar het beoordelen van open vragen. *Nederlands Tijdschrift voor de Psychologie*, 1971, 26, 102-120.
- Merrill, M. D., & R. D. Tennyson, *Teaching concepts: an instructional design guide*. Englewood Cliffs, N.J.: Educational Technology Publications, 1977.
- Mettes, C. T. C. W., & A. Pilot, *Over het leren oplossen van natuurwetenschappelijke problemen; een methode voor ontwikkeling en evaluatie van onderwijs, toegepast op een cursus thermodynamica*. Proefschrift. Enschede: Onderwijskundig Centrum Technische Hogeschool Twente, 1980.

- Mezzick, J. E., & H. Solomon, *Taxonomy and behavioral science*. London: Academic Press, 1980.
- Michener, E. R., Understanding understanding mathematics. *Cognitive science*, 1978, 2, 361-383.
- Mirande, M. J. A., *Studeren door schematiseren*. Groningen: Wolters Noordhof, 1981 (oorspronkelijk uitgegeven als Aula 805, Het Spectrum, Utrecht).
- Mouly, G. J., & L. E. Walton, *Schaum's outline of test items in education*. New York: McGraw-Hill, 1962.
- Murphy, E. A., *The logic of medicine*. London: Johns Hopkins University Press, 1976.
- Naerssen, R. F. van, Simpele items tegenover complexe. *Tijdschrift voor Onderwijsresearch*, 1980, 5, 193-198.
- Nagel, E., *The structure of science. Problems in the logic of scientific explanation*. London: Routledge & Kegan Paul, 1961.
- Nauta, D., *Logica en model*. Bussum: De Haan: 1970.
- Neisser, U., *Cognitive Psychology*. New York: Appleton-Century-Crofts, 1967.
- Newell, A., & H. A. Simon, *Human problem solving*. Englewood Cliffs, N.J.: Prentice Hall, 1972.
- Nickles, Th. (ed.), *Scientific discovery, logic, and rationality*. Dordrecht: Reidel, 1980.
- NIP, *Richtlijnen voor ontwikkeling en gebruik van psychologische tests en studietoetsen*. Nederlands Instituut voor Psychologen, 1978.
- Noordman, L. G. M., & W. Vonk, Lezen: het begrijpen van tekst. *Nederlands Tijdschrift voor de psychologie*, 1981, 36, 385-408.
- Odell, L., Teaching writing by teaching the process of discovery: an interdisciplinary enterprise. In L. W. Gregg & E. R. Steinberg (eds.) *Cognitive processes in writing*. Hillsdale, NJ: Erlbaum, 1980.
- Polya, G., *How to solve it. A new aspect of mathematical method*. Princeton, NJ: Princeton University Press, 1971 (1957).
- Popham, W. J., *Modern educational measurement*. London: Prentice-Hall, 1981.
- Rapoport, A., *Semantics*. New York: Thomas Y. Crowell, 1975.
- Reder, L. M., The role of elaboration in the comprehension and retention of prose: a critical review. *Review of Educational Research*, 1980, 50, 5-53.
- Rickards, J. P., Adjunct questions in text: a critical review of methods and processes. *Review of Educational Research*, 1979, 49, 181-196.
- Roid, G. H., & T. M. Haladyna, The emergence of an item-writing technology. *Review of Educational Research*, 1980, 50, 293-314.

