

Voor mijn ouders

There is a vast amount more in most heads (...) than we are usually aware of, or that we are willing to try to use. You have to convince students (...) of the fact that there are implicit models in their heads which are useful.

J.S. Bruner

In L.S. Shulman en E.R. Keislar (Eds.).
Learning by discovery, a critical appraisal.
Chicago: Rand McNally, 1966, p. 105.

0 1 2 3 4 5 / 87 86 85 84 83 82

© 1982 Van Walraven bv, Apeldoorn, The Netherlands

Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook zonder voorafgaande schriftelijke toestemming van de uitgever.

ISBN 90 6049 351 6

ACTIVATIE VAN VOORKENNIS, INTRINSIEKE MOTIVATIE EN DE VERWERKING VAN TEKST

Studies in probleemgestuurd onderwijs

PROEFSCHRIFT

ter verkrijging van de graad van doctor in de geneeskunde aan de Rijksuniversiteit Limburg te Maastricht op gezag van de rector magnificus, prof. dr. H.C. Hemker, volgens besluit van het College van Dekanen in het openbaar te verdedigen op donderdag, 28 oktober 1982, 's middags om vier uur precies, in de aula van de universiteit

door

Hendricus Gerard Schmidt
geboren te Heerlen

VAN WALRAVEN BV / APELDOORN

Promotoren: Prof. dr. J.C.J. Bonarius en Prof. dr. W.H.F.W. Wijnen

Referenten: Prof. dr. A.P.W.M. Appels en Prof. dr. H.F.M. Crombag

INHOUDSOPGAVE

Voorwoord	7
Inleiding	9
Hoofdstuk 1	
Problem-based learning: rationale and description.	16
Verschijnt in <i>Medical Education</i> , najaar 1982.	
Hoofdstuk 2	
Activatie en herstructurering van voorkennis en hun effect op de verwerking van tekst: gegevens uit recall	31
Ter publikatie aangeboden aan het <i>Nederlands Tijdschrift voor de Psychologie</i> .	
Hoofdstuk 3	
Activation and restructuring of prior knowledge and their effects on text processing: recognition and transfer.	70
Is verschenen in A. Flammer en W. Kintsch (Eds.), <i>Discourse processing</i> . Amsterdam: North-Holland, 1982.	
Hoofdstuk 4	
Intrinsieke motivatie en studieprestatie: enkele verkennende onderzoeken	85
Ter publikatie aangeboden aan <i>Pedagogische Studiën</i> .	
Hoofdstuk 5	
Effecten van structurering van patiëntenproblemen op leerresultaat en satisfactie bij studenten (samen met P.A.J. Bouhuijs)	113
Een bewerking van een bijdrage aan het <i>Congresboek Onderwijsresearchdagen</i> . Amsterdam: Vrije Universiteit, 1977.	
Hoofdstuk 6	
De relatie tussen studiejaar en tentamenresultaat (samen met M.L. de Volder en W.H. Gijsselaers)	123
Ter publikatie aangeboden aan het <i>Tijdschrift voor Onderwijsresearch</i> .	
Samenvatting.	132
Summary.	136
Curriculum vitae	140

VOORWOORD

Dit proefschrift bevat de rapportage van een aantal onderzoeken naar cognitieve en motivationele effecten van probleemgestuurd onderwijs, een benadering van het onderwijsleerproces die aan de Rijksuniversiteit Limburg tot ontwikkeling gebracht wordt.

De vorm waarin dat geschiedt — het proefschrift is een bundeling van zes artikelen over dat onderwerp — brengt met zich mee dat op een aantal plaatsen een zekere herhaling van informatie optreedt. De lezer kan dan ook in de hoofdstukken 3 en 4 alinea's overslaan zonder dat betekenisvolle informatie verloren gaat. Deze hoofdstukken rapporteren data uit twee experimenten die in hoofdstuk 2 uitgebreid beschreven zijn. Het betreft daarbij in het bijzonder passages waarin de methode van probleemanalyse beschreven wordt en passages die betrekking hebben op de in de experimenten gebruikte procedures.

Dit proefschrift zou niet tot stand gekomen zijn zonder de hulp en aandacht van diegenen die in de loop van de tijd betrokken zijn geraakt bij mijn pogingen door middel van onderzoek op bepaalde vragen een antwoord te vinden. In de eerste plaats moeten mijn collega's in de capaciteitsgroep Onderwijsontwikkeling en Onderwijsresearch genoemd worden, die door middel van raad en daad het onderzoek mede gedragen hebben.

Dat geldt speciaal voor Jos Moust, Betsy Stalenhoef en Hetty Snellen die als tutor of proefleider aan verschillende experimenten hun medewerking verleenden, en voor Els Boshuizen, Peter Bouhuijs, Titus Geerligs en Maurice de Volder die als redactieleden van de interne publikatiereeks 'Onderzoek van Onderwijs, Rijksuniversiteit Limburg' vele eerste versies van mijn manuscripten van waardevolle kanttekeningen voorzagen.

Marjo Stevens verdient bijzondere dank. Niet alleen trad zij als tutor op in twee experimenten, maar zij schreef ook de osmosetekst en construeerde de studietoetsen.

Directie en studenten van de scholengemeenschappen Jekerdal in Maastricht en De Wachtendonck in Heerlen verleenden bereidwillig alle medewerking aan het onderzoek. Datzelfde geldt voor enige generaties van medische studenten die zich zonder morren vele malen lieten toetsen of ondervragen.

Bernice Doove, Ronald Linden en Bert Kerkhofs verrichtten monnikenarbeid met de analyse van de recallprotocols ten bate van hoofdstuk 2.

De heer Pieters leverde aanzienlijke bijdragen aan de vormgeving van de verschillende Engelse teksten.

Ellij Meertens, Pauline Schoenmakers en Karin van Ee typten delen van het proefschrift. Ellij Meertens leverde in enkele weken een drukklare versie van de tekst.

Wim Gijselaers verleende assistentie bij enige statistische analyses en Gea Schmidt controleerde de teksten op fouten.

Wim en Rika Reuwer verleenden verschillende malen gastvrijheid op hotel-kasteel 'Neubourg' als ik meende slechts in volstrekte afzondering het werk te kunnen doen.

Ten slotte dank ik promotoren en referenten voor de mogelijkheden die zij mij hebben geboden in de beste Popperiaanse traditie te leren van mijn fouten. In het bijzonder wil ik Wynand Wijnen en Han Bonarius dankzeggen voor de stimulerende wijze waarop zij de leermeesterrol vervuld hebben.

Het thema van dit proefschrift

Richard Anderson constateert aan het slot van de conferentie 'Schooling and the acquisition of knowledge' dat ons onderwijs in veel gevallen het leren van geïsoleerde kennis bevordert (Anderson, 1977). Docenten moedigen dat in feite aan door bij tentamens van hun studenten te verwachten dat zij de tekst die ze bestudeerd hebben, zo getrouw mogelijk reproduceren. Dus leren die studenten de nieuw te verwerven informatie op een wijze die bemoeilijkt dat ze werkelijk geïntegreerd wordt in bestaande cognitieve structuren. Ze hebben die informatie wel beschikbaar in het geheugen, maar als een corpus alienum, zonder relaties met andere kennis. Dat leidt ertoe dat die kennis wel gereproduceerd kan worden, maar niet gebruikt om datgene waarop die kennis betrekking heeft, te begrijpen.

Anderson illustreert dat aan de hand van onderzoek van Driver en Easley (1969). Driver en Easley lieten middelbare scholieren het gedrag beschrijven van ballen die met behulp van een veer werden afgeschoten. Hoewel deze scholieren vertrouwd waren met Newtons mechanica – blijkens het feit dat ze diens terminologie gebruikten – was hun perspectief op wat er met die ballen gebeurde nog Aristotelieans. (Velen van hen waren bijvoorbeeld van mening dat voor een constante beweging een constante kracht nodig is.) Deze scholieren hadden kennis van de moderne mechanica opgedaan, die echter hun visie op de werkelijkheid om hen heen niet veranderd had.

Driver en Easley concluderen dat 'the student, not being a blank slate, has already developed many concepts from his experience with the physical world, which influence his understanding of the new evidence and new arguments. When faced with propositions from teachers, textbooks, or even experiments, which are not in keeping with his own conceptual framework, he either has to modify his own view or construct another conceptual framework, separate from the one generated from his earlier experiences, but which is required for the purpose of "playing the game of school" . . .' (Driver en Easley, 1969, p. 1.).

Dit 'spelen van het spel', zoals de school dat van iemand vraagt, is betreurenswaardig, omdat werkelijk begrijpen, onthouden en gebruiken van nieuwe informatie de interactie met, en de integratie in, bestaande kennisstructuren veronderstelt (Mayer, 1975). En dat niet alleen: iemands voorkennis bepaalt in belangrijke mate wat hij kan leren. David Ausubel zegt het zo: 'If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor influencing learning is what the learner already knows' (Ausubel, 1968, p. VI).

De zes artikelen die in dit proefschrift gebundeld zijn, variëren op het thema dat hierboven aan de orde is gesteld. Probleemgestuurd onderwijs wordt daarin voorgesteld als een benadering van het onderwijsleerproces die de integratie van nieuwe informatie in bestaande kennisstructuren bevordert. Probleemgestuurd onderwijs heeft de volgende kenmerken: Aan een kleine groep studenten wordt een probleem voorgelegd. Dat probleem bestaat uit de beschrijving van een aantal verschijnselen of gebeurtenissen. De taak van de studenten is verklaringen voor deze verschijnselen te bedenken. Zij doen dat door hypothesen te formuleren, gebaseerd op hun voorkennis. Deze activiteit wordt de 'analyse van het probleem' genoemd. Vervolgens krijgen ze nieuwe, probleemrelevante informatie ter bestudering voorgelegd, of zoeken die informatie zelf op in de bibliotheek.

De belangrijkste these van dit proefschrift is nu *dat de analyse van een probleem, voorafgaande aan de bestudering van nieuwe informatie, leidt tot activatie en herstructurering van bestaande cognitieve structuren, die op hun beurt de verwerking van die nieuwe informatie vergemakkelijken.*

In *hoofdstuk 1* wordt een beschrijving van de methode gegeven, en wordt een aantal van de veronderstellingen besproken die aan probleemgestuurd onderwijs – in het Engelse taalgebied problem-based learning genoemd – ten grondslag liggen.

De *hoofdstukken 2, 3 en 4* bevatten reportages van twee experimenten waarmee de volgende vraagstellingen onderzocht werden:

- leidt de analyse van een probleem tot activatie en herstructurering van bestaande cognitieve structuren?
- Wat is de invloed daarvan op de verwerking van nieuwe informatie?
- Welke cognitieve processen kunnen aansprakelijk gesteld worden voor de gevonden effecten?
- Heeft de analyse van een probleem ook motivationele consequenties?
- Welke relaties bestaan er tussen motivationele en cognitieve effecten van probleemgestuurd onderwijs?

Met deze laatste vragen is een neventhema van dit proefschrift aangestipt, namelijk de aard van de invloed die motivationele factoren op het leerproces uitoefenen. Dit onderzoek naar de rol die (intrinsieke) motivatie speelt bij de verwerking van tekst, werd oorspronkelijk begonnen om alternatieve, niet cognitieve, verklaringen voor eventueel te vinden cognitieve effecten onder controle te kunnen brengen, maar ontwikkelde zich alras tot een op zichzelf staand onderzoek, waarover in *hoofdstuk 4* gerapporteerd wordt. *Hoofdstuk 5* beschrijft een experiment waarin de *vormgeving* van aan studenten aangeboden problemen centraal staan.

Een probleem werd hiervoor beschreven als een verzameling verschijnselen of gebeurtenissen die studenten aanspoort verklaringen te bedenken en te zoeken. Vormt een dergelijk probleem alleen wel een voldoende heldere stimulus tot leren? Moet de leersituatie niet meer voorgestructureerd zijn om gewenste en optimale leerresultaten te produceren?

In *hoofdstuk 6* ten slotte wordt de rol van voorkennis (en intrinsieke motivatie) op een wat andere wijze bezien. Men kan zich namelijk afvragen in

hoeverre de belangrijke functie die voorkennis in het leerproces lijkt te spelen niet een typisch laboratoriumeffect is, dat in de praktijk van het onderwijs niet naspeurbaar is. Om deze vraagstelling te onderzoeken werden studieresultaten van vier cohorten Maastrichtse medische studenten verzameld en samenhangen met beide andere onderwijsvariabelen nader onderzocht.

Historisch overzicht

Het kan verhelderend zijn probleemgestuurd onderwijs geplaatst te zien in een bepaalde onderwijskundige traditie. Daarom wordt in deze inleiding een paragraaf gewijd aan een kort historisch overzicht.

Het gebruik van problemen als stimulus voor leren gaat al terug tot de Griekse wijsgeer Socrates. Volgens Plato placht Socrates groepjes leerlingen om zich heen te verzamelen, die hij in een vraag-antwoordspel bijvoorbeeld de stelling van Pythagoras liet bewijzen. Hedendaagse equivalenten kunnen gevonden worden in het werk van Dewey (1929), Katona (1940), Piaget (1969) en verschillende Russische onderwijskundigen (zie Van Oers, 1981). De benaderingen die echter het meest gezien kunnen worden als voorlopers van probleemgestuurd onderwijs zijn ontdekkend leren ('learning-by-discovery': Bruner, 1961; 1971) en de gevalsmethode ('case-study method': Fraser, 1931).

Ontdekkend leren

Learning-by-discovery kan misschien het treffendst gekarakteriseerd worden aan de hand van een citaat uit Bruner (1959): 'One experiment which I can report provides encouragement. It was devised and carried out by the research group with which I am associated at Harvard in collaboration with teachers in the fifth grade of a good public school. It is on the unpromising topic of the geography of the North Central States (. . .). We hit upon the happy idea of presenting this chunk of geography not as a set of knowns, but as a set of unknowns. One class was presented blank maps, containing only tracings of the rivers and lakes of the area as well as the *natural resources*. They were asked as a first exercise to indicate where the principal cities would be located, where the railroads, and where the main highways. Books and maps were not permitted and "looking up the facts" was cast in a sinful light. (. . .)

The discussion was a hot one. After an hour, and much pleading, permission was given to consult the rolled up wall map. I will never forget one young student, as he pointed his finger at the foot of Lake Michigan, shouting, "Yipee, Chicago is at the end of the pointing down lake". And another replying, "Well, OK: but Chicago's no good for the rivers and it should be here where there is a big city (St. Louis)". These children were thinking, and learning was an instrument for checking and improving the process. To at least a half dozen children in the class it is not a matter of indifference that no big city is to be found at the junction of Lake Huron,

Lake Michigan, and Lake Ontario. They were slightly shaken up transportation theorists when the facts were in.

The children in another class taught conventionally, got their facts all right, sitting down, benchbound. And that was that. (. . .) One group learned geography as a set of rational acts of induction — that cities spring up where there is water, where there are natural resources, where there are things to be processed and shipped. The other group learned passively that there were arbitrary cities at arbitrary places by arbitrary bodies of water and arbitrary sources of supply. One learned geography as a form of activity. The other stored some names and positions as a passive form of registration (Bruner, 1959, p. 187-188).'

Samenvattend zou men kunnen zeggen dat ontdekkend leren leerlingen confronteert met problemen, die zij vervolgens proberen op te lossen door middel van discussie met medeleerlingen. Volgens Bruner (1961) stimuleert ontdekkend leren een dieper inzicht in de werkelijkheid, geeft het gelegenheid cognitieve vaardigheden te oefenen, verhoogt het de intrinsieke motivatie tot leren en bevordert het de verwerking en het onthouden van informatie, omdat die informatie betekenis heeft gekregen voor de lerenden. De empirische gegevens die aangevoerd zijn ter ondersteuning van Bruners ideeën zijn echter nogal mager. Ze bestaan voornamelijk uit anekdotes van het type dat hierboven geciteerd is. Onderzoek met betrekking tot de learning-by-discovery-hypothese heeft over de hele linie teleurstellende uitkomsten opgeleverd (Shulman en Keislar, 1966), al worden zo nu en dan positieve resultaten gemeld (Egan en Greeno, 1973).

Volgens Mayer (1975) komt dat waarschijnlijk omdat learning-by-discovery enkel tot activatie van voorkennis leidt. Als daarop geen confrontatie volgt met nieuwe informatie, vindt geen leren plaats. Ontdekkend leren gaat echter impliciet van de veronderstelling uit dat leerlingen in staat zijn zelf de gewenste kennis 'af te leiden'.

De gevalsmethode

Ontdekkend leren is als onderwijsmethode mede gebaseerd op Piagets idee van de fasegewijze cognitieve ontwikkeling van kinderen en op Bruners eigen theoretische werk op het terrein van de begripvorming (Bruner, Goodnow en Austin, 1956).

De gevalsmethode daarentegen, is voortgekomen uit praktische ervaringen van docenten aan de Rechtenfaculteit van de Harvard universiteit en de Harvard Business School (Fraser, 1931). Deze docenten kwamen tot de conclusie dat hun studenten slechts marginaal in staat waren kennis die zij tijdens hun studie verworven hadden ook werkelijk toe te passen en dat curricula leersituaties zouden moeten verschaffen die studenten in staat stellen te leren hoe verworven informatie *gebruikt* kan worden.

De gevalsmethode heeft de volgende vorm: Studenten krijgen een concrete beschrijving van een praktijksituatie of -probleem voorgelegd. Deze beschrijving bevat alle feiten, meningen en verwachtingen die als uitgangspunt kunnen dienen voor een besluitvormingsproces (Grochla en Thom, 1975). Vaak

gaat de ter zake doende informatie vergezeld van irrelevante details om de oplossing van het probleem niet 'weg te geven'. Studenten in de economie krijgen bijvoorbeeld een grote hoeveelheid informatie over een bepaald bedrijf, met als opdracht zich voor te stellen dat zij de managers van die firma zijn die bepaalde beleidsbeslissingen moeten nemen. De taak van de studenten is een optimaal beleid uit te stippelen, rekening houdend met alle beperkende condities, en te laten zien waarom de gekozen aanpak rationeel is.

Cognitieve doelen die volgens Grochla en Thom (1975) bereikt zouden worden met behulp van de gevalsmethode zijn: het kritisch leren analyseren van problemen, het herkennen van complexe samenhangen, versterking van de kennis die voorhanden is door toepassing daarvan, en leren om probleemrelevante informatie te verzamelen.

Probleemgestuurd onderwijs

Aan de medische faculteit van de McMaster universiteit in Hamilton, Canada, ontwikkelde de neuroloog Howard Barrows aan het begin van de jaren zeventig samen met anderen een benadering die zowel gelijkenis vertoonde met ontdekkend leren als met de gevalsmethode en die hij 'problem-based learning' noemde (Neufeld en Barrows, 1974; Barrows en Tamblyn, 1980). Teleurgesteld door de resultaten van het medisch onderwijs — studenten beschikken over veel kennis maar zijn maar beperkt in staat die kennis adequaat te hanteren — kwam hij tot de overtuiging dat medische studenten hun kennis zouden moeten opdoen rondom (praktijk)problemen. Die problemen zouden moeten bestaan uit korte beschrijvingen van klachten en symptomen waarmee patiënten hun arts bezoeken. Studenten zouden dergelijke problemen moeten analyseren en moeten formuleren welke informatie nog nodig is om die problemen beter te begrijpen (en eventueel te verhelpen). Barrows verwachtte dat op deze wijze informatie uit verschillende vakgebieden door de lerende geïntegreerd zou worden en dat de resulterende kennisstructuren 'probleemgecentreerd' zouden zijn. Hij nam verder aan dat deze procedure studenten ertoe zou brengen vertrouwd te raken met het denkproces van de arts. En omdat de verwerving van kennis gedurende de hele studieperiode relatief zelfstandig plaatsvindt, zou deze benadering ten slotte studenten ook nog leren geconstateerde leemten in hun kennis op te vullen door middel van zelfstudie, ook na afloop van hun studie.

Barrows' voorkeur voor concrete beschrijvingen van patiënten als stimulus tot leren lijkt duidelijk beïnvloed door de gevalsmethode, waarin immers ook concrete gevallen object van studie zijn. Datzelfde geldt voor zijn benadrukken van het leren oplossen van problemen als het uiteindelijke doel van medisch onderwijs. Ook in de gevalsmethode neemt de aanpak van complexe problemen een belangrijke plaats in. Echter, de nadruk die Barrows legt op probleemanalyse als instrument voor, en doel van, problem-based learning, herinnert sterk aan Bruners opvattingen van leren vanwege

de rol die door de student geproduceerde verklarende hypothesen in het leerproces spelen.

De eigen bijdrage van Barrows ligt in zijn erkenning dat het verzamelen van nieuwe informatie *uit externe bronnen* deel moet uitmaken van het leerproces in probleemgestuurd onderwijs. In de discussie over ontdekkend leren is al aangeduid dat deze methode van de aanname uitgaat dat lerenden *zelf* de noodzakelijke kennis kunnen produceren die nodig is voor een goed begrip van het probleem waaraan gewerkt wordt. Gewoonlijk worden geen externe bronnen geraadpleegd. Het is echter hoogst onwaarschijnlijk dat zij daarin zullen slagen in geavanceerde kennisdomeinen als de geneeskunde. De gevalsmethode daarentegen is impliciet gebaseerd op de veronderstelling dat studenten reeds de beschikking hebben over alle kennis die nodig is om het probleem effectief aan te pakken en dat zij die kennis slechts hoeven te leren toepassen. In probleemgestuurd onderwijs zijn verwerving en toepassing geïntegreerd.

De belangrijkste bijdrage die de groep waartoe de auteur van deze dissertatie behoort, aan de verdere ontwikkeling van probleemgestuurd onderwijs heeft geleverd, is, naast het benadrukken van de rol van de probleem-analyse in het proces, waarschijnlijk de relativering geweest van de veronderstelling dat alleen *praktijk* problemen zich voor de aanpak zouden lenen. Ook andere, eenvoudiger verschijnselen dan de klachten en symptomen van een patiënt, kunnen het startpunt zijn voor leren; zoals bijvoorbeeld het gedrag van een bloedcel in water en in een zoutoplossing, waarvan in de volgende hoofdstukken sprake zal zijn.

Literatuur

- Anderson, R.C. The notion of schemata and the educational enterprise: general discussion of the conference. In R.C. Anderson, R.J. Spiro en W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Ausubel, D.P. *Educational Psychology, a cognitive view*. New York: Holt, Rinehart and Winston, 1968.
- Barrows, H.S. en R.M. Tamblyn. *Problem-based learning*. New York: Springer, 1980.
- Bruner, J.S. Learning and thinking. *Harvard Educational Review*, 1959, 29, 184-192.
- Bruner, J.S. The act of discovery. *Harvard Educational Review*, 1961, 31, 21-32.
- Bruner, J.S. *Toward a theory of instruction*. Cambridge, Mass.: Harvard University Press, 1971.
- Bruner, J.S., J.J. Goodnow en G.A. Austin. *A study of thinking*. New York: Wiley, 1956.
- Dewey, J. *The quest for certainty*. New York: Minton, 1929.
- Driver, R. en J.A. Easley. Autonomous dynamic thinking of young adolescent physics students. In J.A. Easley (Ed.). *The use of mathematics in science teaching*. Urbana, Ill.: University of Illinois, 1969.
- Egan, D.E. en J.G. Greeno. Acquiring cognitive structure by discovery and rule learning. *Journal of Educational Psychology*, 1973, 64, 85-97.
- Fraser, C.E. *The case method of instruction*. New York: McGraw-Hill, 1931.

- Grochla, E. en N. Thom. *Fallmethode und Gruppenarbeit in der betriebswirtschaftlichen Hochschulausbildung*. Hamburg: Arbeitsgemeinschaft für Hochschuldidaktik, 1975.
- Katona, G. *Organizing and memorizing*. New York: Norton, 1940.
- Mayer, R.E. Information processing variables in learning to solve problems. *Review of Educational Research*, 1975, 45, 525-541.
- Neufeld, V.R. en H.S. Barrows. The McMaster philosophy: an approach to medical education. *Journal of Medical Education*, 1974, 49, 1040-1050.
- Oers, H.J.M. van. Zelfstandige kennisverwerving als aspect van wetenschappelijk denken. *Pedagogische Studiën*, 1981, 58, 420-432.
- Piaget, J. *Psychologie et pédagogie*. Parijs: Denoel, 1969.
- Shulman, K.S. en E.R. Keislar. *Learning by discovery: a critical appraisal*. Chicago: Rand McNally, 1966.

PROBLEM-BASED LEARNING: RATIONALE AND DESCRIPTION

H.G. Schmidt

Problem-based learning claims that it integrates, more fully than other instructional strategies, acquisition and application of knowledge, by using a problem-solving format for learning. In this contribution the validity of this assertion is examined. Three conditions that facilitate learning are discussed and a short description of the process of clinical reasoning is given. The question is answered to what extent problem-based learning meets the conditions set for learning and simulates the clinical reasoning process. Some empirical evidence regarding the approach is presented.

In the literature on the concept, problem-based learning is presented as a useful didactic procedure which solves some major problems of medical education (Barrows and Tamblyn, 1980; Neame, 1981; Neufeld and Barrows, 1974; Schmidt and Bouhuijs, 1980).

The first is the problem of *irrelevance* of part of the knowledge medical students are forced to acquire in the course of their education. Even where there is agreement about the goals of medical education, e.g. training students to become physicians who are well adapted to the practical needs of society, it is often not easy to decide which subject-matters are important in reaching these goals and which are not. The use of practical problems as a starting point for learning is supposed to provide a means of deciding on that matter (Barrows and Tamblyn, 1980).

A second problem is the *lack of integration* between the materials from different disciplines which contribute to the curriculum. Physicians confronted with a patient complaining of pain on the chest must be able to simultaneously and efficiently use information from anatomy, (patho)-physiology, biochemistry, psychology, in conjunction with several practical clinical skills in order to understand the etiology of the complaint and treat the cardiologic problem. The classical curriculum, however, provides few opportunities to medical students to integrate knowledge from different domains into coherent cognitive structures.

A third problem of medical education is *the need for continuing education*. Few to no measures are usually taken to prepare students for lifelong learning. By giving them responsibility for formulating their own learning goals — as is done in problem-based learning — teachers create conditions favourable for the development of competence and a predisposition to-

wards learning in students (Neufeld and Barrows, 1974).

This paper concentrates on a fourth problem.

Educational institutions are providing students with many opportunities to acquire knowledge. Training in the use of that knowledge is often given less room. Sometimes, educators even do not seem to make a distinction between the *acquisition* of knowledge and the *use* of that knowledge. If a person is in the possession of information, he is assumed to be able to use it. If you pass your examinations in medicine, you are a doctor ready for practice.

This conclusion, however, can be seriously questioned. Gonella, Goran, Williamson and Cotsonas (1970), for instance, monitored the care of a group of patients suspected of having a urinary tract infection. They noted in which of the cases the physicians (residents or fully trained) conducted critical evaluations with respect to the problems and what additional information was acquired, given the suspicion of pyelonephritis. After completing this monitoring of care in an unobtrusive way, Gonella and colleagues asked the physicians involved to complete a multiple-choice examination dealing with diagnosis, treatment and follow-up of patients with urinary tract infection.

The results showed that in only about 50% of the cases had the staff acquired information from the patients concerned relevant to bacteriuria. On the other hand, the mean score of correct answers on the multiple-choice test pertaining to that same subject was 82%. The main conclusion that has to be drawn from these results is that these medical people had knowledge they did not use.

Other examples can be drawn from the non-medical educational literature. Driver and Easley (in: Anderson, 1977) for instance, conducted research among high school physics students. Interviews were conducted regarding the movements of balls along a horizontal track. According to Driver and Easley many of the students manifested 'an Aristotelian notion that constant force is required to produce constant motion ... while they used the language of Newtonian mechanics, such as force, momentum, impulse, etc.'. These students were acquainted with the concepts of modern mechanics, but were unable to use them in understanding real-world phenomena.

Willems and his co-workers (Willems, 1978; 1981) established essentially the same phenomenon in investigations among students of law, social geography, planning science and sociology. When writing a report on some topic or when solving problems, these students seem to be unable to use the knowledge they had acquired.

The main conclusion that has to be drawn from these results is that people can possess knowledge they seem unable to apply. They know things but cannot use them. This conclusion is in accordance with findings in educational psychology. Mayer and Greeno (1972) for instance, show that different instructional methods produce different learning outcomes. The way in which a topic is taught determines what students can do with the

information acquired. In short, conventional instructional procedures do not always enable students to learn how to make an appropriate use of their knowledge (e.g. to solve problems).

Problem-based learning claims that it integrates, more fully than other instructional strategies, acquisition and application of knowledge, because problem-solving and learning go hand in hand in the process of problem-based learning (Barrows and Tamblyn, 1980; Engel and Clarke, 1979).

In this contribution the validity of this assertion will be examined. In order to be able to do so I will first make a few remarks about human learning in general. In the next section three conditions that facilitate learning are discussed. The extent to which instructional methods meet these requirements determine their effectivity in fostering the acquisition and use of knowledge.

Next, recent conceptions of medical problem-solving are reviewed.

Finally, the process of problem-based learning will be described in some detail and discussed with reference to the remarks made in the preceding sections.

Three Conditions That Facilitate Learning

In order to be able to evaluate the characteristics of problem-based learning one needs a frame of reference. This frame of reference can be provided by the information processing approach to learning (Anderson, 1977). Three principles deduced from the information processing conceptualization are relevant here: activation of prior knowledge, encoding specificity and elaboration of knowledge. First, I will discuss the role of prior knowledge in learning.

Learning, by its nature, has a restructuring character. It presupposes earlier knowledge that is used in understanding new information and that is changed by this new information. For instance by studying an article about defence mechanisms against infections, a student applies knowledge he already possesses in order to understand the new information. A first-year medical student, while reading and interpreting the article, will probably make use of his secondary-school knowledge of biology, perhaps supplemented with common sense knowledge on bacteria, viruses, influenza and vaccinations. A fourth-year student will, in studying this article, use his earlier acquired immunological and microbiological knowledge. This prior knowledge and the kind of structure in which it is available in long-term memory, will determine what is understood from the article and this, in turn, will define what is learned from it (Rumelhart & Ortony, 1977). This relationship between prior knowledge and learning could be demonstrated with first- and fourth-year medical students, who read the same article and have the same amount of reading time available to process it. The fourth-year students will show better learning results because their

more elaborated prior knowledge will enable them to process the new information easier. Experimental evidence for this assertion is provided by Dooling and Lachman (1971) and Bransford and Johnson (1972).

Instructional methods, however, differ in their capacity to induce activation of relevant prior knowledge (Mayer and Greeno, 1972). Only in as far as they help students in activating relevant knowledge they will facilitate the processing of new information (Mayer, 1982).

The second condition that facilitates learning has been derived from the work of Tulving and Thomson (1973). They state that successful retrieval of information on some point in the future is promoted when the retrieval cues that are to reactivate that information are encoded together with that information. In other words: The closer the situation in which something is learned resembles the situation in which the learned material should be applied, the better the performance. This phenomenon is called *encoding specificity*. It is practiced in medical education during clinical lectures or clerkships, where students acquire knowledge related to patient problems that have characteristics in common with what students have to expect in later professional life.

The third condition has to do with elaboration of knowledge. Not until recently psychologists discovered that information is better understood, processed and retrieved if students have an opportunity to *elaborate* on that information (Anderson and Reder, 1979). In education there are many ways in which this condition can be fulfilled. Students can elaborate on information by answering questions about a text (Anderson and Biddle, 1975), by taking notes (Peper and Mayer, 1978), by discussing subject-matter to be learned with other students (Rudduck, 1978), by teaching peers what they have first learned themselves (Bargh and Schul, 1980), by writing summaries (Wittrock, 1974), and by formulating and criticizing hypotheses, given a problem (Schmidt, 1982). According to Reder (1980) elaborations provide redundancy in the memory structure. Redundancy can be viewed as a safeguard against forgetting and an aid to fast retrieval. In conclusion: education should, in order to optimize learning, help students in activating relevant prior knowledge, provide a context that resembles the future professional context as close as possible and stimulate students to elaborate on their knowledge. If one or more of these conditions are not met, instruction will produce suboptimal results.

Medical Problem-solving

Ever since Rimoldi (1961) published his article about the measuring of diagnostic skills, physicians and psychologists concerned with the training of medical students have been pondering ways of teaching medical problem-solving. These efforts have been paralleled by an increased interest in the way experienced physicians solve clinical problems. Much of the research on medical problem-solving has been reviewed in Elstein et al.

(1978), a landmark in this field. The available evidence on the process of clinical reasoning converges to the type of cybernetic model summarized in table 1 below. It is explained in table 2 and illustrated by a naturalistic description of a physician's thought processes, adapted from Barrows and Tamblyn (1980).

Table 1: The course of a problem-solving process

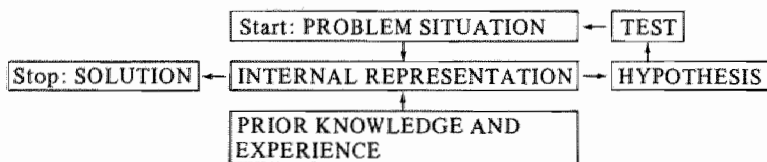


Table 2: Explanation and illustration of a cybernetic model of medical problem-solving

Description of the process	Example, adapted from Barrows and Tamblyn (1980)
There is a problem, or a PROBLEM SITUATION , perceived as such by the thinker.	The 67-year-old man, sitting forward in his chair, explains the reason for his visit. 'Doctor, since Saturday morning I've been very short of breath.'
On the basis of information directly deriving from the problem situation and on the basis of PRIOR EXPERIENCE the thinker forms a mental image, the INTERNAL REPRESENTATION of the problem.	'An old man' the doctor thinks, 'short of breath; sitting forward in his chair; slightly red in the face.'
This internal representation automatically generates a number of expectations or implicit HYPOTHESES about the nature of the problem and how it might be solved.	'Congestive failure?' 'Pneumothorax?' 'Pulmonary embolus?' 'Heart?'
Next the thinker TESTS these hypotheses, one after the other or simultaneously. He attempts to discover whether his hypotheses fit the problem situation. (He does so only for the two or three most likely hypotheses.)	Doctor: 'Did it start suddenly?' Patient: 'Yes.'
The outcome of the tests is fed back into the INTERNAL REPRESENTATION of the problem.	'Old man, short of breath, sitting forward, sudden onset, red in the face.'
Resulting in new HYPOTHESES .	'Arrhythmia?'
And new TESTS .	Doctor takes pulse and asks: 'Were you doing something when it started?' Patient: 'Sawing wood.'
This process is repeated until the thinker finds (or thinks he has found) the SOLUTION to the problem.	Doctor formulates final diagnosis.

Medical problem-solving might thus be defined as the transformation of a problem into a solution by means of the testing of hypotheses. In the area of medicine, problem-solving is the transformation of a problem situation, consisting of complaints and directly discernible symptoms, into a solution, namely the identification of the disease or other factors responsible for those complaints and symptoms. The transformation takes place with the help of tests in the form of questions, physical examination, and laboratory investigation, on the basis of the physician's hypotheses concerning the problem and its solution. These hypotheses are in turn based on prior knowledge and experience. Problem-solving is a dynamic process, involving continual cognitive interaction between problem situation, internal representation, memory and hypotheses arising from the combination of these factors.

Medical problem-solving requires great skill. If problem situations in this field were clear and unambiguous, computers might be able to take over the work, as indeed is sometimes proposed. However, the problem situation is nearly always obscured by the fact that patients react individually both to the doctor and to the signals of their own bodies. What someone experiences as a dim pain in the loins, someone else will report as a stinging feeling in the stomach. Moreover, people tend to suppress some signals and to overreact to others. As a result, the complaints they take to the doctor are often interpretations rather than objective representations of their problems.

Useful analyses of the medical problem-solving process are given by Gorry (1970), Elstein et al. (1972), and Barrows et al. (1978). The present state of affairs in research on problem-solving has been reviewed by Kleinmuntz (1966), Newell and Simon (1972), and Weisberg (1980).

Problem-based Learning: Description of the Process

In order to be able to assess to what extent problem-based learning simulates the clinical reasoning process and meets the requirements stated above, it is necessary to take a closer look at it. That will be done by means of an example that is borrowed from Schmidt and Bouhuijs (1980). The process of problem-based learning starts with a problem like the following: A plumber sees his doctor with the following complaint: 'During a hard cough this morning I suddenly tasted blood in my mouth. As this has occurred more often these past few weeks, I'm becoming a bit anxious.' The task of a group of students is to explain these phenomena in terms of underlying processes. Explanations take the form of physical, biological or psycho-social mechanisms that may be underlying the phenomena concerned.¹

¹ Of course, also another task can be formulated for the group, namely to ask themselves what they would *do* if they were the doctor of the person described. What

While working on the problem, the study group is using a systematic procedure to analyze the problem, to formulate learning objectives and to collect additional information. This procedure consists of seven steps² and is summarized in table 3.

Table 3: Steps involved in problem-based learning

Step 1: Clarify terms and concepts not readily comprehensible.
Step 2: Define the problem.
Step 3: Analyze the problem.
Step 4: Draw a systematic inventory of the explanations inferred from step 3.
Step 5: Formulate learning objectives.
Step 6: Collect additional information outside the group.
Step 7: Synthesize and test the newly acquired information.

The first activity in relation to any problem is clarification of terms and concepts not understood on first sight. Use may be made of group members' relevant knowledge. In some instances also a dictionary lends a helping hand, whereas in others the first step consists in reaching concord of opinion regarding the meaning to be attached to the various terms. In the latter case, therefore, the aim is attunement of interpretation rather than clarification of terms. If a group of students dealing with the 'oral blood loss' problem should happen to be unfamiliar with a plumber's duties, the dictionary will tell them that the man engages in the 'installation and maintenance of water-conduits and sewerage systems in and about the building'. The second step produces an exact definition of the problem. The group, as a whole, is now to reach agreement as to which inter-related phenomena should be explained. More often than not, problems offered do not present difficulties in this respect. Occasionally, however, a problem has been purposely structured to present difficulties to students as far as recognition of relevant symptoms is concerned. Sometimes a problem consists of a series of secondary, independent problems; these must be identified as such prior to further completion being possible. This shows that the question to be answered in order to define the problem is: which phenomena should be explained?

The analysis of the problem is the third step. Careful perusal of the text

kind of questions they would ask, what lab tests they would order, what physical examination they would perform and what treatment they would suggest. This task is secondary to that of explaining the symptoms, because understanding precedes rational action, both in education and in practice.

² Because this procedure consists of seven steps it is called 'Zevensprong' (Seven jumps) after an old Dutch nursery rhyme. This denomination helps students, at least in the Netherlands, in retrieving the procedure for use.

and forming as vivid as possible a representation of the situation described, will soon result in ideas and suppositions about the structure of the problem. These are either based on prior knowledge of students or are produced by rational elaborations. Hence, analysis of the problem substantially consists in recapitulation of group members' opinion, actual knowledge, and ideas with respect to the underlying processes and mechanisms. In this process a group will not confine itself to activation of prior knowledge, but will also try to formulate problem-relevant hypotheses by means of elaboration.

Most of the time a kind of free-association-round is held, in which each individual is allowed to verbalize freely his or her ideas with respect to the problem at hand before the presented ideas, knowledge and suppositions are scrutinized, accepted, completed or modified. This procedure is often referred to as the 'brainstorming' technique. Application of analytical efforts of this kind to the oral blood loss problem may yield the following line of thought.¹

'Could we have a case of lung cancer here? The site of origin of the oral blood need not at all be the lungs. Injury may have occurred at some point in the throat or in the mouth itself. But how? I think that coughing subjects all sorts of blood vessels to a pressure increase, for people who are coughing vigorously often turn red in the face. Does not that point to a broken vessel? For that matter, why does one cough? The reason may be the presence of an irritant in the air passages, an inspired peanut, or phlegm. Coughing is a reflexive mechanism activated to clear the respiratory tract and expel obstructions ...

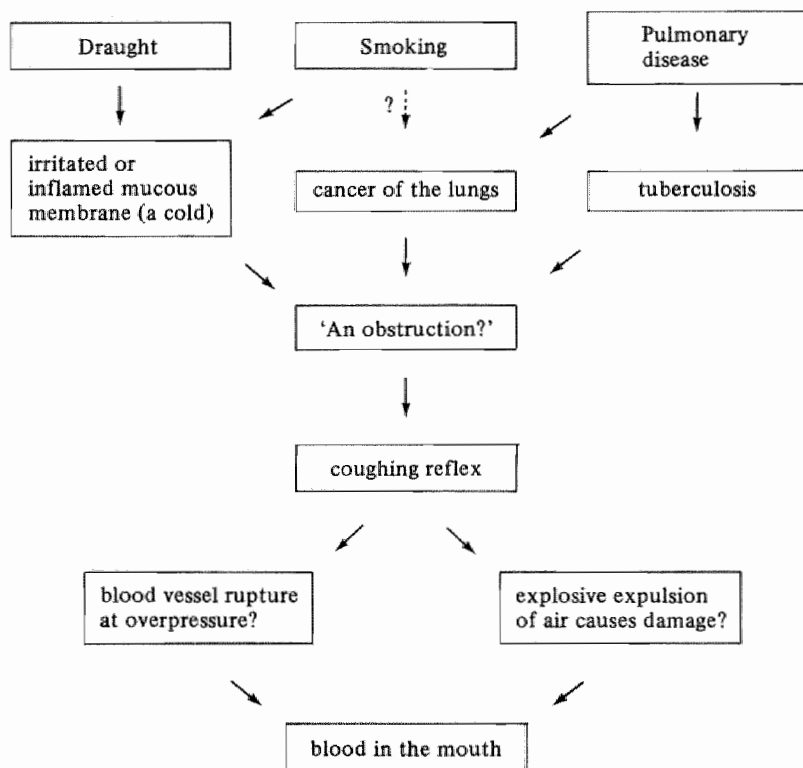
People suffering from tuberculosis sometimes cough up blood. The person figuring in the problem has had this for two weeks now, the oral blood loss, I mean. I wonder if he has been coughing much longer already. Perhaps he is a smoker. But why this blood? I can imagine that the sudden explosive expulsion of a stream of air results in such a high air pressure that something gets damaged. A tumor? Might very well be the obstruction creating the coughing stimulus. The man is a plumber, which means of course, that he frequently works in draughty rooms and spaces. He may also have caught a severe cold then, an ... eh ... inflammation of the mucous membrane in the respiratory system caused by a virus or something similar. In my opinion certain glands in the mucous membrane are secreting an abnormal amount of mucus to protect the body against the virus. But, again, what about the blood, eh? ...

By the way, cancer of the lungs is a growth of malignant cells, malignant lung tissue. It expands in the lungs, so that they can no longer function properly. Well, I am not sure about the exact process. Tuberculosis? That is a disease brought about by the tubercle bacillus. An inflammation of the lungs or something like that.'

¹ The next passage is an adaptation of a verbatim transcript of an actual problem-analysis by a group of Dutch first-year medical students.

In the fourth step a systematic inventory is made of the various explanations of the problem. Considering the above analysis by a group of medical students, one sees that several general descriptions of biological processes were produced. The first one relates to the students' ideas about the origin of the coughing ('a reflexive mechanism activated to clear the respiratory tract and expel obstructions'). The second covers the possible connection between coughing and a bleeding ('I can imagine that the sudden explosive expulsion of a stream of air results in such a high air pressure that something gets damaged'). The other descriptions refer to the many obstructions that may have entered the air passages ('inflammation of the mucous membrane', 'cancer of the lungs', 'tuberculosis'). The systematic inventory might yield the following scheme: see figure 1.

Figure 1: Schematic presentation of the oral blood loss problem



An inventarisation like this acts as a kind of summary and structures the product of the problem-analysis.

The activities have now reached the point that the assumed processes and mechanisms referred to in the analysis are to be studied more extensively. To what degree can the expressed knowledge and ideas be considered correct and complete? Study priorities have to be established because, as a rule, there is not enough time to pay adequate attention to all of the aspects involved. In other words, the fifth step required formulation of learning objectives. Learning objectives are the answers to the questions evoked by the problem-analysis, the questions being directed at provision of a more profound knowledge of the processes forming the crux of the problem. In our example i.e. the following questions might be considered:

- What physiological actions are involved in the coughing reflex?
- In what way does draught affect the air passages?
- How exactly does a bleeding result from coughing?
- What causes mucous membrane to become inflamed and what does the membrane look like in this state?
- What is the exact effect of smoking on the respiratory tract?
- What exactly is lung cancer?
- What is tuberculosis?

The group selects the objectives on which it will concentrate its activities and, if necessary, agrees on a distribution of tasks. Finally it tries to find out which learning resources might supply the required answers.

The sixth step consists of individual study. The group members individually collect information with respect to the learning objectives. The source explored may be literature, but other tools as well. For instance, audio-visual aids might be employed. Also resource persons can be consulted about aspects of the problem not yet clarified.

The process of problem-based learning is completed by synthesizing and testing the newly acquired information. The students inform one another about their individual findings, supplement their knowledge this way, and correct it wherever necessary. It is then established whether the group is capable of giving more detailed descriptions of the processes that play, or may play, a fundamental role in the problem. Often, in the course of this exchange of information, new questions are touched upon. In that case the group may decide to take up the process of problem-based learning – now at a deeper level of understanding – starting with the fourth step. However, it may prove equally effective in that phase to exploit the newly acquired knowledge in tackling a different, though interrelated, problem in order to find out whether the group can now perform the analytical and explanatory task in a faster and better way.

Discussion

The central topic of this contribution was to assess whether the claim was correct that problem-based learning integrates knowledge acquisition and knowledge application in problem-solving.

First some remarks about problem-based learning and clinical reasoning. One can certainly consider problem-based learning as a kind of simulation of the cognitive behavior of physicians solving the problems of their patients. Like physicians, students formulate hypotheses and test them. Both use prior knowledge activated by the concrete medical problem to formulate these hypotheses, and find ways to test them. From this perspective one could say that students working within a medical curriculum that incorporates problem-based learning as its prevalent instructional method train themselves directly for future professional life. There are, however, striking differences between medical problem-solving as practised by the physician and problem-based learning as practised by students.

The first difference pertains to the goal of these activities. The physician solves problems in order to help patients. The student solves problems to gain deeper insight into processes underlying problems and thereby to acquire new knowledge. This is why a student is not in the first place asked *to solve* the problem, that is: to formulate a plan of action for helping the patient. Rather a student is asked *to explain* the problem. Thus, in the case of the activities of a physician, problem-solving means helping people by diagnosis and management; in the case of the activities of a student within the problem-based learning format, problem-solving means analysis and elaboration to acquire increasingly deep understanding of underlying bio- or psycho-social processes.

The second difference relates to the nature of the prior knowledge student and physician use to solve the problem. The prior knowledge of a 'junior' student will be vague, broad, superficial and perhaps incorrect. The knowledge of the physician is the result of years of study and practice. As a result of this, a physician will 'see' the problem of oral blood loss in a different perspective and accordingly analyse it in a different way.

The third difference has to do with the process of hypothesizing itself. Hypothesizing by a physician is often implicit and subconscious. His thinking may not be open to inspection by others or himself. In problem-based learning, students are asked to make their hypothesizing explicit, to 'externalize' their thinking. As a result, their ideas, knowledge, conjectures and thought processes are open to discussion, elaboration and correction by peers. This in its turn can lead to further elaboration of knowledge, to the formulation of better or further questions, to improved thinking and to further self-directed learning. Externalizing of what a student thinks can become the basis for further learning.

Let's look at problem-based learning from the perspective of human learning. In the section on conditions that facilitate learning three principles

were discussed. It was concluded that instructional methods should stimulate students in activating relevant prior knowledge, provide a learning setting that is as similar as possible to the setting in which acquired knowledge is to be used (encoding specificity), and give students opportunities to elaborate on their knowledge.

Does problem-based learning meet these conditions?

With respect to the requirement that problem-based learning should lead to the activation of relevant prior knowledge one might say that it highly depends on characteristics of the problems being used. Written problems will activate relevant prior knowledge only if they have the following features (Schmidt, in: Willems, 1981):

1. They should consist of a neutral description of an event or a set of phenomena that are in need of explanation in terms of underlying processes, principles or mechanisms;
2. They really have to lead to problem-solving activity. If you ask students to study 'the heat regulating mechanism of the human body', this task will not lead to the activities described in the preceding section. But consider the following problem: 'You have been playing a game of tennis. You've a red face and are wet all over your body. How can these phenomena be explained?';
3. Problems have to be formulated as concrete as possible;
4. Problems should have a degree of complexity adapted to students' prior knowledge. If a problem is not complex enough, it will not be recognized as a problem. If it is too complex, students will hardly be able to formulate relevant hypotheses.

The second condition that facilitates learning (encoding specificity) is met by the problem-based learning approach in so far as problems are used that have a close resemblance to problems that will come across students in later professional life, e.g. in the case of medical education: patient or other health problems. Barrows and Tamblyn (1980) and Neame (1981) further specify this condition by proposing that problems be chosen that

5. have the greatest frequency in the usual practical setting;
6. represent life-threatening or urgent situations;
7. have a potentially serious outcome, in terms of morbidity or mortality, in which intervention — preventive or therapeutic — can make a significant difference in prognosis;
8. are most often poorly handled by physicians in the community.

What about the third facilitating condition: giving students opportunities to elaborate on their knowledge? Let's look at the process of problem-based learning again. The reader will remember that during problem-analysis students activate existing knowledge structures by formulating hypotheses regarding the mechanisms underlying the problem. This process can be rephrased as one involving elaboration on prior knowledge, because by 'externalizing' their thinking during group discussion, students provide each other with opportunities for enrichment, amplification and change of existing knowledge structures. When students are trying to explain the

blood loss problem by hypothesizing about possible processes that can be considered responsible for the phenomena observed, they are not merely reproducing knowledge they've acquired at some point in the past. They are using this knowledge as 'stuff for thinking'. In doing so, previously unrelated concepts become connected in memory, newly produced insights change the structure of their cognitions and information supplied by peers is added. The same applies to the last step of the problem-based learning procedure where newly acquired information is exchanged, critically discussed and eventually applied. These are the activities that can be considered as elaboration processes. From these considerations one may conclude that as far as systematic procedures are used for problem analysis and acquisition of problem-relevant knowledge — procedures that force students to process prior and new information as deep as possible (Craik and Lockhart, 1972) — problem-based learning satisfies the third facilitative condition of learning.

Up till here only the *potential* qualities of the problem-based learning approach as an instructional method were discussed. The question is whether there is any *empirical* evidence regarding the effectiveness of the approach. Referring back to the introduction of this article one can ask whether problem-based learning is really promoting the acquisition of knowledge that is applicable. Not very many data are available at the moment. Here I can only give the reader a cursory glance at some of the research on this topic that is executed at the department of educational development and research of the medical faculty at Maastricht in the Netherlands. In one investigation, Boshuizen and Claessen (1982) compared groups of students of the Maastricht medical faculty (where the problem-based learning approach has been implemented) with students from a more traditional medical faculty on a number of problem-solving tasks. Each subject was confronted with case-histories of two patients suffering from pancreatitis and prostatitis. Each of these case-histories was presented on a series of cards, each card containing a piece of information with respect to the patient. The subject had to read these cards and comment on the information. After processing the case-history in this way the subject was asked among other things to write down a differential-diagnosis, to rate the information on the cards in terms of its importance, and to recall the information. Some small but reliable differences were found in favor of the Maastricht medical students. This suggests that a problem-based curriculum may provide better opportunities for learning to solve medical problems.

In another investigation the present writer tested the problem-based learning hypothesis more directly (Schmidt, 1982). Small groups of subjects analyzed a written problem on osmosis. Subsequently, they studied a text containing new information relevant to this topic. As compared with control subjects that only studied the text, the experimental subjects were better in recognizing the information from the text and in using it in order to solve small problems.

My conclusion is that the available theoretical notions and empirical evi-

dence suggest that problem-based learning indeed integrates knowledge acquisition and its application, and can be considered a useful amplification of conventional instructional methods, and perhaps be an alternative.

References

- Anderson, J.R. and L.M. Reder. An elaborative processing explanation of depth of processing. In L.S. Cermak and F.I.M. Craik (Eds.). *Levels of processing in human memory*. Hillsdale: Lawrence Erlbaum, 1979.
- Anderson, R.C. The notion of schemata and the educational enterprise: general discussion of the conference. In R.C. Anderson, R.J. Spiro and W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Anderson, R.C. and W.B. Biddle. On asking people questions about what they are reading. In G.H. Bower (Ed.). *The psychology of learning and motivation* (vol. 9). New York: Academic Press, 1975.
- Bargh, J.A. and Y. Schul. On the cognitive benefits of teaching. *Journal of Educational Psychology*, 1980, 72, 593-604.
- Barrows, H.S., J.W. Feightner, V.R. Neufeld and G.R. Norman. *Analysis of the clinical methods of medical students and physicians*. Hamilton, Canada: McMaster University, 1978.
- Boshuizen, H.P.A. and H.F.A. Claessen. Cognitieve verwerking en onthouden van patiëntgegevens. In H.G. Schmidt (Ed.). *Probleemgestuurd onderwijs*. Stichting voor Onderzoek van het Onderwijs. Harlingen: Flevodruk, 1982.
- Bransford, J.D. and M.K. Johnson. Contextual prerequisites for understanding: some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 717-726.
- Craik, F.I. and R.S. Lockhart. Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- Dooling, D.J. and R. Lachman. Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, 1971, 88, 216-222.
- Elstein, A.S., L.S. Shulman and S.A. Sprafka. *Medical problem-solving, an analysis of clinical reasoning*. Cambridge: Harvard University Press, 1978.
- Elstein, A.S., N. Kagan, L.S. Shulman, H. Jason and M.J. Loupe. Methods and theory in the study of medical inquiry. *Journal of Medical Education*, 1972, 47, 85-92.
- Engel, C.E. and R.M. Clarke. Medical education with a difference. *Programmed Learning and Educational Technology*, 1979, 16, 72-87.
- Gonella, J.S., M.J. Goran, J.W. Williamson and N.J. Cotsonas. Evaluation of patient care, an approach. *Journal of the American Medical Association*, 1970, 214, 2040-2043.
- Gorry, G.A. Modelling the diagnostic process. *Journal of Medical Education*, 1970, 45, 293-302.
- Kleinmuntz, B. (Ed.). *Problem-solving*. New York: Wiley, 1966.
- Mayer, R.E. Instructional variables in text processing. In A. Flammer and W. Kintsch (Eds.). *Discourse processing*. Amsterdam: North-Holland, 1982.
- Mayer, R.E. and J.G. Greeno. Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology*, 1972, 63, 165-173.
- Neame, R.L.B. How to: construct a problem-based course. *Medical Teacher*, 1981, 3, 94-99.
- Neufeld, V.R. and H.S. Barrows. The 'McMaster Philosophy': an approach to medical education. *Journal of Medical Education*, 1974, 49, 1040-1050.
- Newell, A. and H.A. Simon. *Human problem-solving*. Englewood Cliffs: Prentice Hall, 1972.

- Peper, R.J. and R.E. Mayer. Note taking as a generative activity. *Journal of Educational Psychology*, 1978, 70, 514-522.
- Reder, L.M. The role of elaboration in the comprehension and retention of prose: a critical review. *Review of Educational Research*, 1980, 5, 5-53.
- Rimoldi, H.J.A. The test of diagnostic skills. *Journal of Medical Education*, 1961, 36, 73-79.
- Rudduck, J. *Learning through small group discussion*. Guildford: Research into higher education monographs, 1978.
- Rumelhart, D.E. and E. Ortony. The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro and W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Schmidt, H.G. Activation and restructuring of prior knowledge and their effects on text processing. In A. Flammer and W. Kintsch (Eds.). *Discourse processing*. Amsterdam: North-Holland, 1982.
- Schmidt, H.G. and P.A.J. Bouhuijs. *Onderwijs in taakgerichte groepen*. Utrecht: Het Spectrum, 1980.
- Tulving, E. and D.M. Thomson. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.
- Weisberg, R.W. *Memory, thought and behavior*. New York: Oxford University Press, 1980.
- Willems, J. *Probleemgeoriënteerd (groeps)onderwijs*. Nijmegen: IOWO, 1978.
- Willems, J. Problem-based (group)teaching: a cognitive science approach to using available knowledge. *Instructional Science*, 1981, 10, 5-21.
- Wittrock, M.C. Learning as a generative process. *Educational Psychologist*, 1974, 11, 87-95.

**ACTIVATIE EN HERSTRUCTURERING VAN VOORKENNIS
EN HUN EFFECT OP DE VERWERKING VAN TEKST:
GEGEVENS UIT RECALL**

H.G. Schmidt

In dit artikel worden cognitieve effecten gerapporteerd van een methode van probleemanalyse. De taak van de deelnemers aan een dergelijke analyse — die in een kleine groep plaatsvindt — is een verzameling gegeven verschijnselen te verklaren in termen van een onderliggend proces.

Uit Experiment 1 bleek dat probleemanalyse een activerende en herstructurerende invloed uitoefent op probleemrelevante voorkennis die proefpersonen ter beschikking hebben. De probleemanalysegroep produceerde ruim twee maal zoveel proposities in de recall als een controlegroep. Een deel van die proposities bleek van andere deelnemers geleerd te zijn. De probleemanalysegroep produceerde daarnaast drie-en-een-half keer zoveel verklarende proposities.

In Experiment 2 werd de invloed onderzocht van activatie en herstructurering van voorkennis op de cognitieve verwerking van een tekst. De probleemanalysegroep produceerde bijna anderhalf keer meer proposities uit tekst dan de controlegroep. 25% van de variantie in de data werd door de experimentele manipulatie verklaard. Beide groepen verschilden niet in de wijze waarop informatie uit tekst in het lange-termijn-geheugen georganiseerd werd. Uit de resultaten werd afgeleid dat activatie van voorkennis de codering van nieuwe informatie beïnvloedt, eerder dan retrieval van die informatie. Het faciliterende effect van activatie werd toegeschreven aan een capaciteitsvergroting van het werkgeheugen tijdens het verwerkingsproces (Spilich, Vesonder, Chiesi en Voss, 1979). Een verklaring in termen van intensievere elaboratie gedurende het leerproces (Anderson en Reder, 1979) moest op grond van de frequentiepatronen van inferenties in de recall worden afgewezen.

Tekst neemt in de westerse cultuur een centrale plaats in als medium voor kennisverwerving en informatieoverdracht. Met name in de scholen speelt het geschreven of gesproken woord een belangrijke rol als het erom gaat leerlingen vertrouwd te maken met hun cultuurgood: wetenschappelijk, artistiek en maatschappelijk.

Toch is de vraag hoe mensen van tekst leren pas betrekkelijk laat een psychologische vraagstelling geworden. Bartlett (1932) wordt tegenwoordig

meestal aangewezen als degene die als eerste daartoe een poging deed. Het is echter vooral de cognitieve 'revolutie' in de psychologie geweest die een serieuze studie van tekstverwerkingsprocessen op enige schaal entameerde. In het voetspoor van deze revolutie houden onderwijspsychologen en anderen zich de laatste 20 jaar intensief bezig met onderzoek naar hoe leerlingen ertoe gebracht kunnen worden bij het bestuderen van een tekst cognitieve activiteiten te ontplooiën die hen helpen informatie uit die tekst te begrijpen en te onthouden. Het gaat daarbij om onderzoek naar wat genoemd worden: 'leerstrategieën' (O'Neil, 1978) of 'mathemagene activiteiten' (Rothkopf, 1970).

Om een indruk te geven van de aard van de onderzochte leerstrategieën zijn in tabel 1 een aantal voorbeelden opgenomen.

Tabel 1: Leerstrategieën, bedoeld om de bestudering van tekst te vergemakkelijken

-
- beantwoorden van vragen, opgenomen in de tekst (Rickards en Denner, 1979)
 - bestuderen van leerdoelen voorafgaand aan een tekst (Jenkins en Deno, 1971)
 - bestuderen van een advance organizer (Romberg en Wilson, 1973)
 - bestuderen van een metafoor voorafgaande aan tekst (Simons, 1981)
 - maken van een toets na lezing van een tekst (Nungester en Duchastel, 1982)
 - maken van aantekeningen tijdens het lezen (Bretzing en Kulhavy, 1981)
 - parafraseren van de tekst (Glover, Plake, Roberts, Zimmer en Palmere, 1981)
 - bedenken van concretiserende elaboraties (Wouters en Kop, 1981)
 - mobilisatie van tekstrelevante voorkennis (Peeck, Van den Bosch en Kreupeling, 1982)
 - bestuderen van een tekst met macroproposities (Van Oostendorp en Hamaker, 1978)
 - bestuderen van een tekst waarvan de volgorde varieert (Lodewijks, 1981)
 - bestuderen van een tekst waarvan de structuur varieert (Britton, Glynn, Meyer en Penland, 1982)
 - bestuderen van een tekst waaraan typografische signalen zijn toegevoegd (Glynn en Di Vesta, 1979)
-

Lang niet alle pogingen de leerstrategie van leerlingen te beïnvloeden zijn overigens even succesvol (Jenkins en Deno, 1971; Romberg en Wilson, 1973; Van Oostendorp en Hamaker, 1978). Soms worden zelfs negatieve effecten gerapporteerd (Rickards en Denner, 1979).

Een belangrijke oorzaak voor het feit dat studies naar de invloed van leerstrategieën op het verwerken van tekst lang niet altijd de verwachte effecten opleveren, is waarschijnlijk gelegen in het feit dat onderzoekers op dit terrein vaak meer interesse lijken te hebben voor de vraag of een bepaalde benadering succesvol is, dan wáárom hij (wel of niet) succesvol is. Men vraagt zich te weinig af welke cognitieve processen door de gekozen leerstrategie beïnvloed worden en welke resultaten daarvan te verwachten zijn (Bovy, 1981; Winne, 1982).

In de literatuur worden vier cognitieve processen gesuggereerd die mogelijk het coderen van informatie uit tekst ('encoding'), de opslag ('storage') en het terughalen van die informatie uit het lange-termijn-geheugen ('retrieval') reguleren. Die processen worden hier kort besproken omdat ze een rol spelen bij de evaluatie van een benadering van het onderwijsleerproces die het object van dit artikel is.

Het eerste proces kan omschreven worden als *activatie van voorkennis*. Adequate voorkennis lijkt een belangrijke rol te spelen bij het begrijpen en onthouden van nieuwe informatie. Spilich, Vesonder, Chiesi en Voss (1979) gingen bijvoorbeeld na in hoeverre het bezit van domeinspecifieke kennis invloed uitoefent op de verwerking van tekst. Ze lieten proefpersonen met veel, en proefpersonen met weinig kennis van honkbal een tekst over dat onderwerp bestuderen. Proefpersonen met veel voorkennis bleken veel beter in staat te zijn belangrijke informatie uit de tekst te halen en te onthouden.

Het hebben van voorkennis is op zichzelf echter niet voldoende om het leren te bevorderen. Die voorkennis moet ook *geactiveerd* worden door de leersituatie en ingezet bij de verwerking van de nieuwe informatie. Als mensen om de een of andere reden geen relatie leggen tussen het te bestuderen materiaal en hun kennis daarover, dan wordt de verwerving van dat materiaal bemoeilijkt. Bransford en Johnson (1972) lieten proefpersonen enigszins onsamenhangende teksten bestuderen. Sommige proefpersonen kregen voorafgaande aan de tekst een titel voorgelegd, anderen niet. Deze laatste groep leverde slechtere prestaties op maten voor herkenning en herinnering ('recall'). Blijkbaar activeert de voorafgegeven titel voorkennis, die het onthouden van de tekst vergemakkelijkt omdat hij een context biedt waarbinnen die informatie begrepen kan worden.

Men veronderstelt dat voorkennis zijn rol in het tekstverwerkingsproces speelt door de produktie van *inferenties*, die betekenis verlenen aan binnenkomende informatie en fungeren als verwachtingen omtrent nog te verwerken informatie (Schank en Abelson, 1977).

Als die inferenties tot stand komen door bewuste cognitieve activiteit van de persoon spreekt men ook wel van *elaboraties*.

Anderson (1976) liet proefpersonen in een associatietask paren woorden leren, zoals: hond—fiets. Sommige proefpersonen gaf hij de opdracht een betekenisvolle relatie aan te brengen tussen deze woorden, zoals: de hond rijdt op een fiets; anderen niet. Proefpersonen die de woorden ingebed hadden in een betekenisvol geheel, leverden aanzienlijk betere recallprestaties.

Elaboratie is dus het verrijken van informatie door het bedenken van relaties tussen ideeën uit een tekst die in die tekst niet expliciet gegeven zijn. Ook wordt het zoeken naar relaties tussen voorkennis en het in de tekst gegevene als elaboratie beschouwd. Het idee is dat deze extra-verwerkingsactiviteiten resulteren in nieuwe proposities die met de tekst geïntegreerd zijn. Naarmate er meer elaboraties gevormd zijn, wordt de retrieval van informatie vergemakkelijkt, omdat die informatie langs meerdere wegen toe-

gankelijk is geworden. Elaboraties zijn, omdat ze op voorkennis gebaseerd zijn, in hoge mate idiosyncratisch (Anderson en Reder, 1979).

Het tweede hier te bespreken proces is *selectieve aandacht*. Als proefpersonen voorafgaande aan de bestudering van een tekst leerdoelen krijgen voorgelegd die relevant zijn ten aanzien van bepaalde passages uit die tekst, dan leveren zij op een test met betrekking tot die passages een betere prestatie dan proefpersonen die het zonder leerdoelen hebben moeten stellen. Die betere prestatie gaat echter vaak ten koste van de prestatie op leerdoel-irrelevante passages. Daarop leveren proefpersonen aan wier tekst leerdoelen voorafgingen een slechtere prestatie (Gagné en Rothkopf, 1975). Vergelijkbare resultaten worden gemeld met vragen die aan tekst voorafgaan (André, 1979).

Een verklaring voor dit verschijnsel kan waarschijnlijk gezocht worden in de verdeling van leestijd over de verschillende passages in een tekst. Reynolds, Standiford en Anderson (1979) en Rothkopf en Billington (1979) vonden dat proefpersonen meer leestijd besteedden aan passages die relevant waren ten aanzien van een bepaald leerdoel, dan aan passages die dat niet waren. Het lijkt erop dat proefpersonen hun aandacht over een tekst verdelen, afhankelijk van de doelen die zij met het lezen van de tekst voor ogen hebben (Frijda, 1979).

Het *afleiden van macroproposities* uit tekst is een derde proces dat een rol lijkt te spelen bij tekstverwerking. Uit analyses van recallprotocollen blijkt dat mensen de belangrijke ideeën uit een tekst reconstrueren terwijl de details meestal achterwege gelaten worden. Mensen zijn blijkbaar in staat van de concrete tekst te abstraheren naar zijn kern. Zij kunnen de belangrijkste proposities uit een tekst halen onder weglating van de onbelangrijke proposities (Johnson, 1970; Kintsch, Kozminsky, Streby, McKoon en Keenan, 1975). Volgens Kintsch en Van Dijk (1978) komt dat omdat zogenaamde macro-operatoren (een verzameling abstraherende cognitieve processen) de proposities waaruit een tekst is opgebouwd transformeren in een set macroproposities, die de centrale boodschap van de tekst vormen. De toepassing van deze macro-operatoren staat onder controle van de doelstellingen waarmee de lezer de tekst bestudeert.

Meyer (1975) legt daarentegen ter verklaring van het feit dat belangrijke ideeën beter onthouden worden de nadruk op de structuur van de tekst zélf. Een tekst bestaat uit een verzameling proposities die onderling logisch of anderszins gerelateerd zijn. Voor zover deze relatie een dominantie-relatie is (dat wil zeggen dat sommige proposities onder andere geordend kunnen worden) vormen de proposities een hiërarchische structuur. De belangrijkste proposities nemen posities in boven in deze structuur. Meyer en haar collega's (Meyer en McConkie, 1973; Britton, Meyer, Hodge en Glynn, 1980) vinden dat deze centrale proposities twee of drie keer frequenter worden gereproduceerd dan proposities die een lagere positie in de tekststructuur innemen. Als een propositie gereproduceerd wordt, dan wordt de propositie direct boven deze propositie in de tekststructuur in 70% van de gevallen ook gereproduceerd, wat erop duidt dat mensen ken-

nis van de tekst organiseren rond centrale proposities die als ankers dienen bij de retrieval van die kennis (Meyer en McConkie, 1973).

Het vierde cognitieve proces ten slotte, dat een rol speelt bij de verwerking van tekst, en wel met name bij de retrieval, is het *toevoegen van retrieval-signalen* ('retrieval cues') aan de informatie.

Retrievalsignalen kunnen worden beschouwd als geheugensteuntjes die tegelijk met de te onthouden informatie worden opgeslagen (Tulving en Thomson, 1973). Als een retrievalssignaal geactiveerd wordt, activeert dit op zijn beurt de met dit signaal geassocieerde informatie. De context waarin een bepaalde tekst bestudeerd wordt, levert vaak de stimuli die de retrieval van die informatie vergemakkelijkt (Wittrock, Marks en Doctorow, 1975). Retrievalsignalen worden verondersteld van cruciaal belang te zijn, wanneer informatie uit het lange-termijn-geheugen moet worden teruggehaald voor gebruik in transfersituaties.

Activatie en herstructurering van voorkennis door middel van probleem-analyse

In deze bijdrage wordt verslag gedaan van onderzoek naar cognitieve effecten van een benadering van het onderwijsleerproces die 'probleemgestuurd onderwijs' genoemd wordt (Schmidt, 1982a; Schmidt en Bouhuijs, 1980). In het Engelse taalgebied heeft die benadering enige bekendheid onder de naam 'problem-based learning' (Barrows en Tamblyn, 1980; Neame, 1981). De methode heeft de volgende vorm. Aan een kleine groep leerlingen wordt de beschrijving van een aantal concrete verschijnselen uit de werkelijkheid voorgelegd. Hun taak is voor die verschijnselen verklaringen te bedenken in termen van een aan die verschijnselen ten grondslag liggend proces, principe of mechanisme. De beschrijving van de verschijnselen is het probleem; de verklaringen die de groep bedenkt, zijn de 'oplossingen' van dat probleem.

Door middel van discussie probeert de groep voorlopige hypothesen te produceren omtrent het proces of de processen die aan het probleem ten grondslag zouden kunnen liggen. Vervolgens wordt naar nieuwe, probleem-relevante informatie gezocht of de deelnemers krijgen die informatie ter bestudering voorgelegd. Een tutor heeft bij dit proces een verhelderende functie. Voor een uitgebreidere beschrijving: zie Schmidt (1982b).

Drie elementen uit dit onderwijsarrangement behoeven enige toelichting: de aard van de gebruikte probleembeschrijvingen, de taak van de groep om 'verklaringen' te bedenken, en de rol van discussie bij het produceren van die verklaringen.

Een probleem werd gedefinieerd als een concrete beschrijving van een verzameling verschijnselen uit de werkelijkheid. De nadruk ligt daarbij op concreetheid, om leerlingen in staat te stellen zich zonder al te veel cognitieve inspanning voor te stellen wat verklaard moet worden. Concreetheid wordt algemeen gezien als een belangrijk kenmerk van materiaal dat geconstrueerd wordt om het leerproces te vergemakkelijken. Royer en Cable

(1976) gebruikten illustraties en concrete analogieën om het leren van tekst te vergemakkelijken. Simons (1981) bood proefpersonen metaforen van de te bestuderen leerstof aan, voorafgaande aan die bestudering. Mayer (1980; 1982) lijkt het falen van sommige advance organizer studies te verklaren uit het feit dat proefpersonen de organizers gewoonweg niet begrepen vanwege hun abstractieniveau.

Het tweede element dat toelichting behoeft, is de nadruk die gelegd wordt op het vinden van 'verklaringen' voor de beschreven verschijnselen als doel van de activiteiten van de groep. Daarmee wordt geduid op een oude tegenstelling in de leerpsychologie, die de Gestaltpsychologen omschreven als 'rote learning' versus 'meaningful insight learning' (Wertheimer, 1959). Doel van scholing is niet in de eerste plaats leerlingen feiten zonder enige betekenis en samenhang te leren, maar hen te helpen de wereld om hen heen te begrijpen. Bij dat laatste speelt kennis van processen, principes en mechanismen een belangrijke rol. Katona (1940) liet bijvoorbeeld zien dat proefpersonen die zich bij het leren concentreerden op de structuur onder een hoeveelheid informatie, beter dan anderen in staat waren die informatie in transfertaken te gebruiken.

Bromage en Mayer (1981) lieten proefpersonen een tekst bestuderen over het gebruik van camera's. Die tekst bevatte een hoeveelheid feitelijke gegevens over camera's, naast informatie over hoe een camera werkt en op welke wijze een foto tot stand komt. Proefpersonen die in de recall relatief meer verklarende proposities reproduceerden, bleken betere prestaties te leveren op transfertaken dan proefpersonen die relatief meer descriptieve proposities reproduceerden. Begrip van onderliggende processen speelt dus blijkbaar een belangrijke rol bij het oplossen van transfervraagstukken.

De verklaringen die voor het probleem bedacht worden, worden geproduceerd in een discussie. Die discussie bestaat uit *explicitering* van de voorkennis die deelnemers aan de probleemanalyse met betrekking tot het probleem beschikbaar hebben. Datgene wat geproduceerd wordt, kan uit twee bronnen afkomstig zijn. Deelnemers kunnen de beschikking hebben over feitelijke voorkennis ten aanzien van het aan het probleem ten grondslag liggende proces; of zij beschikken over algemene wereldkennis die hen in staat stelt door middel van *elaboratie* iets over het probleem te berde te brengen.

De probleemanalyse activeert dus voorkennis, relevant ten aanzien van het probleem. Men mag echter veronderstellen dat gedurende de discussie meer gebeurt dan alleen maar activatie van aanwezige cognitieve structuren. De poging verklaringen te bedenken voor het concrete probleem zal naar alle waarschijnlijkheid ook leiden tot *herstructurering* van die cognities. Collins (1977) laat bijvoorbeeld zien hoe een discussie over factoren die het weer in een omgeving bepalen, iemands conceptie daarvan herstructureert. Bargh en Schul (1980) en Allen en Feldman (1973) maken aannemelijk dat het uitleggen aan anderen een herstructureerende invloed heeft op de cognities van degene die uitlegt.

Nu is het niet gemakkelijk onderscheid te maken tussen welke produkten

van een discussie het resultaat zijn van activatie, en welke het resultaat van herstructurering. In het navolgende zal gemakshalve datgene wat de persoon zélf produceert geoperationaliseerd worden als geactiveerde voorkennis, en datgene wat hij in de discussie van anderen leert als geherstructureerde voorkennis. Een dergelijke operationalisatie maakt het mogelijk nader te onderzoeken in welke mate de discussie over het probleem tot herstructurering, dat wil zeggen tot leren leidt.¹

Samenvattend: er wordt verondersteld dat de analyse van een concreet probleem in een kleine groep bij de deelnemers voorkennis mobiliseert en elaboraties daarop stimuleert. Onder invloed van de discussie vindt herstructurering van die voorkennis plaats. De aldus ontstane kennisstructuren faciliteren vervolgens de cognitieve verwerking van probleemrelevante nieuwe informatie uit tekst.

In de hieronder te rapporteren experimenten zijn enkele van de in de inleiding geformuleerde veronderstellingen getest aan de hand van recall-gegevens.

In Experiment 1 werd onderzocht of probleemanalyse voorkennis activeert en zo ja, wat dan de aard van die voorkennis is. In Experiment 2 werd de invloed van probleemanalyse op de verwerking van een relevante tekst onderzocht. Bij de evaluatie van de resultaten van dit experiment werd gezocht naar evidentie voor de werkzaamheid van één of meer van de hiervoor beschreven cognitieve processen die de verwerking van tekst reguleren.

In een deel van de data uit deze experimenten, een deel dat al eerder gerapporteerd werd (Schmidt, 1982c), was enige ondersteuning voor de hier geformuleerde hypothesen te vinden. De aard van de data (het betrof herkennings- en transfergegevens) liet echter geen nadere analyse van de aard van de betrokken cognitieve processen toe.²

Experiment 1

In Experiment 1 werd een poging gedaan een antwoord te vinden op drie vragen:

1. Activeert een probleem en de discussie hierover voorkennis die relevant is ten aanzien van dat probleem?
2. Zo ja, wat is dan de bron van die kennis? Reproduceren proefpersonen voornamelijk kennis die bij henzelf geactiveerd is of reproduceren ze ook kennis die ze van anderen hebben geleerd tijdens de discussie? en
3. Wat is de aard van die kennis?

¹ Natuurlijk kan ook elaboratie op de *eigen* voorkennis al tot herstructurering van die kennis leiden. De hier gekozen operationalisatie is dan ook niet meer dan een eerste ruwe benadering.

² Het betreffende artikel is in dit proefschrift als hoofdstuk 3 opgenomen.

Om deze vragen te kunnen beantwoorden werd aan kleine groepjes studenten een beschrijving voorgelegd van het gedrag van een bloedcel in water en in een zoutoplossing, met het verzoek dat gedrag te verklaren. Na afloop hiervan werd hen gevraagd alles op te schrijven wat zij zich van het onderwerp osmose herinnerden. Deze recallprotocols werden vergeleken met die van een controlegroep.

Methode

Proefpersonen

Aan Experiment 1 werd door 39 proefpersonen deelgenomen: 31 vrouwen en 8 mannen. Hun gemiddelde leeftijd was 19,4 jaar met een standaarddeviatie van 1,1. De proefpersonen waren studenten aan een instelling voor hoger beroepsonderwijs; een opleiding voor kinderverzorging en -opvoeding. Allen waren in het bezit van een HAVO-diploma. Een verdere voorwaarde voor deelname was dat proefpersonen geëxamineerd waren in het vak biologie. Homogeniteit van de proefgroep met betrekking tot voorkennis op het terrein van de biologie was in relatie tot de vraagstellingen gewenst.

De proefpersonen ontvingen een vergoeding voor hun medewerking.

Materiaal

Het materiaal bestond uit een probleembeschrijving en een aantal vragenlijsten.

Het probleem bestond uit de volgende tekst: 'Een rode bloedcel (een rood bloedlichaampje) wordt onder een microscoop in zuiver water gebracht. De bloedcel zwelt in korte tijd op en springt uiteindelijk kapot. Een andere bloedcel wordt in een oplossing van zout in water gebracht. Hij schrompelt ineen. Hoe zijn deze verschijnselen te verklaren?'

Proefpersonen kregen dus tot taak het beschreven gedrag van de bloedcel te verklaren in termen van een aan dit gedrag ten grondslag liggend proces, principe of mechanisme. In dit geval is dat het osmotisch proces.

Osmose is gekozen, omdat het als dynamisch proces aan een groot aantal in de natuur waar te nemen fenomenen ten grondslag ligt. Daarnaast wordt het in op de HAVO veelgebruikte biologie-handboeken op een min of meer gelijke wijze behandeld, zoals bij inspectie van vier van deze boeken bleek. (Met gelijk wordt hier bedoeld dat de teksten een vergelijkbare diepgang hadden en vergelijkbare voorbeelden presenteerden.)

Als afhankelijke variabele figureerde het *recallprotocol* bestaande uit datgene wat proefpersonen zich in samenhang met de onderwerpen osmose en diffusie herinnerden.

Voor een van de analyses werd een *verbatim transcriptie* gemaakt van geluidsopnamen van de probleemanalyse.

Ten slotte werden enkele *biografische gegevens* verzameld, zoals leeftijd, geslacht en eindcijfer voor biologie op de HAVO.

Procedure

De proefpersonen werden aselekt toegewezen aan één van de twee condities van het experiment. De experimentele groep (verder de probleemanalysegroep genoemd) bestond uit 20 proefpersonen; de controlegroep bevatte 19 proefpersonen.

Na de randomisatieprocedure werd gecontroleerd in hoeverre beide groepen vergelijkbaar waren voor zover het biologische voorkennis betrof. Het gemiddelde eindcijfer voor biologie in de probleemanalysegroep was gelijk aan 6,30, met een standaardafwijking van 0,73. De controlegroep had een gemiddelde van 6,47, met een standaardafwijking van 0,69. Deze kleine verschillen gaven geen aanleiding om een matching-achteraf uit te voeren, of een covariantie-analyse van de recallgegevens te overwegen.

De probleemanalysegroep werd vervolgens aselekt opgedeeld in drie subgroepen, bestaande uit respectievelijk 7, 7 en 6 proefpersonen. In deze groepen vond de analyse van het bloedcelprobleem plaats.

Elke groep kreeg aselekt een vrouwelijke tutor toegewezen. Deze tutores hadden ervaring met de hiervoor beschreven methode van probleemanalyse en waren vertrouwd met het onderwerp osmose. Tutores en proefpersonen hadden elkaar voor de experimentele sessie nooit eerder ontmoet.

De probleemanalyse vond plaats in klaslokalen, waarbij stoelen en tafels in een vierkant gerangschikt waren.

De tutor ging op de volgende wijze te werk. Ze besteedde eerst enige tijd aan een kennismaking en aan het op hun gemak stellen van de proefpersonen. Vervolgens introduceerde ze de methode aan de hand van een voorbeeld. De proefpersonen hadden namelijk geen ervaring met de van hen gevraagde aanpak van het probleem. Ze reikte daartoe de volgende probleembeschrijving uit: 'Waterpest wordt in een bekerglas met water geplaatst dat rijk is aan kooldioxide. Als dit glas aan het licht wordt blootgesteld komen er gasbelletjes vrij. Als deze worden opgevangen in een reageerbuisje kan men met een gloeiende houtspaander aantonen dat het ontwijkende gas zuurstof is. Wordt het licht afgeschermd dan houdt de zuurstofproductie op.'

Aan de hand van deze probleembeschrijving lichtte ze de verschillende stappen van de analyse toe. Op een schoolbord had ze daartoe de volgende tekst geschreven: Probleemanalyse: 1. Nadenken. 2. Definitie van het probleem. 3. Verklaringen bedenken. 4. Uitdieping. De proefpersonen werden bij de uitleg van de methode actief betrokken.

Voordat het bloedcelprobleem werd aangeboden vertelde de tutor dat de probleemanalyse niet langer dan 15 minuten zou duren. Het probleem werd uitgedeeld en de groep kreeg 1 minuut voor lezen en nadenken. Daarna begon de discussie. De tutor trad daarbij als gespreksleider op; dat wil zeggen dat zij datgene wat door de studenten naar voren gebracht werd, ordende. Zij had expliciet de opdracht geen informatie over het probleem te verschaffen en erop toe te zien dat ook haar samenvattingen niet impliciet informatie bevatten, waaraan proefpersonen inzichten met betrekking tot het probleem zouden kunnen ontleenen.

Er werd een geluidsopname van de discussie gemaakt om achteraf te kunnen controleren of de betrokkenen zich aan de afgesproken aanpak hadden gehouden. Dit bleek het geval. Alle groepen hadden minder dan 10 minuten nodig voor de analyse.

Vervolgens werd aan probleemanalyse- en controlegroep gevraagd op te schrijven wat zij zich van het onderwerp osmose en diffusie herinnerden. Benadrukt werd dat datgene wat werd opgeschreven niet per se steeds een samenhangend geheel hoefde te zijn, maar dat het belangrijk was dat alles werd opgeschreven, desnoods in losse zinnen.

Identificatie van proposities in de recall

De door de proefpersonen geproduceerde teksten werden in proposities opgedeeld. Propositionen zijn (meestal) onderwerp-gezegde-combinaties, die één bepaald idee uitdrukken. Men kan ze in teksten identificeren door voegwoorden (als 'en', 'of' en 'want'), bijwoorden (als 'wanneer' en 'waarvoor'), betrekkelijke voornaamwoorden, punten en puntkomma's te markeren. De tekst delen tussen deze linguïstische markeringspunten kunnen als proposities beschouwd worden. Soms scheiden ook komma's of haakjes proposities van elkaar.

Een voorbeeld geeft de analyse van de volgende tekst. Verticale strepen scheiden de proposities van elkaar.

Osmose is het verschijnsel | waarbij een sterkere oplossing water aantrekt uit een zwakkere oplossing, | wanneer deze oplossingen gescheiden zijn door een half-doorlaatbare wand. | Het Griekse woord 'osmos' betekent aanzuiging. | De aanzuiging van watermoleculen kan tot stand komen door middel van diffusie (diffundere betekent uitbreiden).

Met behulp van deze regels werden de door proefpersonen geproduceerde teksten door twee onafhankelijke beoordelaars opgedeeld in samenstellende proposities. De overeenstemming tussen de beoordelaars was 92%. Verschil van mening werd door middel van discussie opgelost. De hier beschreven methode om teksten te analyseren leidt in grote lijnen tot dezelfde resultaten als de methode beschreven door Meyer (1975).

Van elk van de proposities werd vervolgens vastgesteld of hij correct dan wel incorrect was.

Ten slotte werden de proposities onderverdeeld in verklarende en descriptieve proposities (Bromage en Mayer, 1981). Dit onderscheid werd van belang geacht omdat men zou kunnen verwachten dat in de protocollen van de probleemanalysegroep, vergeleken met de controlegroep, relatief meer verklarende proposities voorkomen. Verklarende proposities werden gedefinieerd als proposities die een proces beschrijven ('een hoeveelheid watermoleculen verplaatst zich van de vloeistof met de hoge concentratie naar de vloeistof met de lage concentratie'), dan wel proposities die de causale factoren van het proces beschrijven, of de condities waaronder dat proces

plaatsvindt ('Wanneer 2 oplossingen met een verschillende concentratie aan opgeloste stoffen gescheiden zijn door een semi-permeabel vlies, dan ...').

Alle overige proposities werden als descriptief gelabeld. (Voorbeeld: 'Osmose komt zowel bij mensen als bij dieren en planten voor.') De interbeoordelaars-overeenstemming bij deze analyse was gelijk aan 79%.

Identificatie van proposities in de verbatim transcripties van de probleem-analyse

De verbatim transcripties van de probleemanalyse werden op dezelfde wijze behandeld als de recallprotocols. De teksten werden opgedeeld in proposities. Voor elk van de proposities werd vastgelegd door wie hij geproduceerd was.

Resultaten en discussie

In tabel 2 zijn voor beide groepen gemiddelde aantallen geproduceerde proposities weergegeven. Het gaat daarbij om het *totaal aantal* in de recall teruggevonden proposities; correcte en incorrecte proposities zijn zonder onderscheid bij elkaar opgeteld. Als men ervan uitgaat dat ook incorrecte kennis kennis is, dan geeft tabel 2 een goed overzicht van de verschillen in kennis tussen beide groepen na de probleemanalyse.

Tabel 3 bevat alleen de correcte proposities.

Tabel 2: Gemiddelde aantallen proposities, correcte en incorrecte (met standaarddeviaties), geproduceerd in Experiment 1

	$\bar{X}$	SD	N
Probleemanalysegroep	36,35	11,13	20
Controlegroep	16,95	6,07	19
Totaal	26,90	9,03	39

Tabel 3: Gemiddelde aantallen correcte proposities (met standaarddeviaties), geproduceerd in Experiment 1

	$\bar{X}$	SD	N
Probleemanalysegroep	27,20	9,13	20
Controlegroep	11,84	5,37	19
Totaal	19,72	7,54	39

Op beide variabelen werd een variantieanalyse uitgevoerd.

De verschillen tussen gemiddelden, zowel voor het totaal aantal proposities als voor het aantal correcte proposities, zijn in hoge mate statistisch significant. Totaal aantal proposities: $F(1,37) = 44,98$; $p = 0,00$. Aantal correcte proposities: $F(1,37) = 40,43$; $p = 0,00$.

Tabel 4 bevat het gemiddeld aantal incorrecte proposities dat beide groepen produceerden.

Tabel 4: Gemiddelde aantallen incorrecte proposities (met standaarddeviaties), geproduceerd in Experiment 1

	$\bar{X}$	SD	N
Probleemanalysegroep	9,15	4,56	20
Controlegroep	5,11	2,13	19
Totaal	7,18	3,59	39

Ook deze gemiddelden verschillen van elkaar. De probleemanalysegroep produceert significant meer incorrecte proposities: $F(1,37) = 12,37$; $p < 0,002$. De aantallen incorrecte proposities zijn echter proportioneel aan de totale aantallen geproduceerde proposities: $z = 0,38$; $p < 0,36$. Men zou dus kunnen zeggen dat wie meer weet, blijkbaar ook meer fout weet.

De conclusie die uit deze analyses getrokken kan worden, is dat probleem-analyse bij diegenen die daaraan deelnemen verrassend veel voorkennis activeert. De probleemanalysegroep produceert ruim twee maal zoveel proposities als de controlegroep. De puntbiseriële correlatie verschaft een beeld van de grootte van de invloed die de experimentele manipulatie op de afhankelijke variabele uitoefent. $r_{pbis} = 0,72$. Dat betekent dat het wel of niet deelnemen aan de probleemanalyse 52% van de variantie in prestatie verklaart. Probleemanalyse lijkt een krachtige methode te zijn om kennis die in een eerder stadium verworven is, uit het lange-termijn-geheugen terug te halen.

De vraag doet zich nu voor wat de bron is van de in de recall geproduceerde proposities. Meer concreet: zijn de in de recall gevonden proposities vooral een functie van wat proefpersonen gedurende de discussie over het probleem zelf geproduceerd hebben, of neemt ook van anderen verworven kennis een aandeel daarin? In de inleiding is dit de operationalisatie van de dichotomie 'geactiveerde' versus 'geherstructureerde' voorkennis genoemd. Ten einde deze vraag te kunnen beantwoorden is een van Bretzing en Kulhavy (1981) afgeleide aanpak gekozen. Bretzing en Kulhavy onderzochten welke relatie er bestond tussen de aantekeningen die proefpersonen gemaakt hadden tijdens het doornemen van een passage en datgene wat zij zich van die passage herinnerden. Zij berekenden daartoe conditionele proporties — zelf noemen ze dat conditionele waarschijnlijkheden — $p(R|N)$. Deze parameter duidt de relatieve frequentie aan waarmee proposities in

de recall teruggevonden worden, gegeven het feit dat die proposities in de aantekeningen aanwezig waren ($R = \text{Recall}$, $N = \text{Notes}$).

In het hier beschreven experiment werd de volgende procedure toegepast. Verbatim transcripties van de discussies en recallprotocols werden met elkaar vergeleken. Voor iedere proefpersoon in de probleemanalysegroep werden twee conditionele proporties berekend. De eerste, $p(R|Z)$, duidt de relatieve frequentie van herinnerde proposities aan, gegeven het feit dat die proposities door de persoon zelf ($Z = \text{Zelf}$) in de probleemanalyse geproduceerd zijn. De tweede conditionele proportie, $p(R|A)$, duidt de relatieve frequentie van herinnerde proposities aan die door anderen ($A = \text{Anderen}$) geproduceerd zijn. Deze conditionele proporties zijn vervolgens gemiddeld over alle proefpersonen. In tabel 5 zijn de resultaten samengevat.

Tabel 5: Gemiddelde conditionele proporties, $p(R|Z)$ en $p(R|A)$, (met standaarddeviaties) voor Experiment 1

	Gemiddelde	SD
$p(R Z) *$	0,42	0,10
$p(R A) *$	0,19	0,06

* Voor toelichting, zie de tekst.

Deze resultaten geven aan dat ongeveer 40% van de in de recall voorkomende proposities beschouwd kan worden als door de procedure *geactiveerde* voorkennis, terwijl bijna 20% bestaat uit voorkennis die onder invloed van door anderen geproduceerde kennis *herstructurering* heeft ondergaan: de persoon integreert blijkbaar een deel van wat anderen in de discussie te berde brengen in zijn eigen cognitieve structuren en reproduceert die kennis in de recall. Merk op dat bijna 40% van de herinnerde proposities niet in de discussie naar voren gebracht is. Tenzij de recallprocedure zelf nog weer nieuwe voorkennis activeert, moet worden aangenomen dat proefpersonen niet alles wat tijdens de discussie geactiveerd wordt, expliciteren; mogelijk omdat ze er stilzwijgend van uitgaan dat andere deelnemers al over die informatie beschikken (Schmidt, 1979).

Ten slotte is in Experiment 1 nog onderzocht wat de aard van de in de recall geproduceerde proposities is. Verondersteld werd dat de aard van de taak die aan de groep verstrekt werd (namelijk de verschijnselen in termen van onderliggende processen te verklaren) invloed zou uitoefenen op het soort proposities dat in de recall gevonden zou worden. De probleemanalysegroep werd verondersteld meer verklarende proposities te produceren dan de controlegroep. Dat wil zeggen dat de probleemanalysegroep over osmose en diffusie meer in procesmatige termen zou schrijven en meer causale en anderzins conditionele factoren voor die processen zou benoemen.

Daartoe zijn de proposities onderverdeeld in verklarende en descriptieve proposities (Bromage en Mayer, 1981), en zijn voor beide categorieën proporties berekend. In tabel 6 zijn de proporties weergegeven.

Tabel 6: Proporties verklarende en descriptieve proposities in Experiment 1

	Verklaringen	Descripties	n*
Probleemanalysegroep	0,53	0,47	544
Controlegroep	0,37	0,63	225

* n is het totaal aantal correcte proposities.

De proporties verklarende proposities verschillen significant van elkaar: $z = 4,03$; $p < 0,0001$. Het lijkt er dus op dat de probleemanalyse proefpersonen richt op de procesmatige aspecten van osmose en diffusie en de recall van procesrelevante proposities stimuleert. Bromage en Mayer (1981) vonden een correlatie van $r = 0,59$ tussen de aard van de geproduceerde proposities (verklarend versus descriptief) en de prestaties van proefpersonen op een transfertoets. Goede probleemoplossers (zij die hoog scoorden op de transfertoets) produceerden meer verklarende proposities. Bromage en Mayer concluderen daaruit dat goede probleemoplossers zich blijkbaar een dieper inzicht verworven hebben in de leerstof — tot uitdrukking komend in meer verklarende proposities — en *daardoor* beter in staat zijn problemen op te lossen.

Samengevat: de analyse van een probleem lijkt een krachtige procedure te zijn om voorkennis te activeren en te herstructureren. De probleemanalysegroep produceerde ruim twee maal zoveel correcte proposities als de controlegroep. Veertig procent daarvan kan beschouwd worden als geactiveerde proposities, twintig procent bestaat uit proposities die proefpersonen van anderen overnemen en in hun eigen cognitieve structuur integreren. De helft van de proposities in de protocols van de probleemanalysegroep bestond uit verklaringen, bij de controlegroep maakte deze categorie iets meer dan een derde deel van de proposities uit. Als men met de conclusie van Bromage en Mayer (1981) meegaat dat meer verklarende proposities duiden op een dieper inzicht in de materie, kan vastgesteld worden dat probleemanalyse proefpersonen stimuleert tot een beter begrijpen van de processen die aan het bloedcelprobleem ten grondslag liggen.

Experiment 2

In Experiment 1 is aangetoond dat pogingen om in een probleem beschreven fenomenen te verklaren als een functie van onderliggende processen, een sterke activator zijn van de cognities die personen over een bepaald onderwerp beschikbaar hebben.

De vraag doet zich nu voor wat de invloed van een dergelijke activatie is op de cognitieve verwerking van een nieuwe, doch verwante tekst. Is het effect zo robuust dat het de verwerking van probleemrelevante informatie faciliteert? En zo ja, op welke wijze? Deze vragen zijn verre van triviaal, al was het alleen maar omdat onduidelijk is of en hoe de geactiveerde voorkennis met de nieuwe informatie interacteert. Mayer (1975) zegt bijvoorbeeld dat activatie van voorkennis op zichzelf onvoldoende is om extra leerresultaten te verkrijgen. Het geactiveerde kennisbestand moet ook daadwerkelijk worden ingezet bij het begrijpen van de nieuwe informatie. Het is dus zeer wel mogelijk dat in de recall van teksten geen effecten van de probleemanalyse waarneembaar zijn omdat proefpersonen domweg geen relatie leggen tussen datgene wat geactualiseerd is en de inhoud van de tekst. Of — andere mogelijkheid — omdat de tekst op zichzelf al zoveel voorkennis activeert bij lezers van die tekst, dat het door de probleemanalyse geïnduceerde effect verloren gaat.

In Experiment 2 bestudeerde daarom een experimentele groep een tekst over osmose en diffusie, nadat eerst het bloedcelprobleem geanalyseerd was. De recallprestatie van deze probleemanalysegroep werd vergeleken met die van een controlegroep die enkel de tekst bestudeerd had.

In de inleiding tot deze bijdrage is betoogd dat het vaststellen van leerwinsten bij het bestuderen van tekst als gevolg van ondersteunende leerstrategieën weinig informatief is, als niet ook een poging gedaan is te onderzoeken welke intermediaire cognitieve processen voor die leereffecten verantwoordelijk gesteld kunnen worden. In diezelfde inleiding zijn ook een aantal processen beschreven die in principe als verklaringsgrond voor mogelijke effecten in aanmerking komen. In het navolgende zullen die processen opnieuw kort besproken worden, maar nu vanuit het perspectief van mogelijke invloeden die de activatie van bestaande cognitieve structuren op die processen uitoefent.

De literatuur over de rol die voorkennis speelt bij de verwerking van tekst (e.g. Anderson, Spiro en Anderson, 1978; Ausubel, 1963; Bower, Black en Turner, 1979; Peeck, 1979; Pichert en Anderson, 1977; Spilich, Vesonder, Chiesi en Voss, 1979) levert weinig restrictieve voorspellingen op ten aanzien van de effecten die van activatie van bestaande kennisstructuren te verwachten zijn. De teneur is dat van een dergelijke ingreep een *algemeen* faciliterend effect zal uitgaan (Gagné, 1978). Buiten een studie van Peeck, Van den Bosch en Kreupeling (1982), waarvan verderop nog sprake zal zijn, is echter feitelijk nog nooit expliciet geprobeerd bestaande kennis-schema's voorafgaande aan een tekstbestuderingstaak te mobiliseren bij proefpersonen, en de invloed daarvan op de verwerking van die tekst na te gaan. Afgewacht moet dus worden of de algemeen-positieve effecten van het aanbieden van een titel voordat een tekst bestudeerd wordt (Dooling en Lachman, 1971), het verschaffen van relevante voorkennis (Morris, Stein en Bransford, 1979) of het bezitten van expertise in een bepaald domein (Spilich e.a., 1979), gegeneraliseerd kunnen worden naar de hier onderzochte situatie.

Peeck e.a. (1982) lieten leerlingen een tekst bestuderen over een fictie dier: de Amerikaanse heidevos. Een deel van de proefpersonen kreeg voor af een voorkennisactivatietaak: zij moesten opschrijven wat zij wisten over het uiterlijk en de eet- en leefgewoonten van vossen. Nóch in de recall nóch op een meerkeuzetoets over de tekst (die een week later werd afgenomen) waren effecten van voorkennisactivatie terug te vinden.

Indien er echter voorlopig gemakshalve van uitgegaan wordt dat er een faciliterend effect optreedt, dan kunnen die effecten langs verschillende wegen tot stand komen.

De *eerste mogelijkheid* is dat proefpersonen wier kennisbestand door de activatieprocedure geactualiseerd is, sneller in staat zijn de tekst te begrijpen. Dat wil zeggen dat zij meer informatie-eenheden per tijdseenheid kunnen verwerken. Spilich e.a. (1979) veronderstellen dat dat zou komen omdat de voorkennis makkelijker toegankelijk is. De belasting van het werkgeheugen neemt daardoor af (de capaciteit daarvan neemt toe) er meer informatie per tijdseenheid kan verwerkt worden. Een dergelijk proces zou moeten resulteren in méér proposities in de recall, omdat de informatie een intensievere verwerking heeft ondergaan.

De *tweede mogelijkheid* is dat geactiveerde voorkennis het produceren van betekenisvolle elaboraties stimuleert (Mayer, 1980; Reder, 1980). Dit verkrijgen van de informatie uit tekst met op voorkennis gebaseerde elaboraties leidt volgens Anderson en Reder (1979) tot wat zij noemen meervoudig overvloedige retrieval-paden ('multiple redundant retrieval paths') in de resulterende cognitieve structuren, die het terughalen van de gecodeerde informatie bij recall vergemakkelijken. Ook deze theorie voorspelt dus het optreden van méér proposities in de recall, net als de opvatting die zegt dat de informatie sneller verwerkt wordt. Het elaboratieproces laat echter bijzondere sporen na in de recall. Doordat proefpersonen, als gevolg van de wijze waarop ze de informatie verwerkt hebben, geen goed onderscheid meer kunnen maken tussen wat feitelijk in de tekst stond en hun eigen elaboraties daarop, neemt het aantal *inferenties* in de recall toe (Bartlett, 1932; Frederiksen, 1975). Inferenties zijn tekstrelevante proposities die beschouwd kunnen worden als door de lezer geproduceerde, zinvolle aanvullingen op de informatie uit tekst.

De *derde mogelijkheid* is dat proefpersonen informatie uit tekst op een andere wijze in het lange-termijn-geheugen organiseren, als gevolg van de activatieprocedure. Matthews (1982) liet proefpersonen voorafgaande aan een criteriumtekst een relevante prepassage bestuderen. Vervolgens ging hij in de recall na hoe deze proefpersonen de informatie uit de criteriumtekst organiseerden, vergeleken met anderen die geen prepassage gelezen hadden. Hoewel in de protocols van beide groepen de structuur van de oorspronkelijke criteriumtekst nog wel terug te vinden was, verschilden de groepen *van elkaar* in de wijze waarop zij die tekstinformatie georganiseerd hadden. Het is dus niet uitgesloten dat de probleemanalysegroep de informatie uit tekst rondom andere, bijvoorbeeld aan het probleem ontleende, concepten organiseert. De wijze van organisatie van informatie be-

invloedt op zijn beurt in hoge mate de toegankelijkheid ('accessibility') van die informatie in het lange-termijn-geheugen (Gardner en Schumacher, 1977). Verschillen in organisatie worden verondersteld zich in de recall te weerspiegelen.

Er is nog een *vierde verwerkingsstrategie* die ten aanzien van de rol van voorkennis in het tekstverwerkingsproces enige a priori plausibiliteit bezit. Het is mogelijk dat de probleemanalyse proefpersonen ertoe brengt hun aandacht selectief op bepaalde, bijvoorbeeld probleemrelevante, passages in de tekst te richten. Deze passages worden vervolgens intensiever verwerkt. Men zou zich dus, naar analogie van de functie van leerdoelen of toegevoegde vragen vooraf (Anderson en Biddle, 1975; Melton, 1978), kunnen voorstellen dat de probleemanalyse een sturende functie vervult ten aanzien van het cognitieve systeem. Deze selectieve-aandachts-hypothese voorspelt dat de probleemanalysegroep aan probleemrelevante informatie meer aandacht geeft (meestal uitgedrukt in studietijd), maar dat dit ten koste gaat van de bestudering van informatie die niet relevant is ten aanzien van het probleem. Gegeven een standaard hoeveelheid beschikbare studietijd wordt aan deze informatie namelijk minder aandacht gegeven (Reynolds, Standiford en Anderson, 1979).

Er zijn aanwijzingen die deze veronderstelling niet geheel waarschijnlijk maken. In het onderzoek van Peeck e.a. (1982) werd een tekst gebruikt die, naast voorkennis-relevante, ook voorkennis-neutrale en voorkennis-incongruente informatie bevatte. De selectieve-aandachts-hypothese voorspelt dat aan deze laatste informatie minder aandacht besteed wordt. Dit bleek echter in het geheel niet het geval. De groep wier kennisbestand gemobiliseerd was, leverde juist op dit deel van de informatie een significant betere prestatie, hoewel, zoals gezegd, de totale prestatie niet verschilde. Schmidt (1982c) vond een vergelijkbaar effect. Toch lijkt het zinvol de invloed van probleemanalyse op de verwerking van probleemrelevante en niet-probleemrelevante informatie uit tekst nader te onderzoeken.

Een *laatste punt* van discussie wordt in de literatuur gevormd door de vraag waar activatie van voorkennis in het cognitieve proces aangrijpt, bij de codering en opslag van de informatie of bij de retrieval. In het voorgaande is er stilzwijgend van uitgegaan dat activatie van bestaande kennis-schema's de wijze van *codering* van de nieuwe informatie beïnvloedt. Dat lijkt ook wel een correcte veronderstelling, gegeven de beschikbare informatie op dit punt. Dooling en Mullet (1973) lieten proefpersonen een enigszins onsamenhangende tekst bestuderen. Sommigen kregen een faciliterende titel vooraf; anderen na het lezen van de tekst. Alleen de titelvooraf vergemakkelijkte de recall van die tekst, wat erop duidt dat voorkennis-activatie de codering van informatie beïnvloedt en niet de retrieval. Mayer en Bromage (1980) vonden een zelfde resultaat bij de aanbieding van een organizer voorafgaande aan een tekst of daarna. Anderson enichert (1978) stelden echter vast dat activatie van relevante cognitieve structuren ná lezing van een tekst niettemin de recall van die tekst bevorderde, en concludeerden daaruit dat activatie van voorkennis óók bij

de retrieval van informatie een rol speelt.

Uit dit laatste resultaat zou men kunnen afleiden dat dus toch niet is uitgesloten dat de probleemanalyse (ook) als een 'retrieval aid', een geheugensteun, zou kunnen functioneren. Een dergelijke stand van zaken heeft repercussies voor het gemak waarmee na enige tijd probleemrelevante versus niet-probleemrelevante informatie in het lange-termijn-geheugen kan worden teruggevonden. Men moet aannemen dat, als de probleemanalyse een retrievalsignaal (Tulving en Thomson, 1973) voor de informatie vormt, na enige weken probleemrelevante informatie makkelijker toegankelijk is dan niet-probleemrelevante informatie.

Ter wille van de overzichtelijkheid zijn de verschillende mogelijkheden in tabel 7 samengevat.

Tabel 7: Te verwachten effecten van activatie van voorkennis op recall, gegeven verschillende cognitieve processen

Op welke wijze beïnvloedt activatie van voorkennis het tekstverwerkingsproces.	Te verwachten effect op de recall.
1. Per tijdseenheid wordt meer informatie verwerkt.	Probleemanalysegroep produceert meer proposities dan controlegroep.
2. Produktie van elaboraties leidt tot 'multiple redundant retrieval paths'.	Probleemanalysegroep produceert meer proposities, waaronder meer inferenties.
3. Organisatie van kennis in het lange-termijn-geheugen vindt op andere wijze plaats.	Verschillen in organisatie van de recall.
4. Probleemanalyse richt de aandacht van de lezer op probleemrelevante delen van de tekst.	Probleemanalysegroep produceert meer probleemrelevante proposities (en minder niet-probleemrelevante).
5. Probleemanalyse fungeert als retrievalsignaal.	Probleemrelevante informatie vertoont minder verval na twee weken dan niet-probleemrelevante informatie.

Methoden

Proefpersonen

48 studenten van een opleiding voor diëtetiek en toegepaste huishoudkunde namen deel aan Experiment 2: 42 vrouwen en 6 mannen.

Alle proefpersonen hadden een vergelijkbare opleiding: HAVO, met biologie als eindexamenvak. De gemiddelde leeftijd van de proefpersonen was 19,4 jaar met een standaardafwijking van 1,6.

De proefpersonen ontvingen een vergoeding voor hun medewerking.

Materiaal

Het materiaal dat bij Experiment 2 gebruikt werd was identiek aan dat van Experiment 1. Daarnaast werd een tekst gebruikt.

De tekst was getiteld 'Osmose en diffusie'. Hij had een lengte van ongeveer 2220 woorden en was zodanig geschreven dat hij aansloot bij datgene wat op de HAVO aan de orde komt. In de tekst werden de volgende onderwerpen behandeld: diffusie, diffusiesnelheid, (semi)permeabiliteit van de celwand, concentratieverschillen, osmose, osmotische druk, osmotische waarde, structuur van plantencellen, turgor en plasmolyse.

Onderstrepingen, tussenkopjes en andere typografische signalen waren achterwege gebleven, opdat de proefpersonen geen ongewenste informatie zouden krijgen over het relatieve belang van de verschillende tekstdelen.

Constructie van een tekststructuur

Om de invloed van de organisatie van de tekst op de organisatie van de recall te kunnen onderzoeken werd de tekst eerst in proposities opgedeeld met behulp van dezelfde regels als in Experiment 1 gebruikt voor de identificatie van proposities in recall.

De interbeoordelaarsovereenstemming bij deze taak was 94%. Verschil van mening werd door middel van discussie opgelost.

De tekst bestond uit 241 proposities met een gemiddelde lengte van 9,2 woorden. In tabel 8 zijn bij wijze van voorbeeld de eerste elf proposities van de tekst afgedrukt.

Tabel 8: Propositions waarmee de tekst 'Osmose en diffusie' begint.

-
1. Osmose is het verschijnsel
 2. waarbij een sterkere oplossing water aantrekt uit een zwakkere oplossing,
 3. wanneer deze oplossingen gescheiden zijn door een half-doorlaatbare wand.
 4. Het Griekse woord 'osmos' betekent aanzuiging.
 5. De aanzuiging van watermoleculen kan tot stand komen door middel van diffusie
 6. (diffundere betekent uitbreiden).
 7. Diffusie is een proces
 8. waarbij moleculen bewegen van een plaats met een hoge concentratie naar een plaats met een lage concentratie.
 9. Stoffen zijn opgebouwd uit moleculen.
 10. De moleculen zijn niet statisch in de stof verankerd
 11. maar zijn steeds in beweging.
-

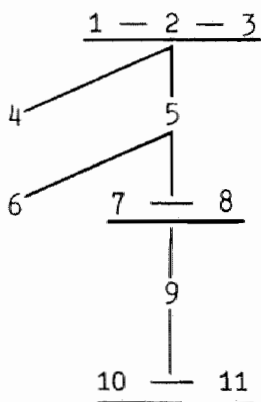
Vervolgens werd voor elke propositie nagegaan aan welke andere propositie(s) hij gerelateerd is. De meeste proposities bevatten expliciete verwijzingen naar andere proposities. Ze hebben één of meerdere begrippen gemeen of verwijzen naar elkaar door middel van aanwijzende voornaamwoorden. Wanneer een propositie geen expliciete verwijzing bevatte, was uit de context op te maken aan welke andere propositie hij gerelateerd was.

Twee soorten relaties tussen proposities werden onderscheiden: dominantierelaties en sequentierelaties. In een dominantierelatie is de dominante propositie algemener dan de ondergeordende propositie. De laatste geeft vaak een specificatie of een nadere uitwerking van de dominante propositie. Een sequentierelatie tussen twee proposities duidt een relatie aan waarbij de ene propositie de antecedent is en de andere de consequent. Ook temporele relaties tussen proposities werden als sequentierelaties beschouwd.

Een analyse als hierboven beschreven levert resultaten op die in belangrijke mate vergelijkbaar zijn met de produkten van de analysemethoden van Meyer (1975) en Grimes (1975). De aanpak is minder gedetailleerd — en dus minder bewerkelijk — dan de benadering van Frederiksen (1972) of die van Kintsch en Van Dijk (1978) die een tekst in microproposities opdelen.

Ten slotte werden de relaties tussen de proposities afgebeeld in een deels hiërarchische, deels seriële structuur, die door Meyer (1975) de inhoudsstructuur genoemd wordt. (Hier zal hij verder eenvoudigweg de tekststructuur genoemd worden.) Deze laatste procedure leverde een overeenstemming tussen beoordelaars op van 85,5%. In figuur 1 is een klein deel van de tekststructuur van de tekst 'Osmose en diffusie' ter illustratie afgebeeld.

Figuur 1: Fragment van de tekststructuur van 'Osmose en diffusie'



Ten slotte werd de tekst onderverdeeld in probleemrelevante en niet-probleemrelevante proposities (Interbeoordelaarsovereenstemming: 73%).

Procedure

De procedure kwam in grote lijnen overeen met die in Experiment 1. De experimentele conditie bestond uit 23 proefpersonen en de controleconditie uit 25 proefpersonen. Na randomisering werd gecontroleerd in hoeverre beide groepen vergelijkbaar waren in biologische voorkennis. Het gemiddelde eindcijfer voor biologie was voor de controlegroep 6,36 (standaarddeviatie 0,81); voor de probleemanalysegroep was dat 6,39 (standaarddeviatie 0,94).

De probleemanalysegroep werd opgedeeld in drie kleinere groepen. Elk van deze groepen kreeg aselect een ervaren tutor toegewezen. Tutor en groep werkten op dezelfde wijze als beschreven voor Experiment 1. Alle groepen hadden minder dan 10 minuten nodig voor de analyse.

Vervolgens kregen probleemanalyse- en controlegroep de tekst voorgelegd, met de opdracht deze te bestuderen. De studietijd was 15 minuten.

Ten slotte schreven de proefpersonen op wat zij zich van de tekst herinnerden. Hun recallprotocols werden verder geanalyseerd.

Na twee weken vond een tweede meting plaats. Om te voorkomen dat proefpersonen in de tussenliggende periode extra studie aan het onderwerp zouden besteden werd hen na afloop van de eerste meting meegedeeld dat bij de tweede bijeenkomst nog enkele persoonlijkheidsvragenlijsten zouden worden afgenomen (Frederiksen, 1975).

Identificatie van proposities in de recall

Twee beoordelaars splitsten de door de proefpersonen geproduceerde teksten op in proposities (interbeoordelaarsovereenstemming: 92%). Van deze proposities werd vervolgens aangegeven of het om een letterlijke weergave (of parafrase) van een tekstpropositie handelde, of dat het een inferentie uit de tekst betrof. Als het om een letterlijke weergave (of een parafrase) ging, werd aangegeven welke propositie in de tekst ermee correspondeerde. Letterlijke weergaven of parafrases (in de verdere tekst zal nog alleen dit laatste begrip gebruikt worden omdat letterlijke weergaven van proposities bijna niet voorkomen) konden correct, half-correct (bijvoorbeeld als maar een deel was gereproduceerd), of incorrect zijn. Inferenties werden als correct of incorrect gecodeerd.

De interbeoordelaarsovereenstemming voor deze procedure was 76%. Met name bleek het soms moeilijk overeenstemming te bereiken over de vraag of een bepaalde propositie een inferentie was of eigenlijk een parafrase van meerdere proposities tegelijk.

Vergelijking van de organisatiestructuur van recall

Om na te kunnen gaan in hoeverre de organisatie van recall van verschillende groepen gelijkenis vertoont met de tekststructuur en met elkaar werd gebruik gemaakt van de Kwadratische Toewijzingsprocedure ('Quadratic Assignment Procedure': QAP) van Schultz en Hubert (1976). QAP is een methode waarmee de gelijkenis van twee matrices bepaald kan worden. De loopindex van beide assen van een QAP-matrix wordt gevormd door de

tekstproposities. In de resulterende cellen wordt aangegeven of er een relatie tussen twee proposities bestaat of niet.¹

De tekststructuur van 'Osmose en diffusie' werd ondergebracht in een 241-bij-241-matrix. In figuur 2 zijn de relaties tussen de eerste elf proposities in een QAP-matrix opgenomen.

Figuur 2: Tekststructuurmatrix voor figuur 1

	1	2	3	4	5	6	7	8	9	10	11
1		1	1	1	1						
2			1	1	1						
3				1	1						
4											
5						1	1	1			
6											
7								1	1		
8									1		
9										1	1
10											1
11											

Een lege cel duidt aan dat er geen relatie tussen beide proposities opgemerkt werd. Een cel met de waarde 1 erin duidt aan dat er wel een relatie was.

Vervolgens werden twee *datamatrices* gevormd; één voor de probleemanalysegroep en één voor de controlegroep. De cellen van beide matrices werden gevuld met *frequenties*; namelijk de frequentie waarmee een bepaalde relatie tussen twee proposities gevonden werd (in de recall). Dat konden dus ook andere relaties zijn dan gedefinieerd in de tekststructuurmatrix, wanneer de recall van proefpersonen daartoe aanleiding gaf.

Drie vergelijkingen werden uitgevoerd. De *datamatrices* werden ieder afzonderlijk vergeleken met de tekststructuurmatrix en vervolgens met elkaar.

¹ QAP maakt geen onderscheid mogelijk naar de aard van de relaties: dominantie of sequentie.

Mathematisch is de procedure relatief eenvoudig. Een statistiek Gamma wordt berekend door corresponderende cellen uit twee matrices met elkaar te vermenigvuldigen en deze produkten over alle cellen heen op te tellen. Gamma kan beschouwd worden als een niet-genormaliseerde correlatie tussen de beide matrices. Deze statistiek wordt vervolgens getest tegen een verwachte waarde van Gamma die berekend is onder de aanname dat de beide matrices van elkaar onafhankelijk zijn. De uitslag van deze test maakt dus de beslissing mogelijk of de beide relatiepatronen vergelijkbaar zijn of niet. Zie voor technische details Schultz en Hubert (1976).

Resultaten en discussie

Maten voor bruto en netto output; eerste afname

De eerste analyse had betrekking op het totaal aantal geproduceerde proposities. In tabel 9 zijn gemiddelden en standaarddeviaties opgenomen.

Tabel 9: Gemiddelde aantallen correcte en incorrecte proposities (met standaarddeviaties), geproduceerd in Experiment 2

	$\bar{X}$	SD	N
Probleemanalysegroep	45,87	12,64	23
Controlegroep	32,28	12,89	25
Totaal	38,79	12,77	48

Het gevonden verschil is statistisch significant. De variantieanalyse leverde het volgende resultaat op: $F(1,46) = 13,56$; $p < 0,0007$. De punt-biseriële correlatie tussen de conditie waaronder de proefpersonen gewerkt hebben (probleemanalyse versus geen probleemanalyse) en het totaal aantal door hen geproduceerde proposities is $r_{pbis} = 0,48$.

Ook het aantal incorrecte proposities verschilt per conditie. De probleemanalysegroep produceert significant meer fouten: $F(1,46) = 7,65$; $p < 0,009$. Dit verschil is echter proportioneel gelijk aan het verschil in het totaal aantal geproduceerde proposities.

In tabel 10 zijn de gemiddelde aantallen correcte proposities opgenomen, onderverdeeld in gemiddelde aantallen parafrases en inferenties.

Tabel 10: Gemiddelde aantallen correcte proposities, parafrases en inferenties (standaarddeviaties tussen haakjes), geproduceerd in Experiment 2

	Proposities	Parafrases	Inferenties	N
Probleemanalysegroep	38,63 (10,87)	32,76 (9,54)	5,87 (3,33)	23
Controlegroep	27,58 (11,55)	22,42 (9,43)	5,16 (3,72)	25
Totaal	32,88 (12,44)	27,38 (9,48)	5,50 (3,54)	48

De probleemanalysegroep produceert significant meer correcte propositie dan de controlegroep: $F(1,46) = 11,59$; $p < 0,002$ ($r_{pbis} = 0,45$). Dat komt vooral omdat deze groep meer parafrases reproduceert: $F(1,46) = 14,26$ $p < 0,0006$ ($r_{pbis} = 0,49$) dan de controlegroep, want de gemiddelde aantallen geproduceerde inferenties verschillen niet van elkaar: $F(1,46) = 0,48$ $p = 0,49$.

Deze resultaten laten een aantal conclusies toe.

De eerste is dat activatie van voorkennis blijkbaar de verwerking van een tekst, relevant ten aanzien van die voorkennis, aanzienlijk vergemakkelijkt. Proefpersonen die aan het probleem gewerkt hebben, produceren 1,4 keer meer proposities dan proefpersonen die dat niet hebben gedaan. De experimentele manipulatie verklaart tussen de 23% en de 25% van de variatie in prestatie tussen de proefpersonen. Dat is tamelijk veel, als men het vergelijkt met de effecten van andere leerstrategieën die de verwerking van tekst beïnvloeden. Deze verklaren door de bank genomen niet meer variatie dan ongeveer 10% à 15%.

De tweede conclusie is, dat het feit dat het aantal geproduceerde inferenties niet verschilt, consequenties heeft voor een van de veronderstellingen omtrent het onderliggende cognitieve mechanisme dat voor de gevonden effecten verantwoordelijk zou kunnen zijn. Een van de in de inleiding to dit experiment geformuleerde hypothesen was dat de probleemanalysegroep meer elaboraties zou produceren tijdens het verwerkingsproces (Reder, 1980) die op hun beurt bijdragen tot het ontstaan van 'multiple redundant retrieval paths' in het lange-termijn-geheugen. Frederiksen (1975) en Mayer en Bromage (1980) constateerden dat elaboratie tijdens het verwerken van de tekst tot inferenties in de cognitieve representatie van die tekst leidt. Dat betekent dat inferenties in de recall teruggevonden moeten worden. De recall van de probleemanalysegroep bevat inferenties; echter niet meer dan die van de controlegroep. De conclusie moet getrokken worden dat de probleemanalysegroep niet méér elaboreert op de tekst dan de controlegroep, en dat de elaboratiehypothese ter verklaring van het gevonden effect dus niet erg plausibel lijkt.

Maten voor de organisatie van recall; eerste afname

Een alternatieve mogelijkheid, hiervoor reeds aangestipt, was dat de probleemanalysegroep de tekstinformatie op een andere manier organiseert dan de controlegroep. Dat zou onder andere tot uitdrukking kunnen komen in de aard van de parafrases die uit de tekst gereproduceerd worden (De hierna volgende analyses blijven beperkt tot parafrases, omdat inferenties zich moeilijk aan specifieke tekstproposities lieten relateren.)

In de eerste twee analyses speelt de hiërarchische structuur van de tekst — waarvan in figuur 1 een fragment is afgebeeld — een belangrijke rol. In de daarop volgende analyses gaat het om de recall van relationele structuren. In de vorige paragraaf is bij de analyse aan iedere parafrase hetzelfde gewicht toegekend. Dat wil zeggen dat er geen onderscheid gemaakt is naar de belangrijkheid van de gereproduceerde proposities. Reproductie van de

tails en bijzaken woog even zwaar als reproductie van de centrale ideeën uit de tekst. Het is echter niet uitgesloten dat de conditie waaronder proefpersonen de tekst bestudeerden nu juist van invloed was op de soort proposities die onthouden werden. Om deze hypothese te onderzoeken is de analyse van de output aan parafrases herhaald, waarbij de positie die de herinnerde proposities in de tekststructuur innemen, in de analyse verdisconteerd is.

Daartoe zijn in de tekststructuur zes niveaus onderscheiden. Parafrases van proposities afkomstig uit de hoogste twee niveaus (de centrale ideeën uit de tekst) werden drie maal zo zwaar gewogen als parafrases uit de laagste twee niveaus (de details en de bijzaken). Parafrases uit de middelste twee niveaus werden twee maal zo zwaar gewogen als die uit de laagste niveaus. Een vergelijkbare aanpak kan worden gevonden in Meyer en McConkie (1973). De resulterende scores werden per protocol, en over proefpersonen, gesommeerd. In tabel 11 zijn de resultaten samengevat.

Tabel 11: Gemiddelde aantallen parafrases (met standaarddeviaties), gewogen naar positie in de tekststructuur, gereproduceerd in Experiment 2

	$\bar{X}$	SD	N
Probleemanalysegroep	48,22	13,73	23
Controlegroep	32,86	13,64	25
Totaal	40,22	13,68	48

Ook deze verschillen zijn statistisch significant: $F(1,46) = 15,09; p < 0,0003$ ($r_{pbis} = 0,50$).

Vergelijkt men dit resultaat met dat van de ongewogen parafrases in tabel 10, dan lijkt het erop dat het gewogen verschil nog een fractie groter is. Dat zou erop kunnen duiden dat activatie van voorkennis vooral invloed uitoefent op het onthouden van de centrale ideeën uit een tekst.

Een meer directe test van deze veronderstelling wordt mogelijk door inspectie van de recallwaarschijnlijkheden van de tekstproposities.

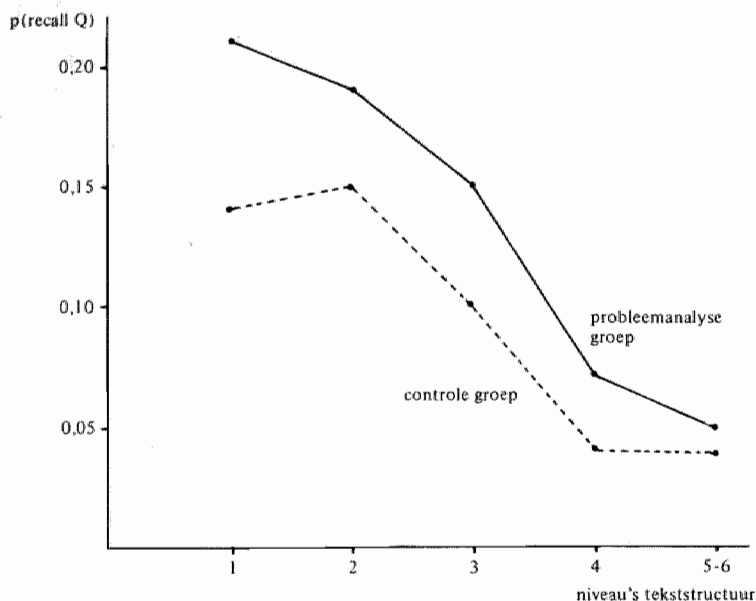
De *recallwaarschijnlijkheid* van een propositie kan worden gedefinieerd als de relatieve frequentie waarmee die propositie in de recall terug te vinden is. Hij wordt berekend met behulp van de volgende formule:

$$p(\text{recallpropositie } Q) = \frac{\text{frequentie van } Q \text{ in recall}}{\text{totaal aantal proefpersonen}}$$

Voor elke van de proposities in de tekst werd een recallwaarschijnlijkheid berekend, en wel voor probleemanalyse- en controlegroep afzonderlijk. Vervolgens werden recallwaarschijnlijkheden binnen elk van de niveaus van de tekststructuur gemiddeld. In figuur 3 zijn de recallwaarschijnlijkheden

heden voor elk van de niveaus en voor beide groepen afzonderlijk afgebeeld (niveau 5 en 6 zijn samengevoegd). Niveau 1 representeert het niveau van de centrale ideeën in de tekst.

Figuur 3: Gemiddelde recallwaarschijnlijkheden voor tekstproposities per niveau van de tekststructuur



Uit deze figuur blijkt dat activatie van bestaande cognitieve structuren het sterkst het onthouden van de belangrijkste ideeën uit een tekst bevordert. De verschillen tussen probleemanalyse- en controlegroep nemen af naarmate parafrases afkomstig zijn uit lagere regionen van de tekststructuur. De conclusie zou dus kunnen zijn dat actualisatie van voorkennis leidt tot een zodanige organisatie van de informatie uit tekst in het geheugen dat vooral de centrale ideeën uit een tekst makkelijker toegankelijk worden bij recall.

Overigens — maar dit is een tussenopmerking — repliceren de hier gevonden resultaten de bevindingen van Meyer (1975; 1977) dat de centrale boodschap van een tekst door proefpersonen beter onthouden wordt dan bijzaken en details. In dit experiment was dat verschil een factor drie tot vier. Proefpersonen zijn blijkbaar in staat macroproposities uit een tekst te abstraheren.

Leggen proefpersonen wier voorkennis geactiveerd is, andere relaties tussen de proposities uit tekst dan anderen? Met andere woorden: doorbreekt activatie het in de tekst gegeven patroon van relaties tussen proposities, en zo ja, op welke wijze? Dat is de volgende vraag die aan de orde is. Bij de beantwoording daarvan staat dus niet centraal welke proposities onthouden worden (zoals in de vorige alinea's het geval was), maar *welke relaties* tussen proposities onthouden worden.

Daartoe werden de datamatrices van beide groepen met elkaar en met de tekststructuurmatrix vergeleken. De resultaten daarvan waren als volgt. De tekststructuurmatrix correleerde significant zowel met de datamatrix van de probleemanalysegroep ($p < 0,03$) als met die van de controlegroep ($p < 0,001$). Ook beide datamatrices correleerden significant met elkaar ($p < 0,05$). Dat betekent dat activatie van bestaande kennisstructuren niet tot een wezenlijk andere organisatie van de relaties tussen proposities uit een tekst leidt.

Beide groepen gebruiken blijkbaar de structuur van de tekst als 'kapstok' om de relaties tussen proposities op te slaan in het geheugen. Dit resultaat is niet in overeenstemming met datgene wat door Matthews (1982) gevonden werd. Matthews vond dat proefpersonen die vooraf tekstrelevante informatie gekregen hadden, de informatie uit de tekst op een andere wijze opsloegen dan proefpersonen die alleen de tekst bestudeerd hadden. De oorzaak van deze discrepantie is niet op het eerste gezicht duidelijk. Mogelijk wordt het verschil veroorzaakt door het verschil in gebruikte procedures.

Ten slotte werd nog onderzocht wat de invloed van de probleemanalyse was op het onthouden van probleemrelevante versus niet-probleemrelevante informatie. Daartoe was de verzameling tekstproposities opgedeeld in een subset probleemrelevante en een subset niet-probleemrelevante proposities (niet-probleemrelevant moet hier verstaan worden als min of meer neutraal ten opzichte van het probleem).

In tabel 12 zijn gemiddelde aantallen relevante en niet-relevante parafrases uit de recall van beide groepen opgenomen.

Tabel 12: Gemiddelde aantallen probleemrelevante en niet-probleemrelevante parafrases, gereproduceerd in Experiment 2

	Probleemrelevante parafrases	Niet-probleemrelevante parafrases
Probleemanalysegroep	23,87	8,89
Controlegroep	14,69	7,73
Totaal	19,09	8,29

In deze data wordt iets zichtbaar wat op een sturend effect lijkt van de probleemanalyse op de verwerking van de tekst. De verschillen tussen de groepen zijn duidelijk groter waar het de probleemrelevante parafrases betreft dan waar het de neutrale parafrases betreft.

De selectieve-aandacht-hypothese voorspelt een intensievere verwerking van probleemrelevante proposities. Echter, deze hypothese voorspelt ook dat gegeven een standaard hoeveelheid beschikbare studietijd aandacht voor relevante proposities *ten koste gaat* van aandacht voor niet-relevante proposities. De probleemanalysegroep zou dus op die proposities een *slechtere* prestatie moeten leveren. Dit is in de hier gerapporteerde data niet het geval.

De onduidelijkheid wordt nog verhoogd als men bedenkt dat wat hier probleemrelevante proposities genoemd zijn, deels samenvallen met de proposities in de eerste drie niveaus van de tekststructuur. Het is dus zeer wel mogelijk dat het in tabel 12 te constateren fenomeen niet zo zeer het gevolg is van selectieve aandacht als gevolg van probleemanalyse, als wel van intensievere verwerking vanwege de positie van de proposities in de tekststructuur. Met andere woorden: het hier geconstateerde fenomeen is in feite een herformulering van wat hiervoor al vastgesteld werd, namelijk dat de verschillen tussen beide groepen met betrekking tot de centrale proposities uit de tekst groter zijn dan met betrekking tot de details. De verheldering van de vraag of de betere prestatie op de probleemrelevante proposities het gevolg is van selectieve aandacht of van het abstraheren van macroproposities kan met behulp van de in dit experiment verzamelde data niet beantwoord worden. Het vereist een experiment waarin de invloed van beide processen van elkaar gescheiden wordt.

Maten voor bruto en netto output: tweede meting

Twee weken na de eerste meting van recall werd een tweede meting gedaan. De resultaten daarvan zijn in tabel 13 en 14 samengevat.

Tabel 13: Gemiddelde aantallen correcte én incorrecte proposities (met standaarddeviaties), geproduceerd in Experiment 2 (tweede meting)

	$\bar{X}$	SD	N
Probleemanalysegroep	37,52	9,05	23
Controlegroep	27,32	7,38	25
Totaal	32,20	8,22	48

Tabel 14: Gemiddelde aantallen correcte proposities, onderverdeeld naar parafrases en inferenties (met standaarddeviaties), geproduceerd in Experiment 2 (tweede meting)

	Proposities	Parafrases	Inferenties	N
Probleemanalysegroep	29,46 (8,15)	24,28 (7,35)	5,17 (4,12)	23
Controlegroep	22,08 (6,22)	19,04 (5,78)	3,04 (2,30)	25
Totaal	25,61 (7,21)	21,55 (6,58)	4,06 (3,44)	48

Op deze gemiddelden zijn variantieanalyses uitgevoerd. De resultaten daarvan zijn als volgt: Totaal aantal proposities: $F(1,46) = 18,43$; $p < 0,0002$. Aantal correcte proposities: $F(1,46) = 12,54$; $p < 0,001$. Aantal parafrases: $F(1,46) = 7,60$; $p < 0,009$. Aantal inferenties: $F(1,46) = 5,02$; $p < 0,04$. Deze resultaten komen niet geheel onverwacht.

De effecten van de probleemanalyse zijn zichtbaar gebleven, zij het wat minder geprononceerd dan direct na de bestudering van de tekst. Op alle hier gerapporteerde variabelen is enig verval opgetreden, met uitzondering van het gemiddeld aantal gemaakte fouten. Dat steeg iets, van 5,92 naar 6,59. Het gemiddeld aantal inferenties bleef min of meer stabiel; zij het dat er nu wél een klein doch significant verschil tussen beide groepen gevonden wordt in het voordeel van de probleemanalysegroep.

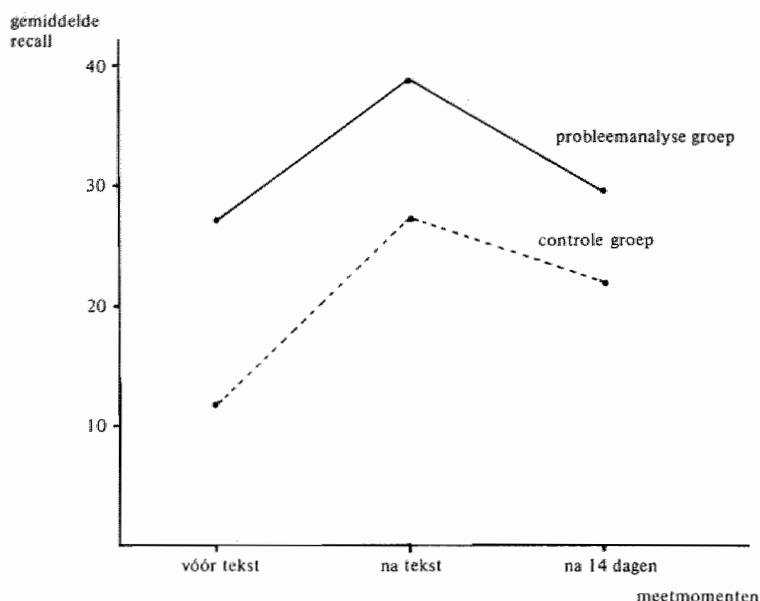
De punt-biseriële correlatie tussen de experimentele manipulatie en het totaal aantal geproduceerde proposities is gelijk aan $r_{pbis} = 0,53$. Tussen de eerstgenoemde variabele en het aantal correcte proposities is hij gelijk aan $r_{pbis} = 0,46$. Het percentage door de experimentele ingreep verklaarde variantie schommelt dus nog steeds rond de 25.

In figuur 4 zijn ter illustratie van het verloop van het leerproces de gemiddelde frequenties correcte proposities uit *beide* experimenten in een grafiek gebracht. (Daarbij is gemakshalve van de veronderstelling uitgegaan dat de recall van de controlegroep uit Experiment 1 een zuivere schatting van de voorkennis van alle groepen vormt.)

Het bezien van deze leercurves maakt twee dingen duidelijk. De probleemanalysegroep heeft op alle meetmomenten de beschikking over een groter kennisbestand. De controlegroep profiteert echter meer van de bestudering van de tekst, en vertoont ook wat minder verval dan de probleemanalysegroep. Dat duidt erop dat een deel van het effect van voorkennis-activatie kunstmatig is en na verloop van tijd verloren gaat. Na 14 dagen is de probleemanalysegroep bijna weer terug op het niveau dat door de activatie alleen al bereikt was.

De proefpersonen in de hier gerapporteerde experimenten hadden nogal veel voorkennis voordat ze aan de verschillende experimentele condities werden blootgesteld. Het zou uitermate interessant zijn na te gaan wat de invloed van probleemanalyse zou zijn op het leerproces van proefperso-

Figuur 4: Gemiddeld aantal correcte proposities met betrekking tot osmose en diffusie vóór, direct na en 14 dagen na het bestuderen van een tekst over dat onderwerp



nen met minder specifieke voorkennis, die bij het analyseren van het probleem meer gebruik zouden moeten maken van algemene wereldkennis. Resultaten die Mayer verkreeg met verschillende typen leerstrategieën (Mayer, 1980; 1982) suggereren dat personen met weinig specifieke voorkennis meer gebaat zijn met faciliterende leerstrategieën dan personen met veel voorkennis.

Maten voor de organisatie van recall: tweede meting

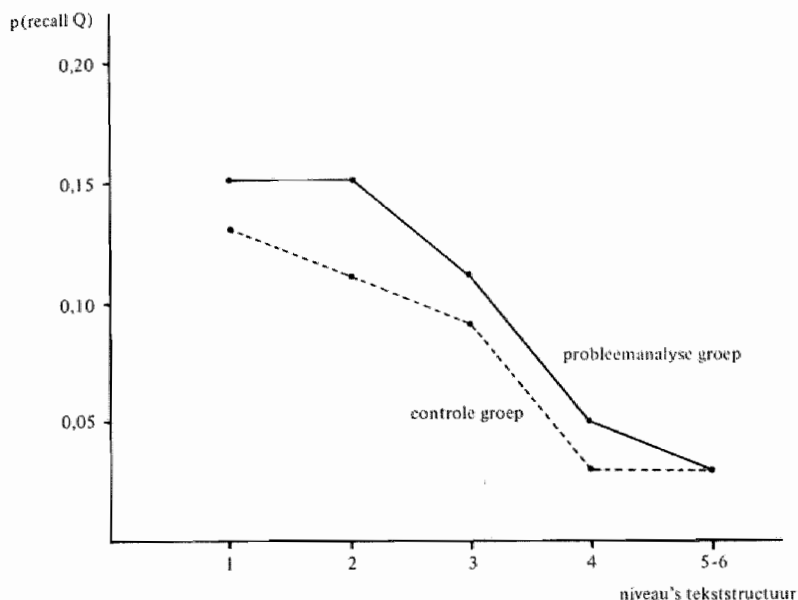
In tabel 15 zijn de gemiddelde aantallen parafrases opgenomen, gewogen naar de positie die corresponderende proposities in de tekststructuur innemen.

Tabel 15: Gemiddelde aantallen gewogen parafrases (met standaarddeviaties), gereproduceerd in Experiment 2 (tweede meting)

	$\bar{X}$	SD	N
Probleemanalysegroep	37,17	9,61	23
Controlegroep	31,26	10,60	25
Totaal	34,09	10,14	48

Het verschil tussen gemiddelden is significant: $F(1,46) = 4,07$; $p < 0,05$, maar minder geprononceerd dan bij de eerste meting. De probleemanalyse-groep vertoont duidelijk meer verval dan de controlegroep. Het feit dat de gewogen score daar sterker onder lijdt dan de ongewogen parameters, duidt erop dat het verval vooral optreedt met betrekking tot de centrale ideeën in de tekst. We kunnen dit directer constateren aan de hand van de gemiddelde recallwaarschijnlijkheden die in figuur 5 zijn opgenomen.

Figuur 5: Gemiddelde recallwaarschijnlijkheden voor tekstproposities per niveau van de tekststructuur (tweede meting)



Vergelijking met figuur 3 leert dat het effect van activatie van bestaande cognitieve structuren op het onthouden van de centrale ideeën uit een tekst enigszins teruggelopen is, vergeleken met het effect dat die activatie heeft op het onthouden van de details.

Hoe staat het nu met het onthouden van de in de tekst aanwezige relaties tussen proposities?

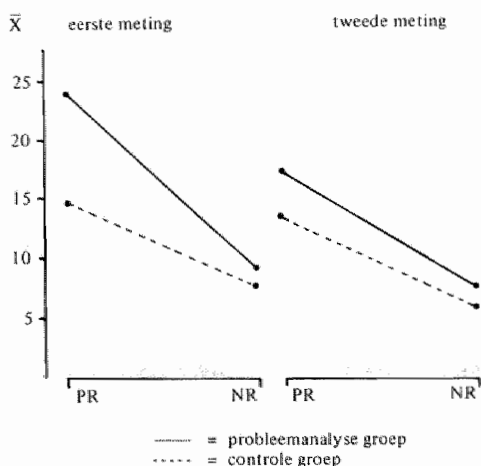
Vergelijking van correlaties tussen de matrices met die van de eerste meting laat zien dat er een zekere convergentie is opgetreden. De correlaties zijn wat hoger. Dat wijst erop dat de vergeleken relatiestructuren (nog) meer op elkaar zijn gaan lijken.

Uit deze gegevens kan afgeleid worden dat kennis van osmose en diffusie bij beide groepen proefpersonen niet wezenlijk anders georganiseerd is in het lange-termijn-geheugen. De relatiestructuur tussen concepten is in belangrijke mate dezelfde. Er is bij de probleemanalysegroep enkel sprake van een zodanige organisatie dat *meer* concepten toegankelijk zijn bij recall. Dit moet welhaast betekenen dat de relaties tussen de verschillende proposities of concepten in het lange-termijn-geheugen van proefpersonen uit de experimentele groepen hechter zijn. Dat wil zeggen dat activatie van het ene concept makkelijker tot activatie van een ander, gerelateerd, concept leidt (Anderson en Bower, 1973). Deze hechtheid zou dan moeten voortvloeien uit de door activatie van voorkennis geboden mogelijkheid meer informatie per tijdseenheid cognitief te verwerken (Spilich e.a., 1979).

De laatste veronderstelling omtrent de rol van probleemanalyse bij de verwerking van tekst die in de inleiding op Experiment 2 naar voren gebracht werd, was dat de probleemanalyse als een retrievalsignaal (Tulving en Thomson, 1973) zou kunnen functioneren. Het idee daarachter is dat beide groepen de tekst op dezelfde wijze coderen, maar dat de probleemanalysegroep bij de retrieval in het voordeel is, omdat tegelijk met de informatie uit tekst contextkenmerken zijn opgeslagen (de analyse van het probleem) die het terughalen van de informatie vergemakkelijken.

De consequentie die uit deze hypothese werd afgeleid is dat probleemrelevante informatie relatief minder snel onbereikbaar zou moeten worden voor retrieval dan niet-probleemrelevante informatie. In figuur 6 zijn gemiddelde scores voor probleemrelevante en niet-probleemrelevante informatie bij de eerste en tweede meting tegen elkaar afgezet.

Figuur 6: Probleemrelevante informatie (PR) vergeleken met niet-relevante informatie (NR) bij twee metingen in Experiment 2



De retrievalsignaal-hypothese voorspelt dat bij de tweede meting de hoek tussen de lijnen die de prestaties van beide groepen aangeven, groter zou moeten worden; dat wil zeggen dat het verval van relevante informatie bij de probleemanalysegroep relatief minder groot zou moeten zijn. Dit blijkt niet het geval. Er zijn dus geen aanwijzingen dat probleemanalyse als retrievalsignaal het terughalen van de informatie bevordert. Het effect van de probleemanalyse moet dus eerder bij de codering van informatie gezocht worden dan bij de retrieval. Dit is in overeenstemming met wat in ander onderzoek gevonden is (Dooling en Lachman, 1971; Dooling en Mullet, 1973; Mayer en Bromage, 1980).

Algemene discussie

De resultaten van twee experimenten waarin de rol van een methode van probleemanalyse (Schmidt, 1982b) bij de verwerking van een probleem-relevante tekst onderzocht werd, kunnen als volgt worden samengevat. Probleemanalyse in een kleine groep leidt tot activatie en herstructurering van voorkennis, relevant ten aanzien van de processen die aan het probleem ten grondslag liggen. In Experiment 1 bleek dat proefpersonen die een probleem met betrekking tot het gedrag van een bloedcel in water en in een zoutoplossing geanalyseerd hadden, ruim twee maal zoveel proposities over het onderwerp osmose en diffusie produceerden dan een controlegroep.

Een analyse van de herkomst van deze proposities, waarbij de inhoud van de recallprotocols vergeleken werd met datgene wat tijdens de discussie over het probleem gezegd was, bracht aan het licht dat gemiddeld ongeveer 40% van de proposities die de persoon zich herinnert, door hemzelf tijdens de discussie geproduceerd is. Ongeveer 20% van wat hij produceert, is tijdens de analyse van anderen geleerd en in de eigen cognities geïntegreerd. Opmerkelijk is dat ongeveer 40% van de proposities die in de recall geproduceerd wordt niet afkomstig is uit de discussie. Blijkbaar wordt lang niet alle voorkennis tijdens de analyse van het probleem geëxpliciteerd; mogelijk omdat personen ervan uitgaan dat andere deelnemers al over die informatie beschikken (Schmidt, 1979). Ook is het niet uitgesloten dat de recalltaak op zichzelf nog weer nieuwe voorkennis mobiliseert. Activatie van voorkennis werd in Experiment 1 geoperationaliseerd als datgene wat de persoon in de discussie zelf produceert, terwijl herstructurering geoperationaliseerd werd als datgene wat tijdens de discussie van anderen geleerd werd. Deze operationalisatie werd als niet volledig adequaat beoordeeld, omdat ook in datgene wat iemand zelf produceert al een element van herstructurering kan zitten. Een zorgvuldiger afweging van het relatieve aandeel van activatie, en herstructurering in de processen die object van deze studie waren, kan verkregen worden door herhaling van het experiment met een extra controlegroep waarin proefpersonen *individueel* het probleem analyseren.

De nadruk in de probleemanalyse ligt op de *verklaring* van verschijnselen

in termen van een onderliggend proces. In de inleiding werd betoogd dat het vooral het begrijpen van aan de werkelijkheid ten grondslag liggende processen of principes, eerder dan de beschrijving van die werkelijkheid, het belangrijkste doel van onderwijs zou moeten zijn, omdat begrip van de wereld mensen meer dan wat ook in staat stelt adequaat te functioneren in die wereld. In die opvatting is het belangrijker dat studenten begrijpen wat het osmotische proces inhoudt en hoe het verloopt, dan dat ze weten dat een 0,9%-NaCl-oplossing in water isotonisch is met de inhoud van een levende cel.

De verwachting was dat de nadruk die in de probleemanalyse op verklaring gelegd werd, ertoe zou leiden dat in de recallprotocols van de probleemanalysegroep relatief meer verklarende — en minder descriptieve — proposities zouden verschijnen dan in de protocols van de controlegroep. Dat bleek inderdaad het geval. In absolute aantallen produceerde de probleemanalysegroep bijna drieëneenhalf keer meer verklarende proposities (en ruim anderhalf keer meer descriptieve). Dit werd als een gunstig effect gezien, vooral omdat Bromage en Mayer (1981) aangetoond hebben dat er een vrij hoge correlatie is tussen het aantal verklarende proposities in recall en prestaties op transferproblemen.

In Experiment 2 werd de invloed van activatie en herstructurering op de cognitieve verwerking van een verwante tekst onderzocht. Uit de resultaten van dit experiment bleek duidelijk dat geactiveerde voorkennis wordt ingezet bij het begrijpen en onthouden van informatie uit tekst en dat het daarbij om een robuust effect gaat. Uit de correlaties tussen de experimentele manipulatie en studieprestatie bleek dat de aan- of afwezigheid van probleemanalyse ongeveer 25% van de variantie in recallscores verklaarde. Een meting na twee weken bracht aan het licht dat het ook om een stabiel effect handelt, al waren de verschillen tussen beide groepen wat teruggelopen.

Het effect van activatie strekt zich uit over de hele tekst, maar lijkt relatief het sterkst te zijn bij de recall van de centrale ideeën en concepten uit de tekst. Met name bij de meting direct na de bestudering van de tekst werd dit effect gevonden. Geconcludeerd werd dat de probleemanalysegroep de informatie uit tekst blijkbaar zodanig organiseerde dat vooral de belangrijkste informatie, de kern van de tekst, bij recall gemakkelijker uit het lange-termijn-geheugen kan worden teruggehaald.

Dit retrievalvoordeel kan niet verklaard worden uit het feit dat personen wier bestaande cognitieve structuren geactiveerd worden, andere relaties tussen de verschillende tekstproposities leggen, en *daardoor* in het voordeel zouden zijn. In tegenstelling tot wat Matthews (1982) constateerde, worden de relaties tussen proposities door voorkennisactivatie niet op een andere wijze in het lange-termijn-geheugen vastgelegd, tenminste niet in het hier gerapporteerde experiment. Beide groepen gebruikten de in de tekst gegeven relatiestructuur als medium om de verwerkte informatie mee te organiseren.

De organisatie van de recall bleek in belangrijke mate overeen te stemmen met de organisatie van de tekst.

Tenzij de structuur van de voorkennis in sterke mate afwijkt van de structuur van de tekst, mag men veronderstellen dat personen de in de tekst gegeven relatiestructuur gebruiken om kennis op te slaan. Dat kan men afleiden uit het feit dat Meyer en haar collega's (Meyer, 1975; Britton, Meyer, Hodge en Glynn, 1980) in de recall zelden ernstige verstoringen vinden van de structuur van de tekst.

De analyse van het onthouden van probleemrelevante versus niet-probleemrelevante informatie, om na te kunnen gaan in hoeverre probleemanalyse de aandacht van proefpersonen richt op bepaalde delen van de tekst, leverde een weinig helder beeld op. Probleemrelevante informatie werd beter herinnerd dan niet-probleemrelevante, maar omdat deze variabele sterk gecorreleerd was aan de positie die proposities in de tekststructuur innamen, kan dit effect ook teruggevoerd worden op het proces dat in de inleiding van deze bijdrage het afleiden van macroproposities uit tekst genoemd werd. Tegen de selectieve-aandacht-hypothese sprak, dat niet-probleemrelevante informatie door de probleemanalysegroep niet *slechter* gereproduceerd werd (Melton, 1978; Rothkopf en Billington, 1979).

Geconcludeerd moet worden dat een definitieve beantwoording van de vraag of probleemanalyse de verwerking van tekst beïnvloedt via het proces van selectieve aandacht, afhankelijk is van een experiment waarin selectieve aandacht en afleiden van macroproposities van elkaar onderscheiden kunnen worden. Dat zou kunnen gebeuren door een probleem zodanig te kiezen dat voorkennis met betrekking tot onbelangrijke details van een tekst geactiveerd wordt. Wordt die informatie relatief beter onthouden, dan kan men concluderen dat selectieve aandacht in het spel is.

Probleemrelevante informatie vertoonde bij de tweede meting niet minder verval dan niet-probleemrelevante informatie. Daaruit werd geconcludeerd dat het effect van activatie van voorkennis niet ligt in de retrieval van informatie maar al in de codering van die informatie. Dit resultaat was in overeenstemming met wat voor andere voorkennis-activerende leerstrategieën gevonden werd (Mayer en Bromage, 1980).

De data, die in beide hier gerapporteerde experimenten verzameld zijn, suggereren dat activatie en herstructurering van voorkennis een algemeen facilitatief effect op de verwerking van tekst hebben. Activatie vooraf leidt er waarschijnlijk toe dat voorkennis tijdens het verwerkingsproces gemakkelijker toegankelijk is en daardoor de verwerkingscapaciteit van het werkgeheugen vergroot (Spilich e.a., 1979). De mogelijkheid meer informatie per tijdseenheid te verwerken bevordert de hechtheid van de relaties in het conceptuele netwerk, dat de interne representatie van de informatie uit tekst vormt (Anderson en Bower, 1973; Rumelhart en Ortony, 1977). Bij recall activeert de ene propositie in een dergelijk geval gemakkelijker de andere, wat de superieure prestatie van de probleemanalysegroep verklaart. Een elaboratie-interpretatie van de gegevens (Anderson en Reder, 1979) moet worden afgewezen op grond van het feit dat de frequentie van infe-

renties in de recall geen aanleiding gaf te veronderstellen dat de probleemanalysegroep meer elaboreerde tijdens het leesproces. Ook gezien de relatiestructuurdata is het onwaarschijnlijk dat de interne representatie van de informatie uit tekst bij de probleemanalysegroep méér, en andere, retrieval-paden bevat.

Literatuur

- Allen, V.L. en R.S. Feldman. Learning through tutoring: Low achieving children as tutors. *Journal of Experimental Education*, 1973, 42, 1-5.
- Anderson, J.R. *Language, Memory and Thought*. Hillsdale: Lawrence Erlbaum, 1976.
- Anderson, J.R. en G.H. Bower. *Human associative memory*. Washington, D.C.: Winston, 1973.
- Anderson, J.R. en L.M. Reder. An elaborative processing explanation of depth of processing. In L.S. Cermak en F.I.M. Craik (Eds.). *Levels of processing in human memory*. Hillsdale: Lawrence Erlbaum, 1979.
- Anderson, R.C. en W.B. Biddle. On asking people questions about what they are reading. In G.H. Bower (Ed.). *Psychology of learning and motivation* (vol. 9). New York: Academic Press, 1975.
- Anderson, R.C. en J.W. Pichert. Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 1978, 17, 1-12.
- Anderson, R.C., R.J. Spiro en M.C. Anderson. Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal*, 1978, 15, 433-440.
- André, E.D. en E.Z. Rothkopf. Text organization and learning goals. *Journal of Educational Psychology*, 1975, 67, 445-450.
- Ausubel, D.P. *The psychology of meaningful verbal learning*. New York: Grune and Stratton, 1963.
- Bargh, J.A. en Y. Schul. On the cognitive benefits of teaching. *Journal of Educational Psychology*, 1980, 72, 593-604.
- Barrows, H.S. en R.M. Tamblyn. *Problem-based learning*. New York: Springer, 1980.
- Bartlett, F.C. *Remembering*. Cambridge: Cambridge University Press, 1932.
- Bovy, R.C. Successful instructional methods: a cognitive information processing approach. *Educational Communication and Technology*, 1981, 29, 203-217.
- Bower, G.H., J.B. Black en T.J. Turner. Scripts in comprehension and memory. *Cognitive Psychology*, 1979, 11, 177-220.
- Bransford, J.D. en M.K. Johnson. Contextual prerequisites for understanding: some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 717-726.
- Bretzing, B.H. en R.W. Kulhavy. Note-taking and passage style. *Journal of Educational Psychology*, 1981, 73, 242-250.
- Britton, B.K., S.M. Glynn, B.J.F. Meyer en M.J. Penland. Effects of text structure on use of cognitive capacity during reading. *Journal of Educational Psychology*, 1982, 74, 51-61.
- Britton, B.K., B.J.F. Meyer, M.H. Hodge en S.M. Glynn. Effects of the organization of text on memory: tests of retrieval and response criterion. *Journal of Experimental Psychology: Human Learning and Memory*, 1980, 6, 620-629.
- Bromage, B.K. en R.E. Mayer. Relationship between what is remembered and creative problem-solving performance in science learning. *Journal of Educational Psychology*, 1981, 73, 451-461.

- Collins, A. Processes in acquiring knowledge. In R.C. Anderson, R.J. Spiro en W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Dooling, D.J. en R. Lachman. Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, 1971, 88, 216-222.
- Dooling, D.J. en R.L. Mullet. Locus of thematic effects on retention of prose. *Journal of Experimental Psychology*, 1973, 97, 404-406.
- Frederiksen, C.H. Effects of task-induced cognitive operations on comprehension and memory processes. In R. Freedle en J.B. Carroll (Eds.). *Language comprehension and the acquisition of knowledge*. Washington, D.C.: Winston, 1972.
- Frederiksen, C.H. Effects of context-induced processing operations on semantic information acquired from discourse. *Cognitive Psychology*, 1975, 7, 139-166.
- Frijda, N.H. Kennisverwerving. In G.J. Mellenbergh, R.F. van Naerssen en H. Wesdorp (Red.). *Rede als richtsnoer*. 's-Gravenhage: Mouton, 1979.
- Gagné, E.D. Long-term retention of information following learning from prose. *Review of Educational Research*, 1978, 48, 629-665.
- Gagné, E.D. en E.Z. Rothkopf. Text organization and learning goals. *Journal of Educational Psychology*, 1975, 67, 445-450.
- Gardner, E.T. en G.M. Schumacher. Effects of contextual organization on prose retention. *Journal of Educational Psychology*, 1977, 69, 146-151.
- Glover, J.A., B.S. Blake, B. Roberts, J.W. Zimmer en M. Palmere. Distinctiveness of encoding: the effects of paraphrasing and drawing inferences on memory from prose. *Journal of Educational Psychology*, 1981, 73, 736-744.
- Glynn, S.M. en F.J. DiVesta. Control of prose processing via instructional and typographical cues. *Journal of Educational Psychology*, 1979, 71, 595-603.
- Grimes, J. *The thread of discourse*. 's-Gravenhage: Mouton, 1975.
- Jenkins, J.R. en S.L. Deno. Influence of knowledge and type of objectives on subject-matter learning. *Journal of Educational Psychology*, 1971, 63, 67-70.
- Johnson, R.E. Recall of prose as a function of the structural importance of the linguistic units. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 12-20.
- Katona, G. *Organizing and memorizing*. New York: Columbia University Press, 1940.
- Kintsch, W. en T.A. van Dijk. Toward a model of text comprehension and production. *Psychological Review*, 1978, 85, 363-394.
- Kintsch, W., E. Kozminsky, W.J. Streby, G. McKoon en J.M. Keenan. Comprehension and recall of text as a function of content variables. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 196-214.
- Lodewijks, J.G.L.C. *Leerstofsequenties*. Academisch Proefschrift. Tilburg, 1981.
- Matthews, S.R. The impact of prior knowledge on accessibility and availability of information from prose. In A. Flammer en W. Kintsch (Eds.). *Discourse processing*. Amsterdam: North-Holland, 1982, in druk.
- Mayer, R.E. Information processing variables in learning to solve problems. *Review of Educational Research*, 1975, 45, 525-541.
- Mayer, R.E. Elaboration techniques that increase the meaningfulness of technical text: an experimental test of the learning strategy hypothesis. *Journal of Educational Psychology*, 1980, 72, 770-784.
- Mayer, R.E. Instructional variables in text processing. In A. Flammer en W. Kintsch (Eds.). *Discourse processing*. Amsterdam: North-Holland, 1982, in druk.
- Mayer, R.E. en B.K. Bromage. Different recall protocols for technical texts due to advance organizers. *Journal of Educational Psychology*, 1980, 72, 209-225.
- Melton, R.F. Resolution of conflicting claims concerning the effect of behavioral objectives on student learning. *Review of Educational Research*, 1978, 48, 291-302.
- Meyer, B.J.F. *The organization of prose and its effect on memory*. Amsterdam: North-Holland, 1975.

- Meyer, B.J.F. The structure of prose: effects on learning and memory and implications for educational practice. In R.C. Anderson, R.J. Spiro en W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Meyer, B.J.F. en G.W. McConkie. What is recalled after hearing a passage? *Journal of Educational Psychology*, 1973, 65, 109-117.
- Morris, C.D., B.S. Stein en J.D. Bransford. Prerequisites for the utilization of knowledge in the recall of prose passages. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, 5, 253-261.
- Neame, R.L.B. How to: construct a problem-based course. *Medical Teacher*, 1981, 3, 94-99.
- Nungester, R.J. en P.C. Duchastel. Testing versus review: effects on retention. *Journal of Educational Psychology*, 1982, 74, 18-22.
- O'Neil, H.F. *Learning strategies*. New York: Academic Press, 1978.
- Oostendorp, H. van en C. Hamaker. De invloed van explicitering van tekststructuur op het onthouden. *Tijdschrift voor Onderwijsresearch*, 1978, 3, 113-123.
- Peeck, J. Voorkennis en tekstbestudering. *Tijdschrift voor Taalbeheersing*, 1979, 1, 301-313.
- Peeck, J., A.B. van den Bosch en W.J. Kreupeling. The effect of mobilizing prior knowledge on learning from text. *Journal of Educational Psychology*, 1982, in druk.
- Pichert, J.W. en R.C. Anderson. Taking different perspectives on a story. *Journal of Educational Psychology*, 1977, 69, 305-315.
- 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.
- Reynolds, R.E., S.N. Standiford en R.C. Anderson. Distribution of reading time when questions are asked about a restricted category of text information. *Journal of Educational Psychology*, 1979, 71, 183-190.
- Rickards, J.P. en P.R. Denner. Depressive effects of underlining and adjunct questions on children's recall of text. *Instructional Science*, 1979, 8, 81-90.
- Romberg, T.A. en J.W. Wilson. The effect of advance organizer, cognitive set and post organizer on the learning and retention of written materials. *Journal of Research in Science Teaching*, 1973, 14, 51-55.
- Rothkopf, E.Z. The concept of mathemagenic activities. *Review of Educational Research*, 1970, 40, 325-336.
- Rothkopf, E.Z. en M.J. Billington. Goal-guided learning from text: inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology*, 1979, 71, 310-327.
- Royer, J.M. en G.W. Cable. Illustrations, analogies, and facilitative transfer in prose learning. *Journal of Educational Psychology*, 1976, 68, 205-209.
- Rumelhart, D.E. en E. Ortony. The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro en W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Schank, R.C. en R.P. Abelson. *Scripts, plans, goals and understanding*. Hillsdale: Lawrence Erlbaum, 1977.
- Schmidt, H.G. *Leren met problemen, een inleiding in probleemgestuurd onderwijs*. In A.G. Vroon (Red.). *Handboek voor de onderwijspraktijk*. Deventer: Van Loghum Slaterus, 1979.
- Schmidt, H.G. (Red.). *Probleemgestuurd onderwijs*. Stichting voor Onderzoek van het Onderwijs. Harlingen: Flevodruk, 1982a.
- Schmidt, H.G. Problem-based learning: rationale and description. *Medical Education*, 1982b, in druk.
- Schmidt, H.G. Activation and restructuring of prior knowledge and their effects on test processing. In A. Flammar en W. Kintsch (Eds.). *Discourse processing*. Amsterdam: North-Holland, 1982c, in druk.
- Schmidt, H.G. en P.A.J. Bouhuijs. *Onderwijs in taakgerichte groepen*. Utrecht: Het Spectrum, 1980.

- Schultz, J.V. en L. Hubert. A nonparametric test for the correspondence between two proximity matrices. *Journal of Educational Statistics*, 1976, 1, 59-67.
- Simons, P.R.J. *Vergelijkenderwijs: onderzoek naar de invloed van metaforen op het leren*. Academisch Proefschrift. Tilburg, 1981.
- Spilich, G.J., G.T. Vesonder, H.L. Chiesi en J.F. Voss. Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, 1979, 18, 275-291.
- Tulving, E. en D.M. Thomson. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.
- Wertheimer, M. *Productive thinking*. New York: Harper & Row, 1959.
- Winne, P.H. Minimizing the black box problem to enhance the validity of theories about instructional effects. *Instructional Science*, 1982, 11, 13-28.
- Wittrock, M.C., C.B. Marks en M. Doctorow. Reading as a generative process. *Journal of Educational Psychology*, 1975, 67, 484-489.
- Wouters, L. en P. Kop. Konkretiserende elaboraties en tekstbestudering. *Tijdschrift voor Onderwijsresearch*, 1981, 6, 113-128.

**ACTIVATION AND RESTRUCTURING OF PRIOR KNOWLEDGE
AND THEIR EFFECTS ON TEXT PROCESSING:
RECOGNITION AND TRANSFER**

H.G. Schmidt

In two experiments effects of a group problem-solving procedure were assessed.

In Experiment 1 small groups of subjects were presented with a problem they had to discuss. Compared with a control condition the experimental subjects showed superior recognition and transfer of information relevant to the problem. These effects were attributed to the activating and restructuring properties of the problem-solving procedure.

In Experiment 2 effects of activation and restructuring of prior knowledge on subsequent text processing were examined. A general facilitative effect was found. This result cannot be explained in terms of selective attention induced by prior problem analysis.

Introduction

Learning by discovery has been considered a useful addition and perhaps even an alternative to existing educational procedures for some time. Jerome Bruner for instance, the first who applied this method on a somewhat broader scale, was of the opinion that discovery learning more than conventional education increases students' intellectual capacities. The method was said to promote the learning of meaningful information, which in turn would have positive effects on long-term retention and transfer. It was also thought to stimulate a student's intrinsic motivation (Bruner, 1961). Bruner's ideas have given impetus to a lot of empirical research, but the yield has not been very significant (Shulman and Keislar, 1966). Learning by discovery does not seem capable of producing the predicted *extra* learning results.

One may wonder why this should be so. After all, the learners in the situations described by Bruner are cognitively active to a high degree and this condition is generally recognized as favoring the learning process. According to Mayer (1975) the disappointing outcome is probably accounted for

by the fact that discovery learning only leads to activation of *existing* knowledge. If no confrontation with new knowledge takes place, learning does not occur (Mayer, 1975, p. 539). Mayer derives this hypothesis from his own assimilation theory, which states that new information is absorbed only if three conditions are fulfilled:

1. a certain amount of prior knowledge about the subject to be studied should be present. This is referred to by Mayer as the *assimilative set*: the body of knowledge into which the information newly to be acquired should be assimilated;
2. the assimilative set should be actually *activated* by the education;
3. during the learning process, knowledge should *interact with new information* to allow development of new knowledge structures (Mayer, 1975; 1979a; 1979b).

As a rule, the third condition is not satisfied by discovery learning. While working on a problem, the student is expected to produce the required additional information *himself*. It does not seem reasonable however to assume that he is able to do so without external assistance.

But what would happen if we let students work on a problem first (as is usual for discovery learning) and *subsequently* confront them with *new* information relevant to the problem? By modifying the discovery learning approach in this direction all three conditions for learning set by Mayer would be fulfilled in principle. In fact, experience with such a modification is being obtained for a few years, notably in medical education (Neufeld and Barrows, 1974; Schmidt and Bouhuijs, 1980; Neame, 1981). A small group of students under the guidance of a tutor is offered a problem description. The problem usually includes a number of phenomena or events which can be observed in reality. Students are asked to *explain* these phenomena in terms of underlying processes, principles or mechanisms (Schmidt, 1979). They do so by utilizing prior knowledge to formulate hypotheses regarding the process or principle that might underlie the phenomena outlined. They are trying, as it were, to give a tentative description of this process. Subsequently, they collect or receive new information relevant to the problem, by which any ambiguities uncovered during the initial analysis of the problem can be clarified. This variant to learning by discovery is called problem-based learning (Barrows and Tamblyn, 1980).

To be able to make predictions about possible effects of problem analysis on the acquisition of additional information, we refer to schema theory, a much used description of human information processing (Bartlett, 1932; Ausubel, 1968; Minsky, 1975; Rumelhart and Ortony, 1977). This theory states that in the encoding of new information existing knowledge schemata are activated that regulate the process of comprehension (Dooling and Lachman, 1971; Bransford and Johnson, 1972; Anderson and Pichert, 1978). A decisive role in understanding new information is probably played by *inferences* generated by an activated schema (Schank and Abelson, 1977). Inferences may be regarded as hypotheses about information still

to be stored that are tested against such information. The outcome of these tests decides on a possible change of existing schemata, which are then said to *accommodate* to the new information (Anderson, 1977). From this perspective learning can be considered a process of differentiation and reconstruction of existing cognitive structures.

Now, if we try with the aid of these notions to describe the cognitive processes taking place in students engaged in problem analysis, the following suppositions might be made. Thinking about the presented problem and discussing it with others activates existing schemata more or less relevant to that problem. These schemata may derive from factual knowledge of the underlying principle or process, from knowledge of analogous processes, or from general world knowledge. The schemata will produce inferences with the aid of which students will try to develop their own cognitive representation of the processes that may be considered responsible for the phenomena contained in the problem description. If the problem cannot be satisfactorily solved with the help of knowledge that students already have available prior to the discussion, the cognitive representation of the processes underlying the problem will take the form of a *new*, more differentiated construction of the prior knowledge of each contributing individual. The tutor stimulates this process of activation and restructuring of prior knowledge by inducing students to explicate their inferences. He will do so by means of Socratic questioning (Collins, 1977).

In summary we may say that analysis of a problem leads to activation and 'recontextualization' of prior knowledge (Anderson, 1977) and, as a result of the production of elaborations, to its restructuring. The availability of more differentiated schemata resulting from problem analysis should in turn facilitate the processing of new information relevant to the problem.

In two experiments we have investigated a number of the hypotheses expressed here. The first was conducted to find out to what extent problem analysis leads to activation and restructuring of people's prior knowledge with respect to the problem and its underlying process. In the second experiment we examined the effect of problem analysis on the processing of new information. In each experiment we investigated effects on two dimensions customary in this kind of research, the degree to which subjects are capable of *recognizing* problem-relevant knowledge and the degree to which they are able to use this knowledge in new situations: the degree of *transfer* of knowledge.

Experiment 1

In Experiment 1 the effect of problem analysis on existing cognitive structures was investigated. The question to be answered was whether analysis of a problem according to the method described in the introduction leads to activation and restructuring of schemata.

Method

Subjects

39 students (8 males, 31 females) of an institute for higher education participated in the experiment. All subjects had previously attended the same type of secondary education (in Dutch: HAVO), with their final examination including biology. They were paid for participation.

Materials

The materials included a problem description, two multiple-choice tests, and a questionnaire. The *problem* was described as follows:

A red blood cell (a red blood corpuscle) is transferred to pure water under a microscope. The blood cell swells rapidly and eventually bursts. Another red blood cell is added to an aqueous salt solution and is observed to shrink.

How can these phenomena be explained?

Subjects were to account for the described phenomena in terms of an underlying process, mechanism, or principle. The problem refers to osmosis. This has been selected because inspection of four biology text books much in use at HAVO schools had shown that this subject was treated in a more or less identical fashion. Homogeneity of the subjects with respect to prior knowledge of biology of course reduces chance variation in the results of the experiment.

Dependent variables consisted of two tests: a recognition and a transfer test. In the construction of the test items use was made of a text about osmosis and related subjects such as diffusion, turgor, and plasmolysis. This text was specifically written for Experiment 2. (Further information is given in the description of Experiment 2.)

The *recognition test* was made up of 33 items for a major part written in agreement with recommendations made by Anderson (1972) and Bormuth (1970). This means that the items resulted from paraphrase and transformation of various passages of the text about osmosis. Except for two, all items were of the true/false type, the two exceptions containing three options for answering.

The *transfer test* consisted of 37 items that were supposed to measure the extent to which subjects were able to apply knowledge about osmosis. Most items had the form of small exercises which could be carried out with the aid of information supplied by the text. However, inferences from the text, i.e., assumptions not included in the text but ensuing therefrom, also formed part of the transfer test. The transfer test consisted of 16 true/false items and 21 items with three answering alternatives. To all test items a question mark category was added, so that subjects could indicate that they did not know the correct answer.

The *questionnaire* concerned such biographic data as age, sex, and final examination mark for biology.

Procedure

Subjects were randomly assigned to one of two conditions. The experimental group consisted of 20 subjects, the control group of 19. Following this randomizing procedure we examined to what degree the groups were comparable in prior knowledge of biology (expressed by the final examination grade). The mean for biology was equal to 6.30 in the experimental group (standard deviation .73) and equal to 6.47 in the control group (standard deviation .69). These small differences did not necessitate subsequent matching.

The experimental group was then randomly divided into three smaller groups ($N = 7, 7$ and 6). The analysis of the blood cell problem took place in these groups. A female tutor experienced in the use of the method and conversant with the osmosis process was assigned to each of these groups. Tutor and subjects had never met before. The tutor proceeded as follows: she explained the various steps of problem analysis by means of an example and had written the following text on a blackboard: Problem analysis: 1. Reflection. 2. Problem definition. 3. Production of explanations. 4. Elaboration. Subjects engaged actively with the explanation of the method. Before the blood cell problem was presented, the tutor mentioned that problem analysis was not to take more than 15 minutes. The problem description was issued and the group given 1 minute for reading and reflection. Next, the discussion was started, with the tutor acting as discussion leader. She had been explicitly instructed not to reveal any information about the problem and to take care that her summaries also did not implicitly contain information from which subjects might derive insights into the nature of the problem at hand. The discussion was recorded on tape for subsequent verification of the tutor's and her group's compliance with the approach agreed upon. It appeared that all tutors had accurately followed their instructions. The groups needed less than 10 minutes for analysis. On completion of the problem analysis both the experimental group and the control group (which had not been presented with a problem) filled out the questionnaire and answered the multiple-choice tests. Each group was told that the tests were not examinations and that nothing was to be gained by good or poor results. Participants were urged not to guess at the correct answer but in case of doubt to indicate the question mark. The answering of the items was anonymous. Recognition and transfer scores were calculated by counting the number of correct answers.

Results and discussion

Items that proved to be too easy were removed from the analysis. As a criterion a p -value $\geq .90$ was used. It proved necessary to remove 10 items from the recognition test and 5 from the transfer test. The alpha reliability coefficient for the resulting 23-item recognition test was equal to .73 and for the 32-item transfer test to .69. These reliabilities compare with those generally obtained in examinations (Wijnen, 1971). The two instruments are considered different on a priori grounds. They were each sup-

posed to measure a different aspect of the subjects' knowledge. Their product-moment correlation was equal to .41. Consequently they have about 16% of variance in common. This means that they each measure the same characteristic to a small degree and a different characteristic to a larger degree. The results of the recognition test are summarized in Table 1 and those for the transfer test are summarized in Table 2. The results on both tests were subjected to a one-way analysis of variance.

Table 1: Means and standard deviations of recognition scores for Experiment 1

	M	SD	N
Experimental group	10.90	4.17	20
Control group	8.63	3.00	19
Total	9.79	3.65	39

Table 2: Means and standard deviations of transfer scores for Experiment 1

	M	SD	N
Experimental group	16.55	5.08	20
Control group	13.26	4.40	19
Total	14.95	4.84	39

The analyses of variance show that the actual difference found between the experimental and control conditions are statistically significant at the 10% level. This applies to both recognition, $F(1,37) = 3.77$, $p < .06$, and transfer, $F(1,37) = 4.96$, $p < .04$. These results support the theory advanced in the introduction with respect to cognitive effects of problem analysis. It seems that written problems indeed activate existing schemata and that the problem-connected inferences based on these schemata lead to their restructuring. Apparently, the restructured schemata enable subjects in the experimental condition to give better answers to the items of the tests, as compared with the control subjects.

The chosen experimental design does not allow us to separate the effects of activation from those of restructuring of schemata. For in the control group activation of prior knowledge also takes place, notably at the moment subjects in that condition answer the tests. Of course, reading the test items activates prior knowledge as well. Whether the combination of problem and test items activates prior knowledge to the same degree as answering the test items only — and whether the established effects should solely be ascribed to restructuring — remains unclear for the time being.

Experiment 2

What happens if you ask students who have just tackled the blood-cell problem to study a text about osmosis? What will be the effect of recontextualized and restructured prior knowledge on the processing of text relevant to the problem?

According to the position defended here, problem analysis functions as a bridge between prior knowledge and knowledge still to be acquired. Elaboration of prior knowledge by means of discussion of a problem causes existing cognitive structures to change in the direction of further differentiation and restructuring. Compared with the situation that exists before problem analysis the subjects' schemata will contain concepts that are clearer and have more and closer mutual relationships. In this way it provides a better ideational scaffolding for new information contained in a problem-relevant text (Anderson et al., 1978). Information from the text will be processed more easily and, consequently, faster. In terms of the second experiment this means that given a standard amount of study time, persons who have first worked on a problem will process a problem-relevant text deeper than others who have only processed the text (Craik and Lockhart, 1972).

If this is the sole effect to be expected from problem analysis, one might just simply extend study time for subjects' text processing (Peeck, 1970) instead of presenting them with a problem-plus-text. However, subjects who have analyzed a problem beforehand will not only process new information faster but also *in a different way*. As a result of the structural changes to which existing schemata have been subjected new information will be comprehended in another, more differentiated manner. This subtler interpretation of the subject-matter will lead to a deeper understanding and consequently to improved application of knowledge in transfer tasks. In other words: schema theory predicts better retention as well as greater transfer. Mayer (1974, 1975) slightly modifies this prediction. He states that different educational procedures often do not lead to an *increase* (or decrease) in learning, but to learning results that are *structurally different*. For instance, education may emphasize the internal structure of information to be acquired, or it may stress relationships with other, already available, knowledge. In the former case, according to Mayer, a cognitive structure develops whose constituent concepts are strongly linked with each other but have few relationships with knowledge elements beyond the domain studied. Mayer refers to this as the 'internal connectedness' of that cognitive structure. In the latter case a cognitive structure develops which has strong connections with knowledge elements outside the specific domain ('external connectedness'). He predicts that educational procedures stressing the internal connections between new concepts lead to improved recognition of such concepts, and that educational procedures emphasizing relationships between new concepts and other already available knowledge lead to better transfer but poorer recognition. Research

conducted by Mayer and Greeno (1972) demonstrated that these 'treatment-post-test' interactions do occur. They gave instruction to subjects on binomial probability by means of two procedures. In one of these the relations between the variables of the binomial formula were emphasized and subjects learned to calculate by using the formula. In the other procedure they tried to establish connections between the variables of the formula on the one hand, and the prior knowledge and experience of subjects with chance, trial, and outcome on the other. The results did not yield an overall difference between procedures on a post test, but did show qualitative differences in the answering of the items. As predicted, the 'internal connectedness' group proved superior in answering items about the relations between variables of the formula, which means that they were better in recognition tasks. The 'external connectedness' group proved more apt in recognizing unsolvable problems and in solving problems whose variables were concealed in a story. This group, therefore, excelled in items involving transfer. It looks as if emphasis on external relationships of concepts to be learned (the relationship between concepts and learners' prior knowledge) is at the expense of the internal connectedness of concepts to be learned (the mutual relationship between the concepts). Assuming that the recognition test applied in the present experiment aims exclusively at concepts from the text proper and at their mutual relationships, one would have to predict on the basis of Mayer's considerations that the control group (text-only) will perform better on the recognition test, whereas the experimental group (problem-plus-text: emphasizing relationships with prior knowledge) will demonstrate a better performance on the transfer test.

The predictions expressed above have all been derived from schema-theoretical notions with respect to how a text is processed. However, there is also another theory about text processing that may be relevant within the scope of this research. This theory is *selective attention theory* (Rothkopf, 1970; Duchastel, 1979; Glynn and Di Vesta, 1979; Reynolds et al., 1979). This theory simply postulates that whatever people learn from a text is a function of the amount of attention paid to various text passages. The amount of attention paid to various passages is in turn a function of the *objectives* of the persons processing the text. Typical for this approach is the following experiment: Two groups of subjects study the same text. One group has been previously given a list of learning objectives or questions relating to the text. Measurement of what has been processed then shows that subjects who have studied the text guided by learning objectives recall goal-relevant information better than the control group. This is achieved however at *the expense* of the storage of information not directly related to the learning objectives.

In an investigation into eye-movements during reading, Rothkopf and Billington (1979) discovered that subjects fixated twice as much on sentences relevant to a learning objective than on sentences which are not. From this they drew the conclusion that what people learn is determined

by the amount of selective attention — expressed in amount of inspection time — paid to the text passages to be learned.

The materials used in our experiment were constructed in such a manner that a selective-attention explanation of possible effects could also be tested. We started from the assumption that the blood cell problem (which was worked on by the subjects of our experiment) might play the same role as the learning objectives in the selective-attention experiments. This means that working on problems would induce subjects to pay more attention to problem-relevant passages of the text at the expense of those passages that are less meaningful in the light of the problem. The text about osmosis was edited to include information of lesser relevance to the blood-cell problem. The selective attention theory predicts that subjects of the experimental group (problem-plus-text) will pay greater attention to problem-relevant text passages compared with the control group (text-only), but also that such attention will reduce the attention paid to less relevant parts of the text. Given an equal amount of study time for both groups, the experimental group will perform better on items concerned with topics relevant to the problem, whereas the control group will do better with non-relevant items. Hence, no overall difference between the two groups is going to be observed in either test. The differential predictions discussed above are summarized in Table 3. The plus-sign means the problem analysis group will perform better on this test (or part of the test) than the control group. A minus-sign means that the control-group performance will be better.

Table 3: Differential predictions about the influence of problem analysis on text processing

		recognition	transfer
Schema theory (Rumelhart and Ortony, 1977)		+	+
Mayer and Greeno's (1972) specification		—	+
Selective attention theory (Rothkopf and Billington, 1979)	items concerned with text relevant to problem	+	+
	items concerned with text not relevant to problem	—	—

Method

Subjects

48 students (42 females, 6 males) of an institute for higher education participated in the experiment. All subjects had attended the same type of secondary education (in Dutch: HAVO), with their final examinations including biology. They were paid for participating.

Materials

The materials used in Experiment 2 were identical to those of Experiment 1. In addition, however, a written text was used.

The 2,220-word *text* was entitled 'Osmosis and Diffusion'. It had been designed to adapt as well as possible to subjects' prior knowledge. The description of the osmosis process was more detailed and precise than is customary for HAVO schools. The text dealt with the following topics: diffusion, diffusion rate, (semi-)permeability of cell membrane, osmosis, osmotic pressure, osmotic value, plant cell structure, turgor, and plasmolysis. Not all of these were directly relevant to the problem. In order to be able to test selective attention theory the two multiple-choice tests were subdivided by two independent judges, including the investigator, into items of immediate importance to the problem and those of lesser relevance. Agreement between judges was 76%. The items about which consensus had not been reached were classified in one of the two categories by mutual agreement. (Removal from the analysis would have left too few items.)

Procedure

The procedure was largely identical to that of Experiment 1. In the experimental condition 23 subjects were tested and in the control condition 25. Following randomization it was established to what extent the two groups were comparable in prior knowledge of biology (expressed by the final examination grade). The average grade for biology in the control group was 6.36 (standard deviation .81) and 6.39 in the experimental group (standard deviation .94). The experimental group was subdivided into three smaller groups ($N = 8, 8$ and 7). Each of them was randomly assigned an experienced female tutor. The tutor and her group proceeded in the way described for the first experiment. All groups needed less than 10 minutes for analysis. After that the experimental group answered the questionnaire. It was next given the text and instructed to study this. The control group answered the questionnaire and studied the text. Each group was allowed 15 minutes for study.

Finally, subjects took the recognition and the transfer tests.

Results and discussion

Items having a p -value equal to or larger than .90 were removed from the analyses as being too easy.

The results for recognition are given in Table 4 and those for transfer in Table 5.

Table 4: Means and standard deviations of recognition scores for Experiment 2

	M	SD	N
Experimental group	16.17	2.46	23
Control group	14.48	2.63	25
Total	15.29	2.55	48

Table 5: Means and standard deviations of transfer scores for Experiment 2

	M	SD	N
Experimental group	20.04	3.01	23
Control group	17.56	3.45	25
Total	18.75	3.25	48

One-way analyses of variance on these results yield the following picture. Recognition: $F(1,46) = 5.26$, $p < .03$. Transfer: $F(1,46) = 7.01$, $p < .02$. These results support the assertion expressed before that analysis of a problem in the way this is prescribed in problem-based learning causes new information to be better understood and remembered. Recontextualization of the assimilative set by means of problem analysis promotes interaction between prior knowledge and new information, which in turn is responsible for the observed gains in the recognition of such information and its application.

The results do not however support the view of Mayer and Greeno (1972) that intensive interaction between prior knowledge and new information should lead to poorer performance on a test measuring knowledge of the 'autonomous' text (Anderson, 1977), whereas transfer should be promoted. The experimental subjects were also better in *recognizing* the text. There is a possibility that the recognition test used in this experiment does not really measure Mayer and Greeno's 'internal connectedness' of the cognitive structure resulting from text processing. Mayer and Greeno do not indicate clearly what exactly they have in mind when using this term, so that our interpretation of the concept — that internal connectedness can be measured by text recognition — was perhaps not quite correct.

However this may be, schema theory appears to allow correct predictions with regard to the influence activation and restructuring of existing schemata have on the processing of new information.

What about the support for selective attention theory? The answer to this question has not been made superfluous by the previous argumentation because, theoretically, it is possible that problem analysis has a schema-restructuring as *well as* a guiding function. In order to investigate this hypothesis we divided as mentioned the items of the tests into a set directly

relevant to the osmotic process and a set of indirect relevance. Items relating to diffusion rate, osmotic value, plant cell structure, turgor, and plasmolysis were regarded as indirectly relevant to the problem. Each of the four subtests resulting from this subdivision has been subjected to an analysis of variance. In addition *all* items of direct relevance from the tests have been combined into a new test and subjected to an analysis of variance. The same was done with *all* indirectly relevant items. The results are given in Tables 6 and 7. These only show mean values and chance probabilities resulting from the analyses.

Table 6: Means and chance probabilities of subtests consisting of items directly relevant to the problem

	recognition	transfer	total
Experimental group	7.43	11.13	18.57
Control group	6.80	10.16	16.96
Chance probability	< .14	< .13	< .08

Table 7: Means and chance probabilities of subtests consisting of indirectly relevant items

	recognition	transfer	total
Experimental group	8.74	8.91	17.65
Control group	7.68	7.40	15.08
Chance probability	< .04	< .02	< .01

In the light of selective attention theory these results can hardly be interpreted. Selective attention theory predicted effects favoring the treatment group with respect to subtests consisting of items directly relevant to the problem, whereas the control group would demonstrate better performance on indirectly relevant items. This was not the case. Statistically significant differences were not found in the case of problem-relevant items, whereas the differences established for indirectly relevant items did favor the treatment group.

If we should nevertheless want to maintain (a combination of schema theory and) selective attention theory in explaining these results, only the following line of thought seems plausible: the way in which the problem analysis group prepares for text study induces this group to pay less attention to problem-relevant information than the control group. Which means that this group has more time available for studying information which is relatively new in the light of the problem. As a result experimental subjects

do better on indirectly relevant items and, while study time and type of preparation counterbalance each other on directly relevant information, they do not show poorer performance on directly relevant items. This explanation should be called a boredom-hypothesis, because it suggests that subjects in the treatment condition become bored earlier with the osmosis topic than the control subjects.

General Discussion

The two experiments discussed here show that analysis of a problem in the way described in this contribution and elsewhere (Schmidt, 1979; Schmidt and Bouhuijs, 1980) causes existing knowledge schemata to be activated and restructured, and that this effect facilitates subsequent processing of a text relevant to that problem. With reference to the introduction of this contribution we may say that Mayer's (1975) diagnosis regarding the effects of discovery learning appears to be correct. Learning by discovery indeed leads to activation of the assimilative set and, as we have demonstrated, to its restructuring. This effect, however, is insufficient to produce the extra learning effects predicted by Bruner, for that calls for actual confrontation with new information.

Of course, the results of the present experiments should to a certain degree be regarded as isolated phenomena. Naturally, it is too early to conclude that elaboration of prior knowledge by means of problem analysis has, in general, a facilitating effect on text processing. We have shown, however, that conditions can be created under which problem analysis effects are demonstrable. Finding out how representative these conditions are and how general the effects certainly requires further research.

Moreover, there are some other possible explanations for the results found, interpretations that could not be ruled out by the chosen experimental design. The first is that the established results could be generated by a *Hawthorne effect*. It is possible that the subjects experienced problem analysis as being so new and interesting that it increased their motivation. This alternative explanation might well hold true for Experiment 1. Not so for the second however. The subjects participating in the latter had, some time beforehand, gained an extensive three-months experience with this type of instruction, so the method was not really new to them. A second alternative derives from motivational descriptions of the learning process (Faw and Waller, 1976; Mayer, 1980). It is not impossible that working on an interesting problem has such a *motivating effect* on subjects that those who have carried out a problem analysis are more interested in a text relevant to that problem than people who have not. There are indications that problem analysis indeed induces a slightly greater motivation. The latter would in turn increase a subject's efforts, so that he/she learns more in less time. However, research on motivation has only rarely established such direct effects on learning. It has been found, though, that motivation influ-

ences the *amount of time* students are prepared to spend on material to be studied. We did keep study time constant in our research, so this alternative is not a very likely one.

For that matter, it would be interesting to examine the influence problem analysis has on study time. Would subjects be inclined to spend more time on studying material relevant to a certain problem? Another question concerns the influence of problem analysis on subsequent retrieval of knowledge stored in long-term memory. One could imagine that storage of information, along with the cues provided by problem analysis, facilitates subsequent retrieval (Tulving and Thomson, 1973).

Finally, interactions between subjects' prior knowledge and the structure and difficulty of the text (Mayer, 1980; Kintsch and Van Dijk, 1978), as well as variations in the nature of the problems presented should be studied.

References

- Anderson, R.C. How to construct achievement tests to assess comprehension. *Review of Educational Research*, 1972, 49, 145-170.
- Anderson, R.C. The notion of schemata and the educational enterprise: general discussion of the conference. In R.C. Anderson, R.J. Spiro and W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Anderson, R.C. and J.W. Pichert. Recall of previously unrecalled information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, 1978, 17, 1-12.
- Anderson, R.C., R.J. Spiro and M.C. Anderson. Schemata as scaffolding for the representation of information in connected discourse. *American Educational Research Journal*, 1978, 15, 433-440.
- Ausubel, D.P. *Educational Psychology, a cognitive view*. New York: Holt, Rinehart and Winston, 1968.
- Barrows, H.S. and R.M. Tamblyn. *Problem-based learning*. New York: Springer, 1980.
- Bartlett, F.C. *Remembering*. Cambridge: Cambridge University Press, 1932.
- Bormuth, J.R. *On the theory of achievement test items*. Chicago: University of Chicago Press, 1970.
- Bransford, J.D. and M.K. Johnson. Contextual prerequisites for understanding: some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 717-726.
- Bruner, J.S. The act of discovery. *Harvard Educational Review*, 1961, 31, 21-32.
- Collins, A. Processes in acquiring knowledge. In R.C. Anderson, R.J. Spiro and W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Craik, F.I. and R.S. Lockhart. Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- Dooling, D.J. and R. Lachman. Effects of comprehension on retention of prose. *Journal of Experimental Psychology*, 1971, 88, 216-222.
- Duchastel, P. Learning objectives and the organization of prose. *Journal of Educational Psychology*, 1971, 71, 100-106.

- Faw, H.W. and T.G. Waller. Mathemagenic behaviors and efficiency in learning from prose. *Review of Educational Research*, 1976, 46, 691-722.
- Glynn, S.M. and F.J. Di Vesta. Control of prose processing via instructional and typographical cues. *Journal of Educational Psychology*, 1979, 71, 595-603.
- Kintsch, W. and T.A. van Dijk. Toward a model of text comprehension and production. *Psychological Review*, 1978, 85, 363-394.
- Mayer, R.E. Acquisition processes and resilience under varying testing conditions of structurally different problem solving procedures. *Journal of Educational Psychology*, 1974, 66, 644-656.
- Mayer, R.E. Information processing variables in learning to solve problems. *Review of Educational Research*, 1975, 45, 525-541.
- Mayer, R.E. Can advance organizers influence meaningful learning? *Review of Educational Research*, 1979a, 49, 371-383.
- Mayer, R.E. Twenty years of research on advance organizers: assimilation theory is still the best predictor of results. *Instructional Science*, 1979b, 8, 133-167.
- Mayer, R.E. Elaboration techniques that increase the meaningfulness of technical text: an experimental test of the learning strategy hypothesis. *Journal of Educational Psychology*, 1980, 72, 770-784.
- Mayer, R.E. and J.G. Greeno. Structural differences between learning outcomes produced by different instructional methods. *Journal of Educational Psychology*, 1972, 63, 165-173.
- Minsky, M. A framework for representing knowledge. In P.H. Winston (Ed.). *The psychology of computer vision*. New York: McGraw-Hill, 1975.
- Neame, R.L.B. How to construct a problem-based course. *Medical Teacher*, 1981, 3, 94-99.
- Neufeld, V.R. and H.S. Barrows. The McMaster Philosophy: an approach to medical education. *Journal of Medical Education*, 1974, 49, 1040-1050.
- Peeck, J. Effects of prequestions on delayed retention of prose material. *Journal of Educational Psychology*, 1970, 61, 241-246.
- Reynolds, R.E., S.N. Standiford and R.C. Anderson. Distribution of reading time when questions are asked about a restricted category of text information. *Journal of Educational Psychology*, 1979, 71, 183-190.
- Rothkopf, E.Z. The concept of mathemagenic activities. *Review of Educational Research*, 1970, 40, 325-336.
- Rothkopf, E.Z. and M.J. Billington. Goal-guided learning from text: inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology*, 1979, 71, 310-327.
- Rumelhart, D.E. and E. Ortony. The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro and W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Schank, R.C. and R.P. Abelson. *Scripts, plans, goals and understanding*. Hillsdale: Lawrence Erlbaum, 1977.
- Schmidt, H.G. Leren met problemen, een inleiding in probleemgestuurd onderwijs. In A.G. Vroon (Ed.). *Handboek voor de onderwijspraktijk*. Deventer: Van Loghum Slaterus, 1979.
- Schmidt, H.H. and P.A.J. Bouhuijs. *Onderwijs in taakgerichte groepen*. Utrecht: Het Spectrum, 1980.
- Shulman, L.S. and E.R. Keislar (Eds.). *Learning by discovery, a critical appraisal*. Chicago: Rand McNally, 1966.
- Tulving, E. and D.M. Thomson. Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 1973, 80, 352-373.
- Wijnen, W.H.F.W. *Onder of boven de maat*. Amsterdam: Swetz en Zeitlinger, 1971.

INTRINSIEKE MOTIVATIE EN STUDIEPRESTATIE: ENKELE VERKENNENDE ONDERZOEKINGEN

H.G. Schmidt

In het in dit artikel beschreven onderzoek werden twee doeleinden nagestreefd. Nagegaan werd òf en in hoeverre een methode van probleemanalyse (Barrows en Tamblyn, 1980; Schmidt, 1982a) invloed uitoefent op de intrinsieke motivatie van degenen die die probleemanalyse uitvoeren. Daarnaast werd onderzocht op welke wijze intrinsieke motivatie met betrekking tot een bepaald onderwerp samenhangt met studieprestaties op dat onderwerp.

In Experiment 1 analyseerden proefpersonen een probleem. Hun taak was verklaringen te vinden voor het gedrag van een rode bloedcel in water en in een zoutoplossing. Na afloop bleken zij meer geïnteresseerd te zijn in nadere informatie met betrekking tot het proces dat aan dat probleem ten grondslag ligt dan een controlegroep. Een verklaring voor dit verschijnsel werd gezocht in Berlyne's (1971) incongruentietheorie van intrinsieke motivatie.

In Experiment 2 kregen proefpersonen die een probleem hadden geanalyseerd vervolgens een probleemrelevante tekst voorgelegd. Voorafgaande aan de bestudering van de tekst werd intrinsieke motivatie gemeten. Nadat de tekst bestudeerd was werden studieprestaties en opnieuw intrinsieke motivatie gemeten. Analyses van de correlatiepatronen suggereren dat intrinsieke motivatie en studieprestatie van elkaar onafhankelijke uitkomsten van een leerproces zijn. Daarbij werd aangetekend dat vanwege tekortkomingen in de onderzoeksopzet (studietijd werd constant gehouden) de hypothese dat intrinsieke motivatie prestatie zou beïnvloeden via de beïnvloeding van de studietijd niet getest kon worden.

Voor de ontwikkeling van de onderwijskunde is onderzoek naar processen die ten grondslag liggen aan intrinsiek gemotiveerd handelen en naar mogelijke relaties tussen die processen en studieprestaties van belang. Immers: (causale) samenhangen tussen deze beide onderwijsvariabelen zijn maar zelden onderzocht, en de schaarse literatuur is nauwelijks eenduidig over de aard van die samenhang.

Nu als gevolg van de veranderende economische omstandigheden extern motiverende factoren in het onderwijs, zoals een diploma of een goede baan, in toenemende mate aan belang inboeten; in andere woorden: nu voor veel lerenden een duidelijk toekomstperspectief begint te ontbreken als drijvende kracht achter het leveren van studieprestaties, krijgt onderzoek naar intrinsieke motivatie ook een groter *praktisch* belang voor het

onderwijs. De behoefte zal toenemen na te gaan welke factoren bevorderen dat lerenden leeractiviteiten gaan waarnemen als *in zichzelf* betekenisvolle en waardevolle bezigheden.

Daarbij kan men zich afvragen of de school niet te zeer gezien wordt als de enige omgeving waar het leren plaatsvindt. Volgens Maehr (1976) zou de school veel sterker een plaats moeten zijn waar het leren slechts begint en waar vooral de belangstelling voor datgene wat geleerd moet worden bevorderd wordt. Dat lijkt hem de enige manier om lerenden voor te bereiden op een levenslang leren, noodzakelijk om het hoofd te kunnen blijven bieden aan de toenemende complexiteit van de samenleving. Bruner (1961) noemt dat 'leren te leren'. Intrinsieke motivatie wordt gezien als een noodzakelijke voorwaarde daarvoor.

In dit artikel zal eerst een beknopt overzicht worden gegeven van het werk dat op het terrein van intrinsieke motivatie en onderwijs gedaan is. Vervolgens worden twee experimenten gerapporteerd. In Experiment 1 werd de effectiviteit geëvalueerd van een didactische methode om intrinsieke motivatie bij lerenden te initiëren. In Experiment 2 werden relaties tussen intrinsieke motivatie en studieprestaties onderzocht.

Onderzoek naar intrinsieke motivatie

Het begrip intrinsieke motivatie werd oorspronkelijk gebruikt als omschrijving van een proces dat verondersteld wordt ten grondslag te liggen aan bepaalde bijzondere vormen van exploratief gedrag. Sommige gedragingen lijken namelijk niet onder directe controle te staan van (fysiologische) drijfveren of van externe beloningen. Die gedragingen worden gesteld omdat ze in zichzelf bevredigend zijn; ze zijn intrinsiek gemotiveerd. Intrinsiek gemotiveerd gedrag is dus gedrag dat niet onder invloed staat van externe factoren. Berlyne (1955) toonde aan dat ratten bijvoorbeeld een voorkeur hebben voor het onderzoeken van nieuwe ruimten en objecten, en dat zij daarmee doorgaan zolang nieuwe stimulatie beschikbaar is. Dat gedrag wordt vertoond onafhankelijk van de aan- of afwezigheid van een beloning. Niet alleen bij dieren komt intrinsiek gemotiveerd gedrag voor. Het kan ook geobserveerd worden bij kinderen die urenlang thuis bezig kunnen zijn met een puzzel, of met datgene wat ze op school geleerd hebben, zoals lezen of sommen maken. Intrinsiek gemotiveerd gedrag is dus gedrag dat geen middel tot een (extern) doel is, maar een doel in zichzelf.

Er zijn verschillende pogingen gedaan om intrinsiek gemotiveerd gedrag te verklaren. De meest plausibele verklaringen worden gevormd door wat genoemd zou kunnen worden: de positief-affect-theorie van Deci (De Charms, 1968; Deci, 1975; Deci en Ryan, 1980) en de incongruentie-theorie van Hunt (Day, Berlyne en Hunt, 1971; Hunt, 1963; 1971).

Volgens Deci (1975) hebben mensen een fundamentele behoefte zichzelf waar te nemen als competent en zelf-bepalend ('self-determining'). De mens wil zichzelf bij voorkeur zien als de *veroorzaker* van zijn gedrag. En

hij wil het gevoel hebben dat hij met zijn gedrag zijn omgeving beheerst, dat hij effectief is in het produceren van door hem gewilde veranderingen in zijn omgeving. Om zich competent te kunnen voelen, zoekt hij naar situaties die een redelijke uitdaging vormen, optimaal in het licht van wat hij ziet als zijn bekwaamheden. Wanneer hij aan die uitdaging het hoofd blijkt te kunnen bieden, brengt dat een gevoel van bevrediging bij hem te weeg. Mensen hebben dus vanuit de behoefte competent te zijn een voorkeur voor activiteiten die deze interne beloning oproepen. Deze activiteiten heten daarom intrinsiek gemotiveerd.

Deci's theorie over intrinsieke motivatie is een affect-theorie omdat het uiteindelijk een emotie is die als drijvende kracht fungeert wanneer de intrinsiek gemotiveerde activiteit uitgevoerd wordt. Hunts incongruentietheorie kan daarentegen omschreven worden als een cognitieve theorie. Volgens Hunt (1971) bezit het organisme intrinsieke behoefte om incongruentie tussen kennis die het heeft van zijn omgeving en binnenkomende informatie uit die omgeving op te heffen. De externe stimulus wordt vergeleken met een interne standaard van die stimulus en eventuele discrepanties worden gereduceerd. Bij de beschrijving van dit proces maakt Hunt gebruik van de terminologie van Miller, Galanter en Pribram (1960). In hun opvatting over hoe mensen (en dieren) informatie verwerken is een centrale plaats ingeruimd voor wat zij een TOTE-eenheid noemen. TOTE staat voor test → operate → test → exit. Het individu vergelijkt binnenkomende informatie met een cognitief schema van die informatie ('test'). Zijn er discrepanties dan is de persoon gemotiveerd tot handelen ('operate'), cognitief of feitelijk, teneinde die discrepantie op te heffen. Vervolgens worden cognitie en werkelijkheid weer met elkaar vergeleken ('test'). Is de discrepantie opgeheven dan stopt de handeling ('exit'). Intrinsiek gemotiveerd gedrag is dus gedrag dat erop gericht is discrepanties op te heffen tussen de kennis die mensen van de werkelijkheid hebben en de werkelijkheid zelf. Volgens Hunt behoeft dit proces geen nadere verklaring: het is *inherent* aan het functioneren van het cognitieve systeem.

Overigens heeft dat systeem nog een andere eigenschap. Bij afwezigheid van incongruentie tussen stimulus en interne representatie heeft het organisme de neiging om zelf incongruentie te creëren — te zoeken naar nieuwe en interessante stimuli — alsof het doende is een optimaal niveau van incongruentie tussen zichzelf en zijn omgeving te handhaven. Dit verschijnsel wordt meestal fysiologisch verklaard. Het centrale zenuwstelsel kent een optimaal niveau van electrofysiologische activiteit ('arousal'). En het organisme probeert met behulp van externe stimulatie dat optimale arousal-niveau te bereiken en in stand te houden (Hebb, 1955).

Berlyne (1971) en Kagan (1972) beschrijven daarentegen intrinsiek gemotiveerd handelen niet zozeer als een proces van reductie van incongruentie, maar als reductie van *onzekerheid*. De behoefte om onzekerheid omtrent de werkelijkheid te verminderen activeert bij de persoon een informatiezoekproces dat door Berlyne 'nieuwsgierigheid naar kennis' of 'cognitieve nieuwsgierigheid' ('epistemic curiosity') genoemd wordt. Dit zoekproces

wordt in plaats van intrinsiek gemotiveerd ook wel *cognitief gemotiveerd* genoemd (McReynolds, 1971).

Voor het onderwijs lijken twee recente onderzoekslijnen met betrekking tot intrinsieke motivatie van belang: het onderzoek naar effecten van de aan- of afwezigheid van externe beloningen op intrinsieke motivatie, en het onderzoek naar de effecten van onderwijsprocedures die proberen cognitieve nieuwsgierigheid bij lerenden te induceren. Eerst enkele zinnen over de invloed van externe beloning op intrinsieke motivatie.

Externe beloning van gedrag blijkt over het algemeen de intrinsieke motivatie te reduceren (Deci, 1971; Lepper, Greene en Nisbett, 1973). Een karakteristiek experiment in dit kader is het volgende. Lepper en Greene (1975) brachten kleuters in een omgeving waarin ze zich bezig konden houden met het oplossen van een aantal puzzels. Van deze activiteit was vastgesteld dat hij door kleine kinderen als intrinsiek motiverend ervaren wordt. De helft van de kleuters mocht als beloning voor het oplossen van de puzzels met zeer aantrekkelijk speelgoed spelen. Aan de andere helft werd geen beloning in het vooruitzicht gesteld, maar beide groepen speelden na afloop met het speelgoed. In de periode na het experiment werd intrinsieke interesse in de puzzels gemeten door de hoeveelheid tijd te registreren die de kleuters spontaan doorbrachten met het werken aan de puzzels. Niet-extrinsiek beloonde kinderen bleken vaker en meer met de puzzels te spelen. Lepper en Greene (1975) concluderen dan ook dat externe beloning intrinsieke motivatie ondermijnt en 'spel in werk verandert' (p. 485). Sarafino en Di Mattia (1978) vonden een (gedeeltelijke) bevestiging van dit effect waar het de invloed van cijfergeven op de interesse in een leertaak betrof. Het geven van cijfers als beloning voor leerprestaties leek de intrinsieke interesse in de leertaak te ondermijnen.

Naast deze interesse voor de invloed van externe beloning op intrinsieke interesse in de leertaak is de aandacht van onderwijsonderzoekers uitgegaan naar procedures die de nieuwsgierigheid van lerenden naar nieuwe informatie bevorderen. Deze onderzoekers baseren zich daarbij voornamelijk op de hiervoor beschreven onzekerheidsreductietheorie van intrinsieke motivatie zoals die door Berlyne (1965, 1971) geformuleerd is. Bij dit onderzoek zal wat langer stilgestaan worden, omdat een aantal van in dat onderzoek gebruikte procedures ook toegepast zijn in de experimenten waarover in het navolgende gerapporteerd wordt.

David Johnson en zijn collega's (Johnson en Johnson, 1979; Lowry en Johnson, 1981; Smith, Johnson en Johnson, 1981) lieten in een serie verschillende experimenten groepen leerlingen uit het secundair onderwijs teksten bestuderen met betrekking tot onderwerpen met een controversiële karakter, bijvoorbeeld: steenkoolwinning in dagbouw. Het materiaal daarover bevatte steeds pro- en contra-argumenten. Met betrekking tot het onderwerp steenkoolwinning in dagbouw was er bijvoorbeeld een tekst over de economische noodzaak van deze activiteit beschikbaar en een tekst die de schade aan de natuur als gevolg van dagbouw benadrukte. De groep

leerlingen werd verdeeld in twee sub-groepen, waarbij de ene subgroep de tekst met pro-argumenten las, terwijl de andere de contra-argumenten bestudeerde. Vervolgens kreeg de groep als geheel de opdracht over het controversiële onderwerp te discussiëren. Daarbij moesten de deelnemers een aantal regels in het oog houden; onder andere: verdedig je eigen opvattingen, maar probeer ook de opvattingen van de tegenpartij te begrijpen. De docent had een verhelderende en stimulerende rol in dit geheel. Controlegroepen bestudeerden de teksten individueel of discussiëerden over het materiaal waarbij het belang van een snel bereiken van overeenstemming benadrukt werd (Smith et al., 1981). Na afloop werden verschillende metingen verricht om het niveau van intrinsieke motivatie vast te stellen. Daartoe werden zowel attitudes als feitelijk intrinsiek gemotiveerd gedrag gemeten. De attitudemetingen bestonden onder andere uit het invullen van Likert-schalen van het type 'Ik vond het onderwerp steenkoolwinning in dagbouw interessant'. Intrinsiek gemotiveerd handelen werd geobserveerd door na te gaan wie uit zichzelf een film ging bekijken die over het controversiële onderwerp handelde, en wie in de bibliotheek aanvullende informatie bestudeerde. Op zowel attitude- als gedragsmetingen van intrinsieke motivatie bleek de controversegroep consistent hoger te scoren. Volgens Johnson en Johnson (1979) komt dat omdat de controverse rondom een bepaald onderwerp in de leerling een — wat zij noemen — conceptueel conflict oproept. (Een conceptueel conflict ontstaat wanneer iemand geconfronteerd wordt met strijdige ideeën.) Dit conceptuele conflict leidt tot onzekerheid die volgens Berlyne (1971) gereduceerd moet worden. De — intrinsiek gemotiveerde — reductiereactie bestaat uit een verhoogde interesse in het onderwerp en een actief zoekproces naar informatie die de onzekerheid kan opheffen (Anderson, 1977).

Inagaki en Hatano (1977) legden aan 9- en 10-jarige kinderen de beschrijving van een conservatieprobleem voor. Dat probleem betrof de vraag of het gezamenlijk gewicht van een hoeveelheid suiker en water verandert als de suiker in het water wordt opgelost. Drie alternatieve oplossingen werden gegeven, ieder ondersteund met op het eerste gezicht plausibele argumenten. De leerlingen moesten, na eerst individueel een keuze te hebben gemaakt, gezamenlijk discussiëren over de oplossingsalternatieven teneinde te beslissen over het juiste antwoord. Vervolgens werd hen gevraagd op een vierpuntsschaal aan te geven hoe graag ze het conservatie-experiment (het oplossen van suiker in water op een balans) met eigen ogen zouden willen observeren. Hun antwoorden op deze vraag werden beschouwd als een maat voor hun intrinsieke interesse, zoals intrinsieke motivatie in de context van het onderwijs ook wel genoemd wordt. De gevolgde procedure — door de onderzoekers hypothese-experiment-instructie genoemd — bleek inderdaad de intrinsieke motivatie te verhogen. Inagaki en Hatano schrijven dat effect toe aan de onzekerheid die opgeroepen werd door het conceptuele conflict tussen even plausibele alternatieve oplossingen.

In het onderhavige artikel wordt onderzoek beschreven naar mogelijke mo-

tivationele effecten van een onderwijsmethode die in een aantal opzichten lijkt op de controverse-methode van Johnson en Johnson (1979) en op de hypothese-experiment-instructie-aanpak van Inagaki en Hatano (1977). Het gaat om een benadering die bekend staat onder de naam 'problem-based learning' (Barrows en Tamblyn, 1980; Neame, 1981) of 'probleemgestuurd onderwijs' (Schmidt, 1979). Deze methode kan beschouwd worden als een mutant van Bruners ontdekkend leren (Bruner, 1961) en de zogenaamde 'case-study'-methode (Fraser, 1931). Hieronder volgt een korte beschrijving van de benadering.

Aan een kleine groep studenten wordt een probleem voorgelegd. Dat probleem bestaat gewoonlijk uit een beschrijving van een verzameling verschijnselen of gebeurtenissen. Het kan daarbij gaan om natuurverschijnselen ('De haven van Bergen in Noorwegen blijft 's winters ijsvrij, terwijl de haven van Nyköping in Zweden, die vele breedtegraden zuidelijker ligt, meestal dichtvriest') of sociale of psychologische verschijnselen ('In gezin A noemen kinderen hun ouders bij de voornaam, in gezin B is dat niet het geval'). Het kan gaan om verschijnselen die iedereen bij zichzelf of anderen kan constateren ('Wanneer iemand enige tijd achter elkaar hard rent, dan versnelt zijn hartslag; zijn ademhaling wordt dieper en sneller, hij gaat transpireren en krijgt een rood hoofd') of om verschijnselen die uit wetenschappelijk onderzoek naar voren gekomen zijn ('Attitudes ten opzichte van minderheids-groepen hangen sterk samen met scores op Adorno's F-schaal').

De taak van de groep is deze verschijnselen te *verklaren* in termen van veronderstelde onderliggende processen, mechanismen, principes of theorieën. Zij doet dat door het probleem te analyseren. Dat houdt in dat de studenten ideeën en veronderstellingen die naar aanleiding van de probleembeschrijving bij hen opkomen in een gestructureerde discussie proberen te *explicitieren*. Die ideeën zijn gebaseerd op feitelijke voorkennis — je weet bijvoorbeeld dat de warme golfstroom langs de kust van Noorwegen stroomt — of zijn het resultaat van een kennisproducerend denkproces. In dat laatste geval *elaboreren* studenten op hun voorkennis (Reder, 1980) — je weet bijvoorbeeld niet dat de warme golfstroom de Oostzee niet bereikt, maar leidt dat af uit je kennis van het feit dat tussen Denemarken en Zweden slechts een nauwe zeeëngte ligt.

Het resultaat van deze discussie is een beschrijving van het proces, principe of mechanisme dat volgens de groep aan het probleem ten grondslag zou kunnen liggen. Het probleem is — als het goed is — echter zodanig dat studenten het met de bij hen aanwezige kennis niet adequaat kunnen verklaren. Dus na afloop van de discussie blijven er dan een aantal vragen en onzekerheden over. Daarom bestuderen de studenten vervolgens individueel probleemrelevante informatie die hen helpt die onduidelijkheden te verhelderen en onzekerheid te reduceren. Een tutor heeft bij de discussie een verhelderende en stimulerende rol.

Voor meer uitgebreide beschrijvingen van het onderwijsleerproces: zie Barrows en Tamblyn (1980) of Schmidt en Bouhuijs (1980).

In deze beschrijving vallen de overeenkomsten met beide andere hiervoor

beschreven benaderingen op. In alle drie de procedures is er sprake van een groep die een bepaald probleem bespreekt. In die discussie worden argumenten en ideeën geactiveerd en bekritiseerd. Dit activeren van ideeën en het elaboreren daarbij induceert een zoekproces naar (meer) probleemrelevante informatie.

Er zijn natuurlijk ook verschillen. In de controverse-methode van Johnson en zijn collega's krijgen leerlingen voorafgaande aan de discussie al enige informatie; bij beide andere methoden is dat niet of nauwelijks het geval. Belangrijker is dat in de controverse-methode en in de Japanse aanpak sterk de nadruk wordt gelegd op het conceptuele conflict dat tot stand wordt gebracht door de leerlingen expliciet te confronteren met elkaar wederzijds uitsluitende verklaringen of argumenten. In de hierboven beschreven methode van probleemanalyse ligt de nadruk juist sterk op de discrepantie tussen de kennis die nodig is om het probleem te begrijpen en de bij studenten al aanwezige *voorkennis*. Het is met name dit laatste verschil dat het interessant maakt na te gaan of de probleemanalyse-methode evenzeer tot een intrinsiek gemotiveerd zoeken naar informatie leidt als blijktbaar bij de beide andere methoden het geval is. Met andere woorden: of conceptueel *conflict* een noodzakelijke voorwaarde voor intrinsieke motivatie is of niet.

Samenvattend: de theorie van Berlyne (1971) met betrekking tot intrinsieke motivatie voorspelt dat probleemanalyse bij studenten tot cognitieve onzekerheid leidt. De behoefte deze onzekerheid te reduceren, doet ontstaan wat Berlyne noemt 'epistemic curiosity'. Er treedt een intrinsiek gemotiveerd zoekproces op naar probleemrelevante informatie. Daaruit kan men afleiden dat studenten die aan een probleem gewerkt hebben een grotere intrinsieke interesse zullen hebben in probleemrelevante informatie dan studenten die geen probleem hebben voorgelegd gekregen.

In Experiment 1 werd nagegaan of en in hoeverre de probleemanalyse-methode intrinsieke motivatie oproept.

In Experiment 2 werden relaties tussen intrinsieke motivatie en studieprestaties onderzocht.

Experiment 1

In Experiment 1 kregen groepjes van 6 of 7 studenten een probleem voorgelegd met het verzoek daarover te discussiëren. Het probleem bestond uit de beschrijving van een aantal (biologische) fenomenen en de studenten moesten proberen die fenomenen te verklaren in termen van een onderliggend proces. Nagegaan werd of deze probleemanalyse-methode – in overeenstemming met Berlyne's onzekerheidsreductietheorie – de interesse voor nieuwe, op dat proces betrekking hebbende, informatie vergroot. Interesse in nieuwe informatie en de bereidheid om een bijeenkomst over het betreffende onderwerp bij te wonen, werden beschouwd als operationalisaties van intrinsieke motivatie.

Methode

Proefpersonen

39 studenten van een instelling voor hoger beroepsonderwijs (een opleiding voor kinderverzorging en -opvoeding) namen deel aan Experiment 1: 31 vrouwen en 8 mannen. De proefpersonen hadden dezelfde vooropleiding: allen waren in het bezit van een HAVO-diploma. Als verdere voorwaarde voor deelname werd gesteld dat de proefpersonen geëxamineerd waren in het vak biologie. De gemiddelde leeftijd was 19,4 jaar met een standaardafwijking van 1,1. De proefpersonen ontvingen voor hun medewerking een vergoeding.

Materiaal

Het materiaal bestond uit een probleembeschrijving, twee vragenlijsten betrekking hebbend op motivationele aspecten van de probleemanalyse, en twee kennistoetsen: een reproductie- en een transfertoets.

De probleembeschrijving bestond uit de volgende tekst: 'Een rode bloedcel (een rood bloemlichaampje) wordt onder een microscoop in zuiver water gebracht. De bloedcel zwelt in korte tijd op en springt uiteindelijk kapot. Een andere bloedcel wordt in een oplossing van zout in water gebracht. Hij schrompelt ineen. Hoe zijn deze verschijnselen te verklaren?'

De taak van de proefpersonen was dus een verklaring te bedenken voor deze fenomenen. Die verklaring moest geformuleerd worden in termen van een verondersteld aan deze verschijnselen ten grondslagliggend proces. In het geval van de opzwellende of ineenschrompelende bloedcel is dat het osmotisch proces. (Voor lezers die niet vertrouwd zijn met dit proces wordt vermeld dat osmose een proces is waarbij een levende cel door een semidoorlaatbare membraan water aantrekt uit een omgeving die een lagere concentratie zouten en andere opgeloste stoffen bevat dan die cel. Omdat de wand van een bloedcel geen eigen stevigheid bezit, is die wand niet bestand tegen de druk van het aangezogen watervolume en scheurt. Als echter de concentratie van stoffen, opgelost in water in de omgeving, hoger is dan in de cel zelf, wordt juist water aan de cel onttrokken. Osmose berust op eigenschappen van het semipermeabele membraan en op het proces van diffusie.)

De afhankelijke variabelen bestonden uit twee metingen betrekking hebbend op motivationele aspecten, twee kennistoetsen en — voor de probleemanalyse-groep — een satisfactiemeting.

Interesse in het onderwerp osmose als maat voor intrinsieke motivatie werd op twee manieren gemeten: de eerste was een attitudemeting; de tweede een gedragsmeting.

De attitudemeting had de volgende vorm: de proefpersonen kregen vier regels informatie over osmose. Vervolgens werd hen gevraagd: Hoe geïnteresseerd ben je op dit moment in nadere informatie over het onderwerp osmose? De proefpersonen konden de mate van hun interesse in dat onderwerp aangeven door omcirkeling van een getal op een 5-puntsschaal. Het

getal 1 stond daarbij voor 'in het geheel niet'. De schaal liep verder via 'nauwelijks', 'wel enigszins' en 'tamelijk veel' naar 'bijzonder veel'. Deze meting van de attitude van proefpersonen ten opzichte van osmose is vergelijkbaar met onder andere datgene wat Inagaki en Hatano (1977) en Smith et al. (1981) gedaan hebben.

De gedragsmeting had de volgende vorm: de proefpersonen ontvingen een schrijven, waarin een lezing-met-discussie over het onderwerp 'osmose en het menselijk lichaam' werd aangekondigd. De proefpersonen konden voor deze lezing intekenen door middel van het zetten van een handtekening op een lijst. Het wel of niet intekenen voor deze — facultatieve — lezing werd beschouwd als een gedragsmaat voor de intrinsieke motivatie van de proefpersonen met betrekking tot het onderwerp osmose. Vergelijkbare benaderingen van de operationalisatie van intrinsieke motivatie zijn gekozen door Maehr (1976) en Lowry en Johnson (1981).

In de experimentele groep — de groep die met het bloedcelprobleem geconfronteerd werd — werd een *satisfactiemeting* gedaan. De vragenlijst die daarvoor gebruikt werd, is ontleend aan Schmidt en Bouhuijs (1977) en bestaat uit 11 items met behulp waarvan de proefpersoon de wijze waarop de groep waarvan hij deel uitmaakt aan het probleem werkte, kan karakteriseren. De items hebben de volgende vorm.

prettig	1	2	3	4	5	onprettig
vervelend	1	2	3	4	5	leuk

Schmidt en Bouhuijs (1977) veronderstellen dat deze vragenlijst een algemeen gevoel van tevredenheid met het functioneren van de groep meet, met name omdat de items prettig-onprettig, aangenaam-onaangenaam en vervelend-leuk het hoogst laden op de (ene) factor die de vragenlijst beschrijft.

De twee *toetsen* waren geconstrueerd aan de hand van een tekst over osmose en diffusie die in Experiment 2 gebruikt werd. Ze zijn voor Experiment 1 van secundair belang en worden daarom uitgebreider beschreven in de methodesectie van Experiment 2. (Zie ook Schmidt, 1982b.)

Ten slotte werden enkele *biografische gegevens* van de proefpersonen vastgelegd: leeftijd, geslacht en eindcijfer voor het vak biologie op de HAVO.

Procedure

De proefpersonen werden aselekt toegewezen aan één van de condities van het onderzoek. De experimentele groep bestond uit 20 proefpersonen; de controlegroep bevatte 19 proefpersonen.

De experimentele groep werd opgedeeld in 3 sub-groepen, waarbinnen de analyse van het bloedcelprobleem zou plaatsvinden. Twee subgroepen bestonden uit 7 personen, één uit 6. Aan elk van deze groepen werd een vrouwelijke tutor toegewezen. Deze tutoren hadden ervaring met de hiervoor beschreven methode van probleemanalyse en waren vertrouwd met het onderwerp osmose. Tutoren en proefpersonen hadden elkaar voor de experimentele sessie nooit eerder ontmoet.

De probleemanalyse vond plaats in klaslokalen, waarbij stoelen en tafels in

een vierkant gerangschikt waren.

De tutor ging op de volgende wijze te werk: ze besteedde eerst enige tijd aan een kennismaking en aan het op hun gemak stellen van de proefpersonen. Vervolgens introduceerde ze de methode aan de hand van een voorbeeld. De proefpersonen hadden namelijk geen ervaring met de van hen gevraagde aanpak. Zij reikte daartoe de volgende probleembeschrijving uit: 'Waterpest wordt in een bekerglas met water geplaatst dat rijk is aan kool-dioxide. Als dit glas aan het licht wordt blootgesteld komen er gasbelletjes vrij. Als deze worden opgevangen in een reageerbuisje kan men met een gloeiende houtspaander aantonen dat het ontwijkende gas zuurstof is. Wordt het licht afgeschermd dan houdt de zuurstofproductie op.'

Aan de hand van dit probleem liet ze de verschillende stappen van de analyse zien. Op een schoolbord dat als hulpmiddel werd gebruikt, stond de volgende tekst: Probleemanalyse: 1. Nadenken. 2. Definitie van het probleem. 3. Verklaringen bedenken. 4. Uitdieping. De proefpersonen werden bij de uitleg van de methode actief betrokken.

Voordat het bloedcelprobleem werd aangeboden, vertelde de tutor dat de probleemanalyse niet langer dan 15 minuten zou duren. Het probleem werd uitgedeeld en de groep kreeg 1 minuut voor lezen en nadenken. Daarna startte de discussie. De tutor trad als gespreksleider op, dat wil zeggen: zij ordende datgene wat naar voren gebracht werd. Ze had expliciet de opdracht geen informatie over het probleem te verschaffen en erop toe te zien dat ook haar samenvattingen niet impliciet informatie bevatten, waaraan proefpersonen inzichten met betrekking tot het probleem zouden kunnen ontleen.

Er werd een geluidsopname van de discussie gemaakt, om achteraf te kunnen controleren of tutoeren en groepen zich aan de afgesproken aanpak hadden gehouden. Dit bleek het geval. Tutoren hielden zich nauwgezet aan de gegeven instructies. Alle groepen hadden minder dan 10 minuten nodig voor de analyse.

Direct na afloop van de probleemanalyse beantwoordden alle proefpersonen schriftelijk de vraag naar hun interesse in nadere informatie over osmose. Vervolgens vulden ze de satisfactievragenlijst in en maakten ze de toetsen. Er was aan de beantwoording van de vragen geen tijdslimiet gesteld. De controlegroep waaraan geen probleembeschrijving was voorgelegd beantwoordde de interessevraag en de toetsen.

Eén dag na de groepssessies en de afname van de vragenlijsten werd op de school waarop Experiment 1 werd uitgevoerd schriftelijk de lezing 'osmose en het menselijk lichaam' aangekondigd. Onder de studenten (proefpersonen en anderen) circuleerden lijsten waarop door middel van het plaatsen van naam en handtekening aangegeven kon worden of de lezing zou worden bijgewoond. De onderzoeker was daarbij niet aanwezig om de relatie tussen lezing en experiment niet te duidelijk te leggen.

Resultaten en discussie

Op de variabele 'behoefte aan nadere informatie over osmose' werd een variantieanalyse uitgevoerd. De resultaten van deze analyse zijn samengevat in tabel 1 en 2. In tabel 1 zijn gemiddelde scores en standaarddeviaties voor probleemanalyse- en controlegroep opgenomen.

Tabel 1: Gemiddelden en standaarddeviaties van interessescores voor Experiment 1.

	$\bar{X}$	SD	N
Probleemanalyse groep	3,20	0,83	20
Controlegroep	2,42	0,84	19
Totaal	2,82	0,91	39

In tabel 2 vindt men de resultaten van de variantieanalyse.

Tabel 2: Variantieanalyse op interessescores voor Experiment 1.

Bron	Kwadraten-som	df	Kwadraten-gemiddelde	F
Experimentele manipulatie	5,91	1	5,91	8,46
Residu	25,83	37	0,70	
Totaal	31,74	38	0,82	

Uit het resultaat van de variantieanalyse op interessescores kan worden afgeleid dat de gebruikte experimentele manipulatie een statistisch significant effect heeft op interesse in osmose: $F = 8,46$, $p < 0,007$. De gezamenlijke analyse van het bloedcelprobleem vergroot dus de belangstelling die proefpersonen aan de dag leggen voor nadere informatie over het onderwerp osmose. Uit de gemiddelden van beide groepen kan men afleiden dat het daarbij gaat om een betrekkelijk subtiel verschil. Proefpersonen die niet geconfronteerd zijn met het bloedcelprobleem zijn tussen 'nauwelijks' en 'wel enigszins' geïnteresseerd in nadere informatie, terwijl de probleemanalysegroep tussen 'wel enigszins' en 'tamelijk veel' geïnteresseerd is. De conclusie moet dus zijn dat probleemanalyse een zekere verhoging te weeg brengt van de intrinsieke motivatie met betrekking tot osmose.

Bij het gebruik van vragenlijsten in psychologisch onderzoek moet men er

altijd alert op zijn dat discrepanties kunnen bestaan tussen wat mensen zeggen en wat ze feitelijk *doen*: tussen hun attitudes en hun gedrag (Fishbein en Ajzen, 1974). In het hier gerapporteerde onderzoek zou er een verschil kunnen zijn tussen wat proefpersonen meedelen over hun interesse in osmose en hun *feitelijke* interesse. Het is daarom interessant na te gaan in hoeverre de geconstateerde verschillen in interesse tussen beide groepen ook optreden wanneer proefpersonen gevraagd wordt een lezing over osmose bij te wonen. In tabel 3 vindt men relatieve frequenties van proefpersonen die wel of niet aangegeven hadden bij de lezing aanwezig te zijn.

Tabel 3: Proporties van proefpersonen in Experiment I die zich wel of niet voor een lezing hebben opgegeven.

	'heeft zich op- gegeven'	'heeft zich niet opgegeven'
Probleemanalysegroep	0,75	0,25
Controlegroep	0,53	0,47
Totaal	0,64	0,36

Het verschil tussen de proporties van diegenen uit beide groepen die zich hadden opgegeven voor de lezing werd statistisch getoetst (Ferguson, 1966). Het resultaat van die test was als volgt: $z = 1,47$, $p < 0,08$, eenzijdig. Geconcludeerd kan dus worden dat er aanwijzingen zijn dat probleem-analyse in een kleine groep de bereidheid vergroot om naar een lezing te komen luisteren waarop probleemrelevante informatie zal worden gegeven. De punt-biseriële correlatiecoëfficiënt tussen attitude- en gedragsmeting is hoog: $r_{pbis} = 0,74$, zeker als men bedenkt dat een dergelijke correlatie in werkelijkheid nooit een maximum van 1 bereikt. Zijn maximum is namelijk onder andere een functie van de scheefheid van de proportieverdeling van de dichotome variabele. Het is daarbij niet onwaarschijnlijk dat deze correlatiecoëfficiënt een onderschatting is van de ware correlatie tussen attitude en gedrag in dit onderzoek. De betrouwbaarheid (in de zin van precisie) van de gedragsmeting werd namelijk aantoonbaar verlaagd door factoren die irrelevant zijn ten aanzien van de vraagstelling, zoals het feit dat enkele proefpersonen achteraf aangaven wel naar de lezing te hebben willen komen maar op het aangegeven tijdstip verhinderd te zijn geweest. Op basis van de hier gerapporteerde gegevens kan gesteld worden dat interessemeting door middel van zelfrapportage een valide maat blijkt te zijn voor intrinsieke motivatie met betrekking tot het onderwerp osmose. Bij de beschrijving van de effecten van probleemanalyse op de intrinsieke motivatie van studenten is er eigenlijk stilzwijgend van uit gegaan dat het *cognitieve* effecten betreft. Dat wil zeggen dat de geconstateerde interesse in osmose beschouwd wordt als 'cognitieve nieuwsgierigheid' (epistemic

curiosity) in de zin door Berlyne bedoeld. Inagaki en Hatano (1977) en McReynolds (1971) gebruiken zelfs de term 'cognitieve motivatie' voor het proces dat aan het interessefenomeen ten grondslag ligt. De persoon wil graag meer gaan begrijpen van het betreffende onderwerp. Hij voelt de drang activiteiten te gaan ondernemen die hem helpen bij het verwerven van een dieper inzicht.

Het is echter niet uitgesloten dat het hier zozeer gaat om een cognitief effect maar eerder om een *affectief* effect. Men zou zich kunnen voorstellen dat de persoon onder invloed van het (succesvol) werken aan de taak in een toestand geraakt van zich prettig voelen, van algeheel welbevinden, die zich onder andere uit in een verhoogde neiging om in positieve zin te reageren op de vraag naar interesse in osmose. De persoon voelt zich lekker, is in een goed humeur en antwoordt daarom dat hij in nadere informatie geïnteresseerd is. In de intrinsieke-motivatietheorie van Deci (1975) wordt er expliciet van uit gegaan dat intrinsieke motivatie een affectieve toestand is. Het positieve affect dat een persoon bij het verrichten van bepaalde taken ervaart, brengt hem ertoe die taken ook weer bij voorkeur op te zoeken. Het in Experiment 1 verzamelde materiaal geeft de gelegenheid na te gaan of 'interesse in osmose' affectief gekleurd zou kunnen zijn. Als dat het geval is dan zou 'interesse in osmose' positief moeten correleren met de resultaten op de satisfactieschaal van Schmidt en Bouhuijs (1977), die immers zo iets meet als een gevoel van tevredenheid met het functioneren van de groep. De berekening van deze correlatie moet plaatsvinden over de helft van de proefpersonen omdat de satisfactieschaal slechts in de probleem-analysegroep kon worden afgenomen. De correlatie tussen de attitudemeting van interesse en de satisfactiescore (= de som van de scores op de afzonderlijke items) was gelijk aan $r = 0,15^1$; die tussen de gedragsmeting en de satisfactiescore was gelijk aan $r_{pbis} = 0,26^2$. Beide correlaties verschillen niet significant van 0.

Daaruit zou men kunnen afleiden dat de gevonden effecten geen uitvloeisel zijn van bepaalde affectieve toestanden, maar een werkelijk cognitief effect beschrijven, in de door Berlyne (1971) bedoelde zin.

Wat betekenen nu deze resultaten voor de in de inleiding geuite veronderstellingen? Het lijkt erop dat analyse van een probleem inderdaad het door Berlyne (1978) en Hunt (1971) beschreven motivationele proces op gang brengt. De explicitering en elaboratie van voorkennis, betrekking hebbend op het probleem (in Hunts terminologie: het testen van hypothesen aan de verschijnselen om na te gaan of er een incongruentie bestaat tussen stimulus en interne representatie van die stimulus) leidt ertoe dat proefpersonen incongruentie ervaren tussen wat ze weten en wat ze zouden moeten weten om de in het probleem beschreven verschijnselen adequaat te kunnen verklaren. Als gevolg van de door deze procedure geïnduceerde onzeker-

¹ De betrouwbaarheid van de satisfactieschaal was gelijk aan $\alpha = 0,98$.

² In Experiment 2 was de correlatie tussen attitude en satisfactie gelijk aan $r = 0,00$.

heid leggen proefpersonen een verhoogde — intrinsiek gemotiveerde — interesse aan de dag voor probleemrelevante informatie en zijn ze bereid activiteiten te ontplooiën die die informatiebehoefte kunnen bevredigen. Cognitieve onzekerheid leidt tot epistemic curiosity (Berlyne, 1978).

Experiment 2

In experiment 2 werd een poging gedaan de resultaten van Experiment 1 te repliceren en mogelijke relaties te onderzoeken tussen intrinsieke motivatie en studieprestaties. Daarnaast werden mogelijke veranderingen in intrinsieke motivatie onder invloed van informatieverwerking onderzocht. In de inleiding tot dit artikel is gesuggereerd dat intrinsieke motivatie belangrijker wordt als motiverende kracht in het onderwijsleerproces naarmate externe beloningen minder een rol gaan spelen. Die suggestie gaat impliciet van de these uit dat intrinsieke motivatie een causale factor is in het onderwijsleerproces en tot betere studieprestaties leidt. De literatuur met betrekking tot de vraag of er bepaalde relaties bestaan tussen beide onderwijsvariabelen, en zo ja: welke, is betrekkelijk schaars. Onderzoek heeft zich voornamelijk bezig gehouden met intrinsieke motivatie als 'outcome' variabele. De vraag was steeds: verhogen bepaalde onderwijsleerprocessen interesse en spontane leeractiviteit bij proefpersonen? en niet: leidt die verhoogde interesse of spontane activiteit ook tot grotere competentie, tot betere prestaties op het terrein waarvoor die interesse aanwezig is? Enige indicaties omtrent wat er te verwachten valt, zijn niettemin toch wel aan die literatuur te ontleen.

In de al eerder geciteerde onderzoeken van Johnson en zijn collega's (Johnson en Johnson, 1979; Lowry en Johnson, 1981; Smith, Johnson en Johnson, 1981) wordt consistent gevonden dat groepen leerlingen die grotere interesse in een bepaald onderwerp vertonen (eerder een film gaan zien; meer gebruik maken van bibliotheekfaciliteiten) op een later tijdstip een betere studieprestatie leveren dan groepen leerlingen die minder interesse tonen. Impliciet lijken deze auteurs daaruit te concluderen dat interesse een determinerende factor is van studieprestaties. Zij rapporteren echter geen correlaties tussen interesse en studieprestaties, die een directe inspectie mogelijk zouden maken van de grootte van de relatie. De experimenten van Johnson laten in feite de mogelijkheid open dat intrinsieke motivatie en studieprestaties van elkaar onafhankelijke grootheden zijn, die wel beide afhankelijk zijn van een derde variabele: het wel of niet gediscussieerd hebben over een controversiële onderwerp. Het is namelijk zeer wel denkbaar dat een controversie-discussie zowel de intrinsieke motivatie stimuleert als bepaalde cognitieve veranderingen bij proefpersonen tot stand brengt die de verwerking van nieuwe probleemrelevante informatie vergemakkelijken. Die processen kunnen onafhankelijk van elkaar verlopen.

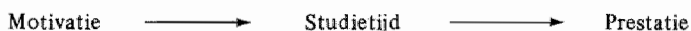
Inagaki en Hatano (1977) rapporteren wel een correlatie tussen interesse en studieprestatie. Die correlatie is gelijk aan $r = 0,12$; $p > 0,10$. Dit resul-

taat komt overeen met de resultaten van een serie onderzoeken die Bloom (1976) geeft. Het ging in die onderzoeken om effecten van mastery learning in een groot aantal verschillende vakgebieden. In dit onderzoek werden correlaties berekend tussen interesse in een onderwerp, gemeten aan het begin van een studietask of serie van taken, en prestatie, gemeten na voltooiing van die taken. De mediane correlatie tussen interesse en studieprestaties was in die studies gelijk aan $r = 0,06$. Daartegenover echter staan resultaten uit de grootschalige studies die de International Association of Educational Achievement (I.E.A.) gedaan heeft naar relaties tussen interesse in bepaalde vakken, zoals wiskunde, en prestaties in die vakken (zie onder andere Carroll, 1975 en Husén, 1967). In die onderzoeken werden met grote regelmaat correlaties in de orde van grootte van 0,30 gevonden.

Uit dit overzicht kan men afleiden dat het verband tussen beide onderwijsvariabelen — zo er een verband is — tamelijk gecompliceerd moet zijn. Het verdient daarom misschien aanbeveling nader in te gaan op de vraag hoe men zich de samenhangen tussen aan deze variabelen ten grondslag liggende motivationele en cognitieve processen zou kunnen voorstellen. Er zijn dan vijf mogelijkheden die zich daarbij voordoen; alle ontleend aan de ruimere motivatieliteratuur.

De eerste mogelijkheid is dat motivatie prestatie beïnvloedt, maar op een indirecte wijze. Carroll (1963) veronderstelt dat motivatie vooral de bereidheid van leerlingen zou beïnvloeden om bij een leertaak te persisteren. De hoog-gemotiveerde leerlingen zouden bereid zijn meer tijd aan hun studie te besteden dan laag-gemotiveerden. Verlenging van de studietijd leidt op zijn beurt tot betere studieprestaties (Bloom et al., 1971). In die opvatting intervieneert studietijd tussen motivatie en prestatie. Zie schema 1; pijltjes geven causale invloed weer.

Schema 1: Relatie tussen motivatie en prestatie volgens Carroll (1963).

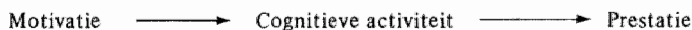


Dit model zou een verklaring kunnen vormen voor het feit dat in het onderzoek van Inagaki en Hatano (1977) geen verband tussen motivatie en prestatie werd gevonden. Het voorspelt immers dat in experimenten waarin studietijd constant gehouden wordt geen verband gevonden kan worden tussen beide variabelen. En in de Japanse studie was studietijd gefixeerd. Met betrekking tot de andere in het overzicht hierboven genoemde empirische bevindingen kan niets zinnigs gezegd worden vanuit Carrolls perspectief omdat bij dat onderzoek niets gemeld wordt over studietijd.

De tweede mogelijke verklaring veronderstelt dat motivatie de aard en intensiteit van de cognitieve activiteit tijdens de leertaak beïnvloedt. Hoog-gemotiveerden leveren dan een betere prestatie dan laag-gemotiveerden

omdat de eersten gedurende het leren meer, en verschillende soorten cognitieve bewerkingen op de te bestuderen informatie uitvoeren. Het is bekend dat een 'diepere' verwerking van informatie tot betere retentie en andere prestaties leidt (Craik en Lockhart, 1972). Elaboreren gedurende het bestuderen van een tekst leidt bijvoorbeeld tot betere studieprestaties (Reader, 1980). Deze opvatting over hoe motivatie het leren beïnvloedt, vindt ondersteuning in onderzoek van Fass en Schumacher (1978). Fass en Schumacher lieten proefpersonen een makkelijke, dan wel een moeilijke versie van een bepaalde tekst bestuderen. Deze proefpersonen werden wel of niet gemotiveerd door het in het vooruitzicht stellen van een som gelds. Degene die voor hun medewerking betaald werden (door de onderzoekers gedefiniëerd als de hoog-gemotiveerden) leverden op de makkelijke en de moeilijke versie van de tekst een even goede prestatie, terwijl de laag-gemotiveerden een betere prestatie leverden op de makkelijke versie van de tekst dan op de moeilijke versie. Fass en Schumacher (1978) verklaren dit door te veronderstellen dat een makkelijke tekst in principe tot betere prestaties leidt dan een moeilijke tekst — een effect dat bij laag-gemotiveerden zichtbaar wordt — maar dat hoog-gemotiveerden voor de moeilijkheidsgraad van een tekst compenseren door verhoogde cognitieve activiteit. Klare (1976) vond resultaten die vergelijkbaar zijn met die van Fass en Schumacher. In schema 2 wordt deze theorie samengevat.

Schema 2: Motivatie en prestatie volgens Fass en Schumacher (1978).



Een derde theorie is dat motivatie een affectief bijproduct zou kunnen zijn van eerdere leerervaringen met een bepaald onderwerp en daaruit voortvloeiende prestaties. Een observatie die regelmatig gedaan kan worden is dat sommige studenten in de sociale wetenschappen laag-gemotiveerd zijn ten aanzien van het vak statistiek op grond van eerdere negatieve leerervaringen met wiskunde. Omgekeerd zouden eerdere positieve ervaringen (en prestaties!) een hogere motivatie tot gevolg hebben. Met andere woorden: de goede studenten zijn niet goed omdat ze op een eerder tijdstip hoog-gemotiveerd waren, maar zijn hoog-gemotiveerd omdat ze eerder goed waren. Hoge motivatie leidt niet tot betere prestaties; prestaties leiden tot hogere motivatie. Hoe meer je van een bepaald onderwerp begrijpt, hoe meer je in het onderwerp geïnteresseerd raakt of bent. Deze opvatting zijn we eerder tegengekomen bij Deci (1975). Deci zegt dat mensen een fundamentele behoefte hebben zich competent te voelen en daarom interesse tonen voor die activiteiten waarin ze zich competent kunnen voelen. Empirische ondersteuning voor dit idee kan gevonden worden bij Bloom (1976). Uit door hem gepresenteerd onderzoek kwam naar voren dat de correlatie tussen studieprestaties op een bepaald onderdeel en interesse

oor een *volgend* — verwant — onderdeel gemiddeld gelijk was aan $r = 0,30$. Dat zou erop kunnen duiden dat prestatie motivatie induceert. In de eerder geciteerde I.E.A.-studies werden prestatie en interesse *tegelijkertijd* gemeten. Dat zou de relatief hoge correlaties kunnen verklaren.

Schema 3: Motivatie en prestatie volgens Deci (1975)



dit verband is een observatie van Bloom (1976) interessant. Zijn studenten vonden over het algemeen geen correlatie tussen eerdere interesse en daaropvolgende prestatie, met uitzondering van een cursus in biologie. Daarbij bleek de correlatie tussen eerdere interesse en latere prestatie gelijk aan $r = 0,31$ te zijn. Die cursus onderscheidde zich echter van de andere cursussen, doordat de proefpersonen met dat vak al eerder ervaringen hadden opgedaan. Aannemelijk is dat de gevonden correlatie van 0,31 niet zozeer een afspiegeling is van de invloed van motivatie voor biologie op prestatie, maar van eerdere prestatie in biologie op motivatie. Ook Ausubel (1968) is van mening dat motivatie beschouwd moet worden als een afgeleide van prestatie.

De vierde theorie zegt dat interesse en prestatie twee van elkaar onafhankelijke uitkomsten van een leerproces kunnen zijn. Sommige leerervaringen leiden bij sommige leerlingen tot een verhoogde interesse en goede leerprestaties. Dat sluit niet uit dat leerervaringen die niet tot een verhoogde interesse leiden niet tot goede leerprestaties zouden kunnen leiden. Eventuele correlaties tussen interesse en prestatie zijn in dat geval spurieus, en terug te voeren op een derde variabele die met de leertaak verbonden is.

Schema 4: Motivatie en prestatie zijn onafhankelijke uitkomsten van een zelfde leerproces.



en slotte zijn er natuurlijk ook combinaties van de hier gegeven modellen mogelijk. Men kan zich bijvoorbeeld een proces voorstellen waarbij beide variabelen elkaar wederzijds beïnvloeden op spiraalvormige wijze. Een hoge motivatie leidt volgens model 1 of 2 tot een goede prestatie, die op zijn beurt volgens model 3 tot een hogere motivatie leidt, die weer tot ..., enzovoorts.

In Experiment 2 zal geprobeerd worden de hier beschreven opvattingen over de relatie tussen interesse in een onderwerp en prestatie met betrekking tot dat onderwerp te evalueren. (Daarbij moet Carrolls opvatting buiten beschouwing blijven, omdat in Experiment 2 studietijd gefixeerd zal zijn.)

De opzet van het experiment is als volgt: Teneinde verschillen in interesse in het onderwerp osmose te induceren, is aan een groep proefpersonen gevraagd het bloedcelprobleem uit Experiment 1 te analyseren; dit werd de experimentele groep. Interesse in osmose is gemeten bij experimentele en controlegroep. Vervolgens is aan beide groepen een tekst over osmose ter bestudering voorgelegd waarbij studietijd constant gehouden werd. Ten slotte werden studieprestaties gemeten en werd opnieuw een interessesmeting gedaan.

Methode

Proefpersonen

48 studenten van een instelling voor hoger beroepsonderwijs (een opleiding voor dietetiek en toegepaste huishoudkunde) namen deel aan Experiment 2: 42 vrouwen en 6 mannen. Alle proefpersonen hadden een vergelijkbare vooropleiding: HAVO, met biologie als eindexamenvak. De gemiddelde leeftijd van de proefpersonen was 19,4 jaar met een standaardafwijking van 1,6. De proefpersonen ontvingen een vergoeding.

Materiaal

Het materiaal dat bij Experiment 2 gebruikt werd, was identiek aan dat van Experiment 1. Om praktische redenen werd de gedragsmeting na de probleemanalyse weggelaten. Daarnaast werd een tekst over osmose gebruikt. De tekst was getiteld 'Osmose en diffusie'. Hij was ongeveer 2220 woorden lang en zodanig geschreven dat hij aansloot bij datgene wat op de HAVO aan de orde komt. In de tekst werden de volgende onderwerpen behandeld: diffusie, diffusiesnelheid, (semi)permeabiliteit van de celwand, concentratieverschillen, osmose, osmotische druk, osmotische waarde, structuur van plantencellen, turgor en plasmolyse.

De afhankelijke variabelen bestonden uit twee kennistoetsen en verschillende interessesmetingen. De kennistoetsen waren dezelfde als die in Experiment 1 gebruikt werden. Ze zullen hier iets uitgebreider beschreven worden.

De *herkenningstoets* bestond uit 33 items. Deze items waren tot stand gekomen door parafrasering en transformatie van de verschillende tekstgedeelten. 31 van de 33 items hadden de juist-onjuist-vorm; 2 items bevatten drie antwoordalternatieven. Na analyse van de data werden 10 items, als zijnde te gemakkelijk (p -waarde $> 0,90$), uit de toets verwijderd. De interne consistentie van de herkenningstoets was gelijk aan $\alpha = 0,73$.

De *transfertoets* bestond uit 37 items (na analyse van de data werden 5 items verwijderd). Het merendeel van deze items bestond uit kleine opga-

ven, die niet in de tekst behandeld werden maar met behulp van informatie uit de tekst wel konden worden opgelost. Het toepassen van informatie uit de tekst over osmose vormde dus het meetdoel van de transfertoets. Echter ook inferenties uit de tekst – beweringen die niet in de tekst zelf voorkomen maar er wel uit afgeleid kunnen worden – maakten deel uit van dit instrument. Zijn betrouwbaarheid was gelijk aan $\alpha = 0,69$.

Aan alle items was een vraagtekencategorie toegevoegd, zodat proefpersonen konden aangeven wanneer ze het juiste antwoord niet wisten.

De interessemetingen bestonden uit de interessevraag die ook in Experiment 1 na de probleemanalyse werd voorgelegd, en uit een attitude- en een gedragsmeting die na de bestudering van de tekst werden afgenomen.

Het *attitudemeetinstrument* bestond uit de volgende 3 items:

1. 'Denk je dat je in de toekomst nog wel eens iets wil lezen over het onderwerp osmose?'
2. 'Zou je geïnteresseerd zijn in toezending door de onderzoeker van schriftelijk materiaal over het onderwerp, zodat je dat thuis zou kunnen bestuderen?'
3. 'Als er van de kant van de deelnemers aan dit onderzoek voldoende belangstelling is, zal de school een bijeenkomst organiseren over "osmose en het menselijk lichaam". Onderstaande vraag is bedoeld om de belangstelling daarvoor te peilen. 'Ben je geïnteresseerd in een dergelijke bijeenkomst?'

Deze items moesten beantwoord worden op een vijf-puntsschaal verlopend van 'nou en of!', via 'ik denk het wel', 'ik weet het niet', 'ik denk het niet' naar 'zeker niet!'.

De *gedragsmeting* bestond uit de volgende procedure: Na afloop van het experiment kregen de proefpersonen een tekst voor ogen die duidelijk met enige haast geschreven en gekopieërd leek. Daarmee werd een poging gedaan bij de deelnemers de indruk te laten ontstaan dat het aanbod dat hen via de kopie bereikte spontaan en geïmproviseerd tot stand was gekomen. De voorgelegde tekst bevatte onder andere de volgende informatie: 'Bij het doornemen van de vragenlijsten die jullie zojuist hebben ingevuld, zagen wij dat er voldoende belangstelling was voor een bijeenkomst die speciaal gewijd is aan het onderwerp "osmose en het menselijk lichaam". Daarom zal die bijeenkomst worden gehouden op 4 juni aanstaande. Je kunt je met dit papiertje opgeven.' De proefpersoon moest daarbij aanduiden of zij of hij wel of niet deel zou nemen aan de bijeenkomst en haar of zijn handtekening zetten.

In navolging van Maehr (1976) is geprobeerd de attitude- en de gedragsmeting te combineren tot een Guttmanschaal van vier items. De items van een Guttmanschaal zijn meestal gerangschikt van makkelijk naar moeilijk of van een pro- naar een contra-attitude. De veronderstelling daarbij is dat als proefpersonen eenmaal een item fout of negatief beantwoorden ze alle volgende items ook fout of negatief zullen beantwoorden. Een schaal die een dergelijk antwoordpatroon oplevert, is een Guttmanschaal (Van der Ven, 1977). De Guttmanschaal die verondersteld wordt interesse in nadere

informatie over osmose te meten is zo geconstrueerd dat de items verlopen van een vage algemene interesse in osmose (item 1) tot de toezegging daadwerkelijk een bijeenkomst bij te wonen (gedragsmeting). Scalogramanalyse van de data leverde het volgende resultaat op: als item 2 uit de schaal verwijderd wordt, vormen item 1, item 3 en de gedragsmeting samen een (één-dimensionele) Guttman-schaal. De reproduceerbaarheidscoëfficiënt van deze schaal is 0,85 en de schaalbaarheidscoëfficiënt is gelijk aan 0,56. Daarmee voldoet de schaal aan de bij deze analyse gestelde minimumcriteria. In het vervolg zal deze variabele gemakshalve interesse-2 genoemd worden, om hem te onderscheiden van de interesse-vraag die ook in Experiment 1 gebruikt werd en die interesse-1 genoemd zal worden.

Ten slotte werden nog enkele *biografische gegevens* gevraagd als leeftijd, geslacht en eindcijfer voor biologie op de HAVO.

Procedure

De procedure kwam in grote lijnen overeen met die in Experiment 1. De experimentele conditie bestond uit 23 proefpersonen en de controleconditie uit 25 proefpersonen. De experimentele groep werd opgedeeld in drie kleinere groepen. Elk van deze groepen kreeg aselekt een ervaren tutor toegewezen. Tutor en groep werkten op de zelfde wijze als beschreven voor Experiment 1. Alle groepen hadden minder dan 10 minuten nodig voor de analyse. Na afloop van de probleemanalyse beantwoordden experimentele en controlegroep interesse-1. Vervolgens kregen ze de tekst voorgelegd, met de opdracht deze te bestuderen. De bestuderingstijd was 15 minuten. Na afloop daarvan kregen beide groepen de drie attitudeitems en de twee toetsen voorgelegd. Ten slotte werd de gedragsmeting gedaan.

Resultaten en discussie

Experiment 2 gaf de mogelijkheid na te gaan in hoeverre de resultaten van Experiment 1 in een andere populatie gerepliceerd konden worden. Dat wil zeggen dat nagegaan werd of de experimentele groep na afloop van de probleemanalyse een hogere intrinsieke interesse in osmose vertoonde dan de controlegroep. De proefpersonen die in Experiment 2 participeerden hadden in tegenstelling tot de proefpersonen in het eerste experiment enige praktische ervaring opgedaan met de methode van probleemanalyse. Een replicatie van de bevindingen uit Experiment 1 zou er dus op kunnen duiden dat het geconstateerde verschil niet terug te voeren is op een Hawthorne-effect.

In tabel 4 en 5 zijn de belangrijkste resultaten samengevat (zie volgende bladzijde).

Uit de in tabel 4 en 5 weergegeven resultaten kan worden afgeleid dat er als gevolg van de probleemanalyse nog wel verschillen in interesse tussen de groepen constateerbaar zijn: $F = 3,03$, $p < 0,09$. Die verschillen zijn

Tabel 4: Gemiddelden en standaarddeviaties van scores op interesse-1 voor Experiment 2.

	$\bar{X}$	SD	N
Probleemanalysegroep	3,35	0,88	23
Controlegroep	2,88	0,97	25
Totaal	3,10	0,95	48

Tabel 5: Variantieanalyse op interesse-1-scores voor Experiment 2.

Bron	Kwadraten-som	df	Kwadraten-gemiddelde	F
Experimentele manipulatie	2,62	1	2,62	3,03
Residu	39,86	46	0,87	
Totaal	42,48	47	0,90	

echter minder duidelijk dan die in Experiment 1 gevonden werden. In feite loopt het verschil met ongeveer de helft terug. Het is echter niet zo dat de probleemanalysegroep in Experiment 2 minder in osmose geïnteresseerd is dan die in het eerste experiment. Het verschil loopt terug omdat de controlegroep geïnteresseerder blijkt te zijn in het onderwerp dan die uit het eerste experiment. Dit resultaat zou misschien verklaard kunnen worden uit het feit dat proefpersonen uit Experiment 2 een wat biologischer getinte opleiding volgen dan die uit Experiment 1, en daarom spontaan al wat meer geïnteresseerd zijn in osmose. In een onderwijssituatie waarin de interesse voor een bepaald onderwerp 'spontaan' al redelijk hoog is, valt met interesseverhogende maatregelen uiteraard niet zoveel winst meer te behalen.

De tweede doelstelling van Experiment 2 was te verhelderen op welke wijzen intrinsieke interesse in osmose en studieprestaties met betrekking tot een tekst over dat onderwerp, met elkaar in verband gebracht kunnen worden. Om die vraag te kunnen beantwoorden zijn correlaties berekend tussen interesse, gemeten op twee momenten, namelijk voorafgaand aan de studietaak en daarna, en studieprestaties gemeten na bestudering van de tekst. Studieprestaties waren geoperationaliseerd als: herkenning van de tekst en transfer van het geleerde. In tabel 6 zijn de resultaten weergegeven.

Tabel 6: Correlaties tussen interesse in osmose en studieprestaties in Experiment 2.

	Interesse-1	Herkenning	Transfer	Interesse-2
Interesse-1	—			
Herkenning	0,32	—		
Transfer	0,24	0,41	—	
Interesse-2	0,16	0,29	0,18	—

Op het eerste gezicht zijn een aantal van de correlaties tussen interesse en prestaties vergelijkbaar met die welke in ander onderzoek op dit terrein gevonden worden: om en nabij de 0,30; niet bijzonder hoog, maar ook niet te negeren. Meer specifiek kan men zich de vraag stellen wat de betekenis van deze correlaties is in het licht van de in de inleiding tot Experiment 2 beschreven opvattingen over de relatie tussen intrinsieke motivatie en prestatie. Het is daarbij nuttig eerst eens de correlatie tussen interesse, gemeten vóór de studietaak, en prestatie, gemeten ná de studietaak, nader te bestuderen. Tussen interesse-1 en herkenning wordt een correlatie van $r_{ilh} = 0,32$ gevonden; tussen interesse-1 en transfer een correlatie van $r_{ilt} = 0,24$.¹ Beide correlaties verschillen significant van nul: $r_{ilh} = 0,32$, $p < 0,025$; $r_{ilt} = 0,24$, $p < 0,05$.

Dat zou erop kunnen duiden dat de cognitieve activiteitshypothese van Fass en Schumacher (1978) in de resultaten van Experiment 2 een bevestiging vindt. Fass en Schumacher (1978) veronderstelden immers dat hoog-gemotiveerden bij de bestudering van een tekst meer en andere cognitieve activiteit ontplooiën dan laag-gemotiveerden, en daardoor een betere prestatie leveren. Een dergelijke theorie voorspelt een correlatie tussen interesse gemeten voorafgaande aan een bepaalde studietaak, en prestaties op die studietaak; een correlatie die in het hier gerapporteerde materiaal gevonden lijkt te worden. Het is echter niet uitgesloten dat de gevonden correlatie tussen eerdere interesse en latere prestatie een artefact is van de gekozen onderzoeksopzet. Indien immers de probleemanalyse zowel een hoge motivatie induceert bij de experimentele groep als *onafhankelijk daarvan* ertoe leidt dat die groep een betere studieprestatie levert, dan zal niettemin een verband tussen motivatie en prestatie gevonden worden. Dit verband is echter spurieus omdat het het gevolg is van de experimentele manipulatie. De eenvoudigste methode om na te gaan of een correlatie tussen twee variabelen spurieus dan wel 'echt' is, is partiële correlatieanalyse. Daarbij wordt de correlatie tussen de twee betreffende variabelen berekend, terwijl een derde variabele, die ervan verdacht wordt de veroorzaker te zijn van het mogelijk spurieus verband, constant wordt gehouden. Deze derde variabele (in het geval van Experiment 2 het lidmaatschap van de experimentele dan wel van de controlegroep) wordt uitgepartialiseerd. De partiële correlatie tussen interesse-1 en herkenning blijkt gelijk te zijn aan 0,23, en die tussen interesse-1 en transfer 0,16. Deze partiële correlaties verschillen niet significant van nul: $r_{ilh.e} = 0,23$: $t = 1,59$, $p > 0,10$; $r_{ilt.e} = 0,16$: $t = 1,09$, $p > 0,20$.²

¹ r_{ilh} betekent: de correlatie tussen interesse-1 en herkenning. r_{ilt} betekent: de correlatie tussen interesse-1 en transfer.

² De statistische