

Expertise in het oplossen van natuurwetenschappelijke problemen bij studenten in drie stadia

I.de Bruijn
Universiteit Twente
Faculteit Technische Natuurkunde

Summary

A study has been done on the observation of expert behavior with physics students in three stadia: grade 10, grade 12 and first year university. Students were given two physics problems, for which the theory had been presented in grade 10 and repeated in two other occasions. A system for scoring audiotaped protocols was developed on theoretical grounds, consisting of elements of forward search, strategic knowledge and rule automation. Scoring with this system provided valid distinctions between students showing elements of expert behavior and students who did not. Expert behavior proved to begin in grade 12 and university students showed this behavior regularly.

A restriction is made with respect to the intersubjective validity of protocol scoring. Only experienced researchers agreed on the classification of subjects, while junior research assistants, after eight hours training, did not reach this proficiency.

1. Inleiding

In de laatste 15 jaar zijn er in het onderwijs in de natuurwetenschappen verscheidene pogingen ondernomen om studenten te leren vraagstukken te maken volgens een systematische methode (Mettes & Pilot, 1980; Van den Berg, 1983; Heller & Reif, 1984; Ferguson-Hessler, 1989). Met 'studenten' bedoelen we daarbij ook leerlingen in het voortgezet onderwijs. De geciteerde onderzoekers hebben een gemeenschappelijke grondvorm in de methode die zij gebruiken: Allen gaan uit van de wenselijke volgorde *analyse — plan — uitwerking — controle*, met in de onderdelen diverse varianten. Het onderzoek dat in dit artikel wordt beschreven gaat over enkele principiële keuzen die bij de uitwerking van deze methode worden gemaakt, over het proces dat studenten doormaken wanneer zij expertise ontwikkelen in het oplossen van natuurwetenschappelijke¹ problemen en over het waarnemen van dit proces.

De vraagstukken waar het hier om gaat, voldoen aan een aantal kenmerken. In de eerste plaats is er een *rijke context*. Dit wil zeggen dat uit de probleemstelling verschillende associaties met de werkelijkheid mogelijk zijn, die niet allemaal nodig zijn bij de oplossing van het vraagstuk. Bijvoorbeeld zal het

bij een vraagstuk over het werpen van een bal vaak niet van belang zijn dat het geworpen voorwerp een ronde vorm heeft. Dit brengt met zich mee dat de student in de *analysefase* een eigen samenhangend beeld van relevante gegevens en van het gevraagde dient op te bouwen. Hierbij moeten gegevens worden geselecteerd, waarbij de *ervaring* van de student, met het vak en met het soort vraagstuk, een belangrijke rol blijkt te spelen.

Een tweede kenmerk van de vraagstukken is dat de oplossing niet in één stap kan worden gevonden en dat er gezocht moet worden naar de gewenste stappen en de volgorde waarin deze moeten worden gezet. Er is dus *geen algoritme* voor de oplossing en het probleem is *heuristisch* van karakter. Daarmee hangt samen dat de student een *plan* dient te ontwikkelen voor de oplossing van het vraagstuk.

Ter toelichting beschouwen we de vraagstukken zoals die voorkomen in een eerder in dit tijdschrift verschenen artikel van Vanderlocht & Van Damme (1989). Het gaat daarbij over het oplossen van kogelbaanproblemen, een relatief contextarme klasse van vraagstukken. Omdat uit een beperkt aantal formules dient te worden gekozen, is er bijna sprake van algoritmische problemen. Dit is waarschijnlijk de reden dat Vanderlocht en Van Damme weinig aandacht schenken aan de heuristiek voor de planning van het oplossingsproces. Er is als het ware één, telkens terugkomende heuristiek: Zet de gegevens om in een aantal vergelijkingen, zodat een oplosbaar stelsel ontstaat. De uitwerking is dan vervolgens geheel wiskundig en alleen bij de interpretatie van antwoorden komt de natuurkunde nog om de hoek kijken. Overigens is de voornaamste bevinding dan dat vaak een onvoldoende probleemanalyse wordt gemaakt en dat in deze fase veel fouten voorkomen.

Experts lossen problemen, van een vakdomein waarin zij expert zijn, anders op dan beginners. Niet alleen vinden zij vaker een oplossing en is deze beter dan bij beginners, ook is hun manier van oplossen kwalitatief anders en maakt deze gebruik van andere wegen. Eén van de kenmerken van het oplossingsproces van een expert is dat deze redeneert vanuit de gegevens en door *voorwaarts redeneren* een oplossingsroute construeert (Larkin, 1979; Larkin, McDermott, Simon & Simon, 1980). Een ander kenmerk is dat de expert geen gedetailleerd maar een *globaal doel* voor ogen heeft bij het construeren van de juiste weg van gegevens naar gevraagde (Greeno, 1976).

Alle systematische aanpakmethoden bevatten tot nu toe de aanbeveling *vanuit het gevraagde te werken* in die gevallen waar de oplossingsroute niet direct wordt gezien. Dit betekent in tenminste twee opzichten een afwijking van het redeneren als een expert: Er wordt geadviseerd *terugwaarts* te werken inplaats van voorwaarts en het doel dient scherp in het oog te worden gehouden inplaats van globaal. Hiervoor zijn overigens goede redenen, zoals in het vervolg uiteengezet zal worden. Niettemin dringt de vraag zich op in welk stadium van de

ontwikkeling van beginner tot expert, geadviseerd kan worden niet meer vanuit het gevraagde te werken maar middels een voorwaartse redeneerweg, vanuit de gegevens.

Ter beantwoording van deze vraag is een onderzoek uitgevoerd, waarbij studenten in drie 'stadia van expertise' dezelfde (natuurkunde)opgaven kregen uit te werken. In eerste instantie was daarbij het doel om na te gaan, wanneer het voorwaarts redeneren spontaan ging optreden. In de loop van het onderzoek kwam mede de vraag naar voren in hoeverre dit kan worden afgeleid uit (de protocollen van) het proces dat de student doormaakt, dus of een procesgerichte beoordeling van de mate van expertise mogelijk is. Dit is belangrijk als een docent met een groep studenten werkt aan probleemaanpak: Zo'n docent zal tijdens het proces - het produceren van de oplossing - immers willen inschatten of de student al vordert op de weg naar de expertise.

Over beide vragen gaat dit artikel. In het onderzoek werden natuurkundevraagstukken gebruikt, waarbij door een aantal transformaties van de gegevens uiteindelijk het antwoord kan worden verkregen (zie de bijlage voor de tekst van vraagstukken).

2. Voorwaarts redeneren, strategische kennis en verkortingen

Waarom lossen experts problemen anders op dan beginners? Oppervlakkig gezien: omdat ze meer over het vak weten, over het *probleemdomein*. Expertise kan dus niet verworven worden door alleen een methode van systematische aanpak te leren, er moet ook domeinkennis worden aangebracht. Vaak wordt die kennis verkregen door het oplossen (of: doorworstelen) van een aantal problemen, waarna de daarmee opgedane ervaring overdraagbaar wordt naar nieuwe problemen (transfer). Met behulp van de cognitieve psychologie kan een wat gedetailleerder beeld worden verkregen welke kennis dat zou moeten zijn en hoe dat de aanpak van problemen beïnvloedt.

Beginnende probleemoplossers kijken naar een aspect van de vraagstelling, pikken een gegeven eruit en gaan wat rekenen, zonder een duidelijk plan voor ogen te hebben. Experts kijken vooral naar de gegevens en met een half oog naar het gevraagde, maken eerst een globaal plan (of *zien* dat in één oogopslag) en gaan dan vanuit de gegevens de oplossing construeren. Het grootste verschil zit in de mogelijkheid een globaal plan te maken; dat kunnen beginners nog niet omdat zij te weinig ervaring met opgaven in het betreffende probleem domein hebben². Om hieraan tegemoet te komen wordt aan beginnende studenten de aanwijzing gegeven vanuit het gevraagde te werken. Het voordeel is dat bij het werken vanuit het gevraagde er minder beslissingen behoeven te worden genomen over vertakkingen in de oplossingsroute.

De domeinspecifieke kennis, die nodig is om bij een nieuw probleem vooraf een plan te kunnen maken, noemen wij *strategische kennis*³. Dit sluit aan bij

de definitie van Taylor en Evans (1985), die deze kennis opvatten als een onderdeel van de declaratieve kennis, meer specifiek het eindproduct dat na diverse herstructureringen tot stand komt. Een citaat:

Strategic knowledge is the final element of cognitive structure which is regarded primarily as a function of the structure of knowledge. Strategic knowledge is essentially concerned with domain-specific strategies for applying the knowledge-base (...) to a variety of domain-specific tasks. These domain-specific strategic processes should not be confused with general strategic processes (Taylor & Evans, 1985, p.350).

Strategische kennis wordt, onder andere, verworven door het zelf maken van vraagstukken of het bestuderen van voorbeelduitwerkingen van vraagstukken. Een nieuw probleem kan opgelost worden door naar *analogie* van één eerder ervaren voorbeeld te werken. Daarbij worden elementen uit het voorbeeld gekopieerd en in lichte mate aangepast. Hiervoor is nodig dat men zich het voorbeeld herinnert als het nieuwe probleem optreedt.

Anders, en veel krachtiger wordt de strategische kennis, wanneer twee of meer voorbeelden met elkaar in verband zijn gebracht en bijvoorbeeld de verschillen in toepassingsvoorwaarden van een regel of principe zijn opgemerkt. Men leert dan welke kenmerken van het probleem zijn gerelateerd aan 'hogere principes', waarmee het probleem kan worden opgelost, een *schema*. De beschreven herstructurering van de kennis wordt *schema-inductie* genoemd (Chi, Feltovitch & Glaser, 1981; Gick & Holyoak, 1983). Zo'n hoger principe is in de natuurkunde meestal een algemeen toepasbare wet (met restricties voor toepasbaarheid!) en in de scheikunde vaak een reactievergelijking. Een nieuw probleem wordt nu dus niet meer door herkenning, maar door *categorisering* opgelost. De hierbij gebruikte kennis is veel beter transferabel naar nieuwe problemen dan de herkenning van een analogie. Het is ook deze beschikbaarheid van schema's die het voorwaarts redeneren mogelijk maakt.

Bij het bestuderen van voorbeelden blijken studenten die later succesvol zijn bij het oplossen van opgaven, anders te werken dan minder succesvolle studenten. De betere studenten leggen meer verbanden tussen onderdelen van het voorbeeld en zij stellen zichzelf meer vragen, bijvoorbeeld waarom een regel in een bepaalde situatie wel of niet wordt toegepast (Chi, Bassok, Lewis, Reimann & Glaser, 1989). Deze structurerende activiteit bij het opnemen van nieuwe informatie wordt wel *elaboratie* genoemd, het is de activiteit die aanleiding geeft tot *schema-inductie*, dus tot beter transferabele strategische kennis. Het is duidelijk dat minder succesvolle studenten zouden moeten leren beter te elaboreren, bij het bestuderen van voorbeelden en het maken van vraagstukken. Er is nog geen onderzoek bekend waarbij dat is gelukt, maar de suggestie is om een systema-

tische aanpak te onderwijzen, met nadruk op een volledige probleemanalyse en het doorlopen van een controlefase.

Behalve krachtige schema's wordt ook de *automatisering van regels* genoemd als een aspect van de expertise in een probleemdomen (Cooper & Sweller, 1987). De theoretische gedachte hierachter is dat geautomatiseerd gebruik van regels minder belastend is voor het werkgeheugen, waardoor typische 'overload-effecten' minder snel zullen optreden. Cognitieve overbelasting leidt er onder andere toe dat het overzicht over de structuur van de -mogelijke- oplossing verloren gaat. Het maken van een globaal plan, en daarmee het voorwaarts redeneren, zal beter gaan wanneer de probleemoplosser intussen geen aandacht hoeft te geven aan de vraag hoe een bepaalde regel moet worden toegepast. Het niet-automatisch toepassen van een regel uit zich vaak in het verbaal repeteren van de informatie. Het al of niet aanwezig zijn van dit verbale aspect, de *verkorting* bij het toepassen van de regel, beschouwen wij als een indicatie voor het optreden van regelautomatisering. De verkortingen treden pas op na een relatief langdurige training, in elk geval langer dan nodig is voor het verwerven van schema's (Cooper & Sweller, 1987). De verkortingen zijn vooral te verwachten bij vaardigheden in de syntactische sfeer: het relateren van een fysische grootheid aan een symbool, het weergeven van een wetmatigheid in de vorm van een formule.

Onze oorspronkelijke vraag krijgt op grond van de voorgaande beschouwing een uitbreiding. Niet alleen willen we weten wanneer het voorwaarts redeneren optreedt, tevens interesseert ons wanneer studenten tekenen vertonen van strategische kennis en van verkortingen bij de regeltoepassing. De theorie geeft aan dat die drie zaken samenhangen.

3. De methode: Tutorinterviews en scores van protocollen

Wij maken in dit onderzoek gebruik van een methode, die wij omschrijven als een *tutor-interview*. In eerste instantie wordt geprobeerd de student een vraagstuk zelf te laten oplossen. Wanneer dit niet lukt, of wanneer er fouten optreden, zal de interviewer als tutor gaan optreden: Er worden hints gegeven of stukjes uitleg. Fouten die niet te maken hebben met de oplossingsweg, zoals rekenfouten, worden door de tutor rechtgezet. De bedoeling is steeds de student de oplossingsroute zelf te laten vinden, deze route wordt dus nooit rechtstreeks 'weggegeven'. Wij hebben ervaren dat op deze wijze protocollen ontstaan met meer informatie dan bij het geheel zelfstandig werken door de student, onder andere omdat de student minder vaak blijft hangen op een detail.

De student lost tijdens een interview twee vraagstukken op (zie bijlage). In de meeste gevallen is de tijdsduur van het interview beperkt tot 50 minuten, wat ertoe leidt dat onderdeel 2c soms niet aan de orde komt. Dit levert weinig verlies van informatie op, omdat dit onderdeel het moeilijkste is: Studenten

die er niet aan toe komen zullen degenen zijn die weinig produceren op dit onderdeel, althans geen zaken waar we in de protocollen naar zoeken. Het interview wordt op geluidsband opgenomen en de uitwerking-op-schrift van beide vraagstukken wordt bewaard. Geluidsband en uitwerkingen vormen samen het protocol van het interview.

Uit de protocollen wordt vervolgens afgeleid in hoeverre het oplossingsgedrag van de student bepaalde, hierna te bespreken, kenmerken vertoont. Een rechtstreekse methode gaat als volgt. Er wordt een oplossingsmodel opgesteld dat stappen in de redenering aangeeft, zodanig dat bij het successievelijk volgen van die stappen de oplossing wordt bereikt. Voor die studenten die redeneren volgens het model, of een variant daarvan, kan dan worden getabelleerd of zij de stappen uitvoeren, of dat spontaan en vlot gaat, of er hulp nodig is, etc. Voor beide vraagstukken is daarom zo'n stappenschema opgesteld.

Bij een alternatieve methode van scoren wordt rechtstreeks gelet op het voorkomen van elementen van (a) voorwaarts redeneren, (b) strategische kennis en (c) verkortingen. Er zijn *criteria* opgesteld, op grond waarvan bij het scoren van de protocollen tot de aanwezigheid van deze elementen kan worden geconcludeerd. Deze criteria zijn hoofdzakelijk afkomstig uit de theorie, maar zijn gevalideerd met enkele proef-protocollen van interviews met ervaren probleemoplossers.

Tot voorwaarts zoeken kan men concluderen wanneer de student (a1) in de redenering uitgaat van de gegevens, (a2) een vage, of een gevariëerde doelmomschrijving hanteert aan het begin van de oplossing, (a3) bij een eerste analyse reeds kennis over het probleemdomein inschakelt en (a4) een kwantitatieve analyse (formules, berekeningen) vooraf laat gaan door een kwalitatieve analyse. Alle criteria van categorie a worden alleen gescoord in het begin van de oplossing. In geval van twijfel vraagt de tutor achteraf of de student een bepaalde gedachtengang heeft toegepast, bijvoorbeeld wanneer de tutor vermoedt dat a3 van toepassing is maar wanneer dat door de student niet hardop is uitgesproken.

Het gebruik van strategische kennis heeft als mogelijke kenmerken dat de student (b1) opmerkt dat dit type probleem meestal op een bepaalde wijze wordt aangepakt, (b2) een algemene wet neemt en nagaat of deze op het specifieke geval van toepassing is, (b3) in het probleem een bekend deelprobleem tracht te onderscheiden en (b4) de status van de variabelen in het probleem (gegeven/-gevraagd/berekenbaar/onbekend) bij het ontwikkelen van de oplossing geregeld onderzoekt (*monitoring van variabelen*). Verkorting en automatisering wordt waargenomen aan het feit dat (c1) een aantal stappen van de oplossing kort na elkaar en in samenhang wordt gemaakt, (c2) een stuk van de oplossing wordt geproduceerd waarbij impliciet één of meer denkstappen juist zijn verricht (de tutor kan dit achteraf navragen), wanneer (c3) het algemene tempo van oplossen

hoog ligt en (c4) de student stukjes van de oplossing uit het hoofd produceert. Bij c1 en c2 is het stappenschema van de 'rechtstreekse scoremethode' een hulpmiddel. Er zijn negatieve scores mogelijk op de kenmerken a1 (bij werken vanuit gevraagde), a3 (domeinkennis blijkt niet aanwezig), b1 (probleem blijkt volkomen onbekend) en c3 (zeer langzaam uitwerken).

Voor deze tweede scoremethode kan er al bij het afnemen van het interview rekening te worden gehouden met de "scoorbaarheid" van het protocol. Een goed voorbereide tutor vraagt expliciet aan de student welke gedachtengang is gevolgd, indien dit niet duidelijk blijkt. Zodoende wordt het ook voor een andere beoordelaar van het protocol duidelijk dat er, bijvoorbeeld, een stap uit het hoofd is gedaan. Verder is in het protocol een volgorde-codering ingebouwd: Wanneer de student in geschreven tekst iets tussenvoegt, geschiedt dat in een afwijkende kleur.

4. De uitvoering: Proefpersonen en tutoren

In eerste instantie is volstaan met een tiental proefpersonen. Vier werden gerecruteerd uit het eerste jaar van de Universiteit Twente (studierichting natuurkunde) waarbij twee studenten zijn gekozen met het cijfer 8 voor het vak Dynamica en twee met het cijfer 6 voor dit vak. Vier studenten waren afkomstig uit 6-VWO-klassen, van twee verschillende scholen, waarbij er wederom twee met een 8 of hoger en twee met een 6 voor het vak natuurkunde werden gekozen. Tenslotte werden er twee studenten van 4 VWO gekozen, van een derde school, die tot de besten van hun klas voor het vak natuurkunde behoorden. Het aantal uit 4 VWO is beperkt gehouden, omdat de verwachting was dat deze categorie zeer veel moeite zou hebben met het oplossen van de gekozen vraagstukken en dat er daarom weinig relevante gegevens zouden worden verkregen. In deze opzet zijn er studenten bij het onderzoek betrokken, die respectievelijk éénmaal (in 4-VWO), tweemaal (6-VWO, examenvorbereiding) en driemaal (eerste jaar universiteit) met de leerstof in aanraking zijn geweest.

In alle gevallen werd een tweetal studenten door de eigen docent aangezocht als proefpersoon, nadat aan de docent was uitgelegd wat de bedoeling was en aan welke voorwaarden de proefpersonen dienden te voldoen. In deze onderzoeksronde trad alleen de auteur op als tutor (gecodeerd als tutor A) en verrichtte hij de scoring van de protocollen zelf. In een tweede ronde werden twee andere tutoren (B en C) ingezet. Zij volgden een training, die zes uur in beslag nam en bestond uit het leren van de stappenschema's en het enkele malen afnemen en scoren van proef-protocollen. Er werden acht andere studenten uit 6-VWO-klassen geïnterviewd, door elke tutor vier. De tutoren scoorden hun zelf afgenomen protocollen en ook die van hun collega, waardoor een schatting van de interbeoordelaarsbetrouwbaarheid van de scoring werd verkregen. Aan het eind van het onderzoek, waar behoefte bleek te bestaan aan een secundaire

analyse van gegevens, werd nog een vierde persoon (D) gevraagd de opgenomen protocollen te beoordelen; deze trad niet als tutor op.

5. Resultaten van de eerste serie interviews

In de eerste ronde zijn tien interviews afgenomen, bij studenten in drie stadia van ervaring met het kennisdomein. In tabel 1 is aangegeven in hoeverre deze proefpersonen de vraagstukken wisten op te lossen, of ze daarbij op essentiële punten hulp nodig hadden en hoeveel tijd ze er voor nodig hadden. Voor het tweede vraagstuk hebben deze gegevens alleen betrekking op de onderdelen a en b. In de tabel is niet weergegeven bij welke stappen de hulp moest worden geboden.

Tabel 1: Gegevens over het oplossen van de vraagstukken door 10 studenten.

Code ppn: 4 = 4e klas VWO, 6 = 6e klas VWO, U = universitair, h = hoog cijfer, l = laag cijfer

pp. code	1 Uh	2 Uh	3 Ul	4 Ul	5 6h	6 6h	7 6l	8 6l	9 4	10 4
vrgst.1										
tijd (min)	17,5	8,5	8,7	8,2	7,4	9,0	22	15	24	23
opgelost	nee	ja	ja	ja	ja	ja	ja	nee	nee	ja
hulpmoment	-	1	0	1	0	1	3	-	-	3
vrgst.2 a + b										
tijd (min)	3,5	4,3	6,5	8,5	7,0	5,8	8,0	9,7	14	16
opgelost	ja	ja	ja	ja	ja	ja	ja	ja	nee	ja
hulpmoment	0	0	1	0	0	1	3	4	-	4

Bij vraagstuk 1 deed zich de bijzonderheid voor dat één van de hoogst gekwalificeerde proefpersonen, een universitaire student met een hoog cijfer, het vraagstuk niet kon oplossen. Dit was te wijten aan het feit dat hij de benadering met de *centrifugale* kracht koos en hierbij niet systematisch te werk ging. Ook enkele andere studenten kozen deze benadering, maar zij losten het vraagstuk wel op⁴. Voor deze personen moest het stappenschema worden aangepast, of anders geïnterpreteerd, teneinde hun oplossingsweg te kunnen coderen. Voor de andere proefpersonen was het stappenschema goed toepasbaar: Per stap kon worden aangegeven of de student de stap had uitgevoerd en of dat vlot, normaal of met hulp was gebeurd. In tabel 2 is dit aangegeven, voor de stappen uit vraagstuk 1b (voor een gedetailleerde uitwerking zie De Bruijn, 1991).

Vervolgens zijn de protocollen opnieuw geanalyseerd, nu met de tweede scoremethode, rechtstreeks gericht op de elementen voorwaarts zoeken, strategische kennis en verkorting, volgens de hiervoor gegeven criteria. Het resultaat is in tabel 3 vermeld.

Tabel 2: Snelheid waarmee de diverse stappen in de oplossing van vraag 1b werden gevonden.
 ++ = spontaan, snel; + = zelf, na enig denken; 0 = moeizaam of met hulp; - = niet gevonden

ppn	1	2	3	4	5	6	7	8	9	10
A. $F_z = mg$	++	++	++	++	+	+	+	+	0	0
B. S ontbinden	-	0	+	++	++	+	-	-	0	0
C. $S_{\text{vert}} + mg = 0$	-	+	+	+	++	+	+	-	+	+
D. $S_{\text{hor}} = F_{\text{mpx}}$	-	+	+	+	+	+	++	0	0	0
E. $S_h = S \cdot \sin \alpha$ etc	0	+	++	+	+	+	+	-	0	0
F. $F_{\text{mpx}} = m \cdot w^2 \cdot R$	+	+	+	+	+	++	0	+	+	+
G. omwerking tan α	-	+	+	+	+	++	0	-	-	+
H. tan α uit fig.	+	0	0	+	+	0	0	0	-	++

Uit de tabellen 1 en 2 is bijvoorbeeld te concluderen of een student het vraagstuk oplost en hoe vlot dat gaat. Over het *proces* van oplossen is met deze analyse-methode slechts een enkel gegeven te vinden. Uit tabel 3 komt een meer consistent beeld naar voren, dat ook betere antwoorden geeft op onze onderzoeksvragen. Zo kan geconstateerd worden dat elementen van expertgedrag voorkomen bij de universitaire studenten en in mindere mate bij studenten van 6-VWO, daar

Tabel 3: Aangetroffen indicaties voor voorwaarts zoeken, strategische kennis en verkorting.
 1, 2 = aangetroffen resp. 1 of 2 maal
 - = indicatie aangetroffen van het tegendeel
 code ppn: zie Tabel 1.

ppn code	1 Uh	2 Uh	3 Ul	4 Ul	5 6h	6 6h	7 6l	8 6l	9 4	10 4
Voorwaarts zoeken										
a1 werkt vanuit gegevens	1	1	1						-	-
a2 hanteert globaal doel			1							
a3 domeinkennis aanwezig	-						-	-		-
a4 kwalitatieve analyse	1-		2	1	1	1	-			
Strategische kennis										
b1 herkent type probleem	1	1			1		-	-		-
b2 verbijzondert alg. wet	1	1	1			1			-	
b3 onderscheidt deelprobl.	1			1						
b4 monit. status variab.					1				-	
Verkorting										
c1 stappen in samenhang	1	2	2	1	1	1				
c2 impliciete denkstap	1	1		1		1			1	
c3 (zeer) vlot werken	1-	1		1	1	1	-	-	-	-
c4 stukken uit 't hoofd			1	1						

alleen bij degenen met hoge cijfers⁵. Ook is in de tabel te zien dat het voorwaarts zoeken alleen voorkomt bij die studenten, waar ook strategische kennis en verkorting wordt waargenomen.

Deze constatering, die met de verwachtingen overeenkomen, dienen met enige reserve te worden gezien, omdat de scoring van de protocollen volgens tabel 3 op persoonlijke interpretaties kan zijn gebaseerd. Wanneer de score-procedure zodanig kan worden geobjectiveerd dat ook anderen tot dezelfde beoordeling van de protocollen kunnen komen, krijgen deze conclusies meer waarde. Teneinde dit na te gaan is de tweede onderzoeksronde uitgevoerd, waarbij met nieuwe proefpersonen en andere tutoren gepoogd is de consistentie van het resultaat aan te tonen. De reeds geanalyseerde protocollen bleken niet bruikbaar voor een tweede beoordelaar, omdat de maatregelen voor verhoging van de scorebaarheid nog niet waren toegepast.

6. Resultaten van de tweede serie interviews, met twee tutoren

Bij de tweede onderzoeksronde werden acht studenten van 6-VWO geïnterviewd, omdat in de eerste ronde de indicatie was verkregen dat juist in dit stadium verschillen in het oplossingsproces optreden. Wederom werden studenten uitgenodigd met hoge (ca. 8) of lage cijfers (ca. 6) voor natuurkunde. De tutoren waren natuurkundestudenten in de eindfase van hun studie, die het prédoctoraal deel van de lerarenopleiding hadden gevolgd. Zij werden gedurende een dag getraind in het afnemen van de interviews en het scoren van de protocollen, aan de hand van een daartoe opgestelde tutorinstructie. De acht interviews werden afgenomen op twee verschillende dagen. Elke tutor nam vier interviews af en scoorde eerst de zelf opgenomen protocollen, daarna die van de collega.

De resultaten van deze scoring zijn in tabel 4 weergegeven. Het valt op dat de ene beoordelaar C veel meer scores toekent dan de de andere beoordelaar, B. Ook lijkt C meer scores toe te kennen aan protocollen die hij zelf heeft afgenomen, hetzelfde geldt voor B.

Bij een dergelijk verschil kan de analyse van overeenstemming tussen beoordelaars het best gebaseerd worden op een niet-parametrische test, waarbij de rangorde die door beide beoordelaars wordt aangebracht bepalend is. De test is dus ongevoelig voor een trend van één der beoordelaar om systematisch vaker te scoren. Voor deze test is hier Kendall's tau gekozen, mede omdat deze een correctie bevat voor samenvallende rangorden ("knopen"). Bij acht te rangschikken objecten ligt de 5%-significantiegrens voor tau bij .48, uitgaande van enkelzijdige toetsing (Blalock, 1960).

De resultaten als weergegeven in tabel 4 kunnen worden onderzocht op basis van afzonderlijke criteria (a1, a2, etc), van categorieën (a, b, c) of van totale scores. Verder is een keuze mogelijk over het meetellen van min-scores. Bij het bepalen van totalen moet een beslissing worden genomen over de wijze

Tabel 4: Indicaties voor voorwaarts zoeken, strategische kennis en verkorting, als gescoord door twee beoordelaars.

l/h = laag/hoog rapportcijfer; 1,2,3 = aantal keren aangetroffen, - = indicatie voor het tegendeel; code ppn: zie Tabel 1.

ppn	1h	2l	3h	4l	5h	6l	7h	8l
tutor	B	B	B	B	C	C	C	C
beoord.	B C	B C	B C	B C	C B	C B	C B	C B
Voorwaarts zoeken								
a1 werkt vanuit gegevens	1 2-	1	2 2	2 2-	2 2	1	3 2	3 -
a2 hanteert globaal doel	1-	1-	- 2	2 1-	3 2	3 2	3 2	2 1
a3 domeinkennis aanwezig	3-	1-3	- 3	1-2	3 1-	1 1	3 1-	3 1
a4 kwalitatieve analyse	1		2	1	1 1	1	1	1
Strategische kennis								
b1 herkent type probleem	2- 3	1 1-	- 3	1	3 -	- 1	1-	3
b2 verbijzondert alg. wet	1 1	-	1	1		1 1	1-	1 1
b3 onderscheidt deelprobl.	1	1	2		1	2	1	1
b4 monit. status variab.	-	1	2	1	3	1-	3	3
Verkorting								
c1 stappen in samenhang	1 1	1	2	1- 1	1 1-	1 1	3 1-	3
c2 impliciete denkstap	1 2	1 2	2	- 1	1 -	1	1 -	2
c3 (zeer) vlot werken	2- 1-	2 -	1-	-	2	- 1-	2	3 1-
c4 stukken uit 't hoofd	1	1	1	1-			1-	

waarop negatieve scores ("minnen") meetellen. Hierbij is de regel gehanteerd dat een min alleen wordt geteld als op hetzelfde criterium niet ook een plus is gescoord. Het gaat immers om het scoren van elementen van expertgedrag en bij aantreffen hiervan is de vraag of ook het tegendeel wordt aangetroffen van minder belang.

Op het niveau van de criteria treden er problemen op met samenvallende rangorden. Criteria b3 en b4 zijn niet analyseerbaar, omdat beoordelaar B nergens een score toekent. Bij a3, a4, b2, c1 en c4 zijn er zoveel samenvallende scores, dat berekening van tau niet zinvol is. Op het oog is te zien dat er geen overeenstemming is tussen de beoordelaars: daar waar C overgaat naar een hogere score is dat voor B niet het geval. Voor de overige criteria is tau berekend, waarbij alleen voor criterium c2 ("impliciete denkstappen") een significant resultaat wordt gevonden ($\tau = .72$).

Op het niveau van de categorieën a, b en c zijn de tau-waarden respectievelijk .13, -.42 en -.31, derhalve geen overeenstemming. Voor de totaalscores van beide beoordelaars blijkt de tau-waarde -.46 te bedragen, wat eerder op tegenspraak dan op overeenstemming wijst.

Op de protocollen van de tweede onderzoeksrunde is vervolgens een secundaire

analyse uitgevoerd. Twee andere beoordelaars (A, de auteur, en D), beide ervaren in onderzoek naar systematische probleemaanpak, scoorden de protocollen opnieuw, uitgaande van de tutorinstructie als gebruikt bij de training van tutoren B en C.

Tabel 5: Vergelijking totaalscores (zie tekst) van vier beoordelaars op dezelfde protocollen, alsmede rangordes. Samenvallende rangordes ("knopen") cursief.
(l/h = laag/hog rapportcijfer)

ppn		1h	2l	3h	4l	5h	6l	7h	8l
Beoordelaar									
A	score	8	6	7	5	6	-2	0	-3
	rang	1	3	2	5	3	7	6	8
B	score	9	7	-2	5	5	6	5	3
	rang	1	2	8	4	4	3	4	7
C	score	15	8	21	12	21	9	23	25
	rang	5	8	3	6	3	7	2	1
D	score	2	-1	1	0	2	-2	-1	-1
	rang	l	6	3	4	l	8	6	6

Alleen het resultaat voor de totale scoring wordt hier vermeld, tezamen met de totalen voor B en C (zie tabel 5). Ook voor deze data is de non-parametrische maat van overeenstemming, Kendall's tau, te bepalen. Deze blijkt alleen voor A en D significant ($\tau = .62$, $p < 0.05$) en voor alle andere paren beoordelaars gering (maximaal bij A en B, $\tau = .27$, n.s.). Aan tabel 5 is verder nog te zien, dat beoordelaar D in het algemeen tot weinig scores komt. Bij nadere inspectie van de scores van D blijkt dit vooral de strategische kennis te betreffen. Tenslotte kan uit tabel 5 nog worden afgeleid dat A en D de studenten met hogere rapportcijfers -met uitzondering van ppn. 7- als de betere probleemoplossers kwalificeren.

7. Discussie en conclusies

De methode voor 'rechtstreekse' scoring van elementen van expertgedrag biedt perspectieven voor het verdere onderzoek naar probleemoplossen. Met name komen er gegevens beschikbaar over het *proces* van oplossen en die gegevens kunnen worden geïnterpreteerd in termen van een cognitieve theorie (als bij Cooper & Sweller, 1987). Met deze methode behoeft men dus het waarnemen van expertgedrag niet meer alleen te baseren op de 'produkten', het al of niet oplossen van problemen door de student. Tevens is er empirische validering gevonden voor de veronderstelling, dat het voorwaarts zoeken alleen optreedt, indien ook strategische kennis en/of verkorting wordt waargenomen. Verder is gebleken, dat het expertgedrag bij het oplossen van natuurkundeproblemen

zich reeds ontwikkelt bij de 'betere' studenten op het niveau van 6 VWO. Het onderzoek is echter beperkt geweest op het niveau van 4 VWO, zodat niet kan worden uitgesloten dat reeds daar, bij een enkele student, expertgedrag wordt aangetroffen. Een hypothese voor verder onderzoek is, dat in de ontwikkelingsgang van de betere studenten zich *eerst* strategische kennis en/of verkorting manifesteert en pas daarna een effectief voorwaarts zoekgedrag ontstaat.

Een voorbehoud dient te worden gemaakt ten aanzien van de overeenstemming tussen beoordelaars bij de beoordeling van protocollen. Alleen tussen ervaren onderzoekers bleek er overeenstemming te zijn betreffende het rangschikken van studenten naar toenemend expertgedrag. Minder ervaren onderzoeksassistenten bleken na de door ons gegeven training van één dag niet tot overeenstemmende beoordelingen te komen. Tevens bestaat de indruk dat de strategische kennis duidelijker dient te worden omschreven of geïnstrueerd in de tutorinstructie, omdat beoordelaars B en D hierop weinig scoren. De training heeft wel voldoende vaardigheid opgeleverd in het afnemen van tutor-interviews, want de verkregen protocollen waren scorebaar voor alle beoordelaars. Niettemin moet sterk betwijfeld worden of een docent in een klassikale situatie expertise bij het oplossingsproces kan waarnemen. Wanneer in groepjes aan vraagstukken wordt gewerkt, zal de terugkoppeling moeten plaatsvinden op basis van de produkten, de (deel)oplossingen die zijn geproduceerd.

De methode van het tutor-interview voor het opsporen van expertgedrag kan nog op enkele punten worden verbeterd. Het is denkbaar dat het interview slechts enkele minuten per vraagstuk omvat bij studenten die geen enkel teken van expertgedrag vertonen. Wanneer de tutor de aanpak van een vraagstuk met een aantal hints op gang moet brengen, wordt zeer sporadisch in het vervolg van het protocol nog een 'score' toegekend. De tijd kan dan beter worden aangewend voor het werken aan een ander vraagstuk, waarbij een nieuwe kans wordt geboden op het vertonen van het gezochte gedrag. Dit beperkt ook het werk van de beoordelaar: Waar nu drie kwartier nodig is voor het beluisteren van het protocol van twee vraagstukken, zou dit in de toekomst bijvoorbeeld tot twintig minuten voor vier vragen kunnen worden beperkt. Dit geeft ook een betere mogelijkheid tot het uitvoeren van een studie naar de overeenstemming tussen beoordelaars, op basis van het vergelijken van rangorden. Het thans gebruikte aantal van acht proefpersonen is aan de lage kant voor vergelijking met Kendall's tau, met kortere protocollen is het minder bezwaarlijk een groter aantal te nemen. De tijd voor een protocol kan ook korter worden door de proefpersonen de opgaven van tevoren te verstrekken ter aandachtige lezing en de leesperiode niet in het protocol te betrekken.

Over de wenselijkheid van de aanwijzing "begin bij het gevraagde" kan het volgende worden geconcludeerd. Degenen die al elementen van expertgedrag vertonen, zouden wellicht beter af zijn met een aanwijzing als "bekijk de ge-

gevens, maak een analyse en probeer op grond daarvan een goed plan voor eerste aanpak van het probleem te vinden". Voor degenen die nog geen strategische kennis of verkortingen op hun repertoire hebben, lijkt een dergelijke aanwijzing niet adequaat en kan beter worden geadviseerd bij het gevraagde te beginnen. In concreto zou deze aanwijzing in 4 VWO raadzaam zijn, maar in universitair onderwijs niet. Voor 6 VWO zou de strategie per student moeten worden vastgesteld.

Noten

1. Het artikel gaat over natuurkundeproblemen. De ervaring is dat één en ander zeker ook op scheikunde van toepassing is.
2. De reden dat beginners niettemin pogingen doen op een primitieve wijze vanuit de gegevens te werken, zou kunnen zijn dat zij hun leermeesters dat vaak zien doen: pseudo-expert-gedrag (Elshout, 1982). Docenten, die vraagstukken voormaken, transformeren vaak gegevens op een wijze die later blijkt de juiste te zijn. Zij zouden er beter aan doen aan studenten te laten blijken dat ze wel degelijk eerst een plan opstellen.
3. Op diverse plaatsen wordt dit begrip afwijkend omschreven, bijvoorbeeld bij Ferguson-Hessler (1989), die met 'strategische kennis' juist de niet-domeinspecifieke kennis bedoelt, die bij probleemoplossen wordt gebruikt. Ook is de benadering anders dan bij De Bruijn (1988), waar de met oefening verkregen of de door oefening veranderde domeinspecifieke kennis wordt opgevat als procedurele kennis.
4. Voor zover dat nagegaan kan worden, werd in alle onderwijssituaties, waarin deze studenten zijn geweest, de benadering van de *centripetale* kracht gebruikt. Het terugvallen op de centrifugale kracht lijkt daarom een gevoelsmatige kwestie.
5. Het verschil tussen beide scoremethoden is vooral bij pp.nr.10 opvallend: hoewel deze beide vraagstukken uiteindelijk oplost, met veel hulp en langzaam, vertoont hij geen enkel teken van expertgedrag.

Literatuur

- Berg, J.S. van den (1983). *Natuurkunde vraagstukken oplossen; een vakdidactische studie van het leren oplossen van natuurkunde vraagstukken in klas vier*. Eindhoven: Technische Universiteit (diss.).
- Blalock, H.M. (1960). *Social statistics (2nd ed)*. Tokyo: McGraw Hill/Kogakusha.
- Bruijn, I. de (1988). *Bijwerken in het Begin*. (Remedial Teaching at the Start). Enschede: Universiteit Twente (diss.).
- Bruijn, I. de (1991). *Cognitieve Elementen bij het Oplossen van Natuurkunde Problemen*. Enschede: Universiteit Twente (Memo Natuurkunde Didactiek nr.2).
- Chi, M.T.H., M.Bassok, M.W.Lewis, P.Reimann & R.Glaser (1989). Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems. *Cognitive Science*, 13, 145-182.
- Chi, M.T.H., P.J.Feltovitch & R.Glaser (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.

- Cooper, G. & J.Sweller (1987). Effects of schema acquisition and rule automation on mathematical problem solving transfer. *Journal of Educational Psychology*, 79, 347-362.
- Elshout, J.J. (1982). Het leren oplossen van problemen. *Losbladig Onderwijskundig Lexicon*.
- Ferguson-Hessler, M.G.M. (1989). *Over kennis en kunde in de fysica*. Eindhoven: Technische Universiteit (diss.).
- Gick, M.L. & K.J.Holyoak (1983). Schema Induction and Analogical Transfer. *Cognitive Psychology*, 15, 1-38.
- Greeno, J.G. (1976). Natures of Problem Solving Abilities. In W.Estes (Ed.), *Handbook of learning and cognitive processes (Vol.5)*. Hillsdale, NJ: Erlbaum.
- Heller, J.I. & F.Reif (1984). Prescribing effective human problem solving processes: problem description in physics. *Cognition and Instruction*, 1, 177-216.
- Larkin, J.H. (1979). Information processing models and science instruction. In J.Lochhead & J.Clement (Eds.), *Cognitive Process Instruction*. Philadelphia: Franklin Institute Press.
- Larkin, J.H., J.McDermott, D.P.Simon & H.A.Simon (1980). Models of competence in solving physics problems. *Cognitive Science* 4, 317- 345.
- Mettes, C.T.C.W. & A.Pilot (1980). *Over het leren oplossen van natuurwetenschappelijke problemen*. Enschede: Universiteit Twente (diss.).
- Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, 4, 457-466.
- Taylor, J.C. & G.Evans (1985). The architecture of human information processing: Empirical evidence. *Instructional Science*, 13, 347-359.
- Vanderlocht, M. & J. van Damme (1989). Fouten en knelpunten bij het oplossen van problemen over projectielbewegingen; een analyse van 152 schriftelijke oplossingen. *Tijdschrift voor didactiek der β -wetenschappen*, 7, 115-137.

Bijlage

De proef-opgaven.

Bij beide opgaven mag voor de gravitatieversnelling $g = 10 \text{ m.s}^{-2}$ worden genomen.

I. De konische slinger.

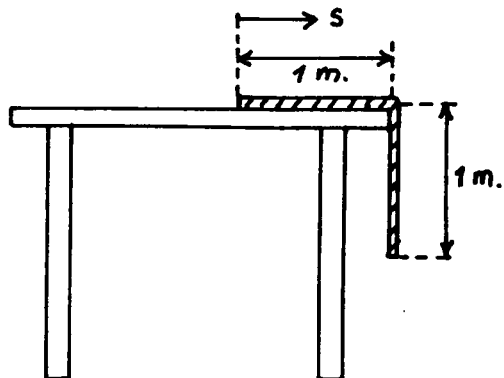
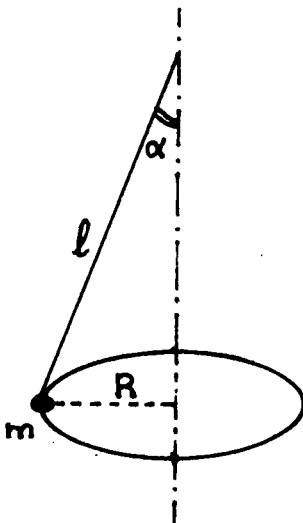
Een massa m is opgehangen aan een draad van lengte l . Op de juiste wijze in beweging gezet beschrijft de massa cirkels met straal R . De hoek tussen de draad en de verticale richting is (steeds) α .

a. Welke krachten werken op m tijdens deze cirkelbeweging? Zijn dit alle 'afzonderlijke' krachten, of is (b.v.) één ervan de resultante van twee (of meer) andere?

b. Bewijs dat voor de hoeksnelheid ω van m geldt:

$$\omega^2 = \frac{g}{\sqrt{l^2 - R^2}}$$

c. Kan m slechts in één cirkelbaan lopen, of zijn er meer mogelijkheden? Indien er meer banen mogelijk zijn, is de hoeksnelheid in een 'hogere' baan dan groter dan, gelijk aan of kleiner dan de hoeksnelheid in de lagere baan?



II. *Het wegglijdende touw.*

Een touw van 2 kg en een lengte van 2 m hangt over de rand van een tafel: half op de tafel, half over de rand. Op het gehele touw werkt de zwaartekracht F_z . Verder werkt op het gedeelte dat op de tafel ligt een wrijvingskracht, althans wanneer het touw beweegt of 'tracht' te bewegen. De (maximale) wrijvingskracht hangt ervan af hoeveel touw er op tafel ligt en bedraagt 3 Newton per meter touw (op de tafel).

a. Bereken met welke versnelling het touw begint te glijden.

Nu wordt ook aan het ondereind van het touw een kracht F , naar beneden, uitgeoefend. De verplaatsing door het wegglijden wordt nu, gedurende enige tijd, gegeven door:

$$s = 3 \cdot t^2$$

b. Bereken welke grootte F dan in het begin van de beweging moet hebben. (aanwijzing: dan is $s=0$ en er ligt nog 1 m touw op de tafel)

c. Bereken F als functie van de tijd, zolang $s = 3 \cdot t^2$ geldt.