- Roid, G., & T. Haladyna, *A technology for test-item writing*. London: Academic Press, 1982.
- Rubinstein, M. F., & K. Pfeiffer, *Concepts in problem solving*. Englewood Cliffs, N.J.: Prentice Hall, 1980.
- Simon, H. A., Problem solving and education. In D. T. Tuma & F. Reif (eds.) *Problem solving and education: issues in teaching and research*. Hillsdale, NJ: Erlbaum, 1980.
- Smith, E. E., & D. L. Medin, *Categories and concepts*. Cambridge, Mass.: Harvard University Press, 1981.
- Sneath, P. H. A., & R. R. Sokal, *Numerical taxonomy, the principles and practice of numerical classification*. San Francisco: Freeman, 1973.
- Sparks, C. P., Open versus secure testing. *Personnel Psychology*, 1980, 33, 1-2.
- Spiegel, M. S., *Advanced calculus*. New York: McGraw-Hill, 1963.
- Stegmüller, W. *Probleme und Resultate der Wissenschaftstheorie und Analytische Philosophie. Band I: Wissenschaftliche Erklärung und Begründung; Band II: Theorie und Erfahrung*. Berlin: Springer Verlag, 1969, 1970.
- Suppe, F. (ed.), *The structure of scientific theories*. Urbana, Illinois: University of Illinois Press, 1977.
- Tarski, A., The semantic conception of truth. *Philosophy and Phenomenological Research*, 1944, 4, 341-375. Reprinted in F. Zabeeh, E. D. Klemke, & A. Jacobson (eds.), *Readings in semantics*. Urbana, Illinois: University of Illinois Press, 1974.
- Terlouw, C., C. T. C. W. Mettes, & F. Roemers, Het leren ontwerpen van overheidsbeleid; een gewenst handelingsverloop. Enschede, Universiteit Twente: Onderwijskundig Centrum, rapport nr. 47, 1981.
- Thorndike, R. L. (ed.), *Educational measurement*. Washington, DC: American Council on Education, 1971.
- Tiemann, P. W., & S. M. Markle, *Analyzing instructional content: a guide to instruction and evaluation*. Champaign, Illinois: Stipes, 1978.
- Tromp, D., Het oordeel van studenten in een individueel studie systeem. *Onderwijs Research Dagen*, 1979.
- Tuma, D. T., & F. Reif (eds.), *Problem solving and education: issues in teaching and research*. Hillsdale, N.J.: Erlbaum, 1980.
- Vaags, D. W., *Over het oplossen van technische problemen*. Proefschrift. Eindhoven: Technische Universiteit Eindhoven, 1975.
- Vroeijenstijn, A. I. (red.), *Kwaliteitsverbetering hoger onderwijs*. Voorburg, CBOWO/Stichting Nationaal Congres, 1981. (Utrecht: VSNU).
- Westrhenen, J. van, De toetsing van onderwijsdoelen. Een empirische studie naar de functie van begrippen en begrippenstructuren in het onderwijsleerproces. Groningen: Wolters-Noordhoff, 1977.

- Wickelgren, W. A., *How to solve problems. Elements of a theory of problems and problem solving*. San Francisco: Freeman, 1974.
- Wickelgren, W. A., *Cognitive psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1979.
- Wilbrink, B., Het verborgen vooroordeel tegen andere dan meerkeuzevragen. Amsterdam: Onderwijs Research Dagen, 1977, a.
- Wilbrink, B., *Cesurbepaling*. Amsterdam: COWO, UvA, 1977 (1e ed.), b.
- Wilbrink, B., *Studiestrategieën*. Amsterdam: COWO, UvA, 1978.
- Wilbrink, B., *Toetsen*. Amsterdam: COWO, UvA, 1979.
- Wilbrink, B., *Cesurbepaling*. Amsterdam: COWO, UvA, 1980 (2e ed.).
- Wittgenstein, L., *Filosofische onderzoekingen*. Meppel: Boom, 1976 (1953).
- Wittgenstein, L., *Zettel*. Oxford: Blackwell, 1967.

Register

[geeft bladzijdennummers van de oorspronkelijke uitgave, gelijk aan de paginering in dit pdf-document]

aannamen 176,181
aanvulvragen 21 e.v.
aanwijzen 92
abstractieniveau 20, 46, 50 e.v., 89, 96, 99, 116, 125, 151
achteraf weglaten van vragen 199
afbeelden 121 e.v.
afleiders 25, 28, 38, 203
algemene beginselen van behoorlijk bestuur 194 e.v.
algoritmen 132 e.v.
alternatieven 30, 203
- , aantal 55
analyse 144, 152 e.v.
antwoordformulieren 23
argumentatie 156, 162
AROB-beroep 194
art. 40 WUB 194

bedenken van vragen 57
begrip 78, 91, 11; zie ook term
begrijpen 71, 79, 100
benoemen 91, 98 e.v.
beoordelingsvoorschrift 36, 189, 195, 196, 199
beroep art. 40 WUB 194
beschrijven 76, 113 e.v.
betekenis 76, 113 e.v.
- , systemische 130
Bloom 81, 93, 152
bonusvragen 43

categorieën, natuurlijke 128
cesuurbepaling 16
classificaties 66, 128 e.v.
classificeren 142
concrete vragen zie abstractieniveau
constructs 62 e.v., 108 e.v.
context 62

definities 57, 64, 70 e.v., 86 e.v.
-, formele 108 e.v.
-, operationele 75
-, ostensieve 72
denkprocessen 8
determineren 87
diagnose 131
doorzichtigheid 35, 37 e.v., 55, 197
dubbele ontkenning 201

eigenschappen 65, 74, 97, 108, 111

fair play 196
familiegelijkenis 76
feed-forward 14, 159
formele uitslag 199
formuleren 117 e.v., 200 e.v., 217

geheimhouding 39
gelijke behandeling 199
gelijkheidsbeginsel 195
gemakkelijke vragen 41
grensgevallen 77, 97, 102

herinneren 24
herkennen 24, 91, 98 e.v.

impliciete context 141
inferentie 144, 158 e.v.
intelligentie 81, 153, 216
invulvraag 21
inzagerecht 196
item form 44

ja/nee-vragen 29, 204

kenbaarheidsbeginsel 197
kenmerken 72, 74, 109, 167, 177
kennis 169
kennisvragen 81
keuzevragen 24 e.v.
kort-antwoordvraag 21
Kunstmatige Intelligentie 191

kwaliteit van toetsvragen 193-219

lay-out 54

leerstofinventarisatie 15, 57-80, 143, 177 e.v.

literatuurtentamen 143-163

matching vraag 28, 106

meerkeuzetoets 215

meerkeuzevraag 12, 21, 26 e.v., 55, 126, 198, 200, 206, 217

meervoudige ja/nee-vraag 28, 105

metaplan 172

modelantwoord 36, 188, 195, 196, 199, 205

moeilijke vragen 43

moeilijkheid 41 e.v., 215

-, van vragen 197

moeilijkheidsvarianten 190

motivering 195

motiveringsbeginsel 195

motiveringsplicht 196

nakijken 23, 35

niet-voorbeeld 97

objectieve vragen 33

objectiviteit 32, 211

onafhankelijke beoordeling 198, 207 e.v.

onbenullige toetsvragen 146

ondergeneralisatie 73, 105

onderscheiden 122 e.v., 142, 206

onderwijsdoelen 13, 14 e.v., 35, 39, 57, 197

ontdekken 191

openbaarheid 197

opinievrage 205

opschonen 103 e.v.

opstel 35

opstelvragen 33, 34 e.v.

overgeneralisatie 73, 105

parafrasen 149

participatiecontrole 144 e.v.

plannen 169 e.v., 179

Plato 110

problemen 164-192

-, aanpak van 121

producties 174
p-waarde 215

raadkans 27, 32, 215
r-bis 215
regel 73
regels bij het examineren 194 e.v.
relaties 57, 64, 65, 66 e.v., 69, 76, 112, 147, 156, 162
-, tussen termen 115-142
representativiteit 14, 36, 197; *zie ook* onderwijsdoelen
richtlijnen 219
r-it, 215 e.v.
rompvragen 44 e.v., 187
rompzin 86
routines 132

schematiseren 66, 79, 122
schemavorm 68
scoren 23, 30, 196
scoringssleutel 32
stappenschema 132 e.v.
strikvraag 205
studierechten 55, 193
studiestrategie 39

take-home-tentamen 94
taxonomie 128 e.v.
techniek 135
tekst, vragen bij 143-163
termen 57, 66 e.v., 78, 81-114, 155, 161
-, abstracte 61 e.v., 97
-, (indirect) waarneembare 59 e.v.
-, theoretische 64 e.v., 97
-, waarneembare 127
terminologie, wetenschappelijke 58
thema 156
toetsanalyse 16
toetsen 174, 179
typologie 128 e.v.

universele quantoren 204

verklaren 139
-, van verschijnselen 7
verknoptheid van termen 66
verschillen tussen studenten 13, 197
vertalen 82 e.v., 117 e.v.
vertrouwensbeginsel 197

voorbeelden 47, 73, 91 e.v., 98 e.v., 111 e.v.

voorspellen 139

vraagvormen 20

vrijgeven van vragen 39, 55

vuistregels 16, 56, 78

waarnemen 92, 93

werkstuk 35

wetenschapsfilosofie 65, 191

wetten 69, 137 e.v.

zak/slaag-grens 16, 30

<http://www.benwilbrink.nl/projecten/83ToetsvragenAula.pdf>

Het boek is in herziening, tussenresultaten zijn beschikbaar op de projectenpagina's

<http://www.benwilbrink.nl/projecten/toetsvragen.1.htm>

(hoofdstuk 1),

<http://www.benwilbrink.nl/projecten/toetsvragen.2.htm>

(hoofdstuk 2) etcetera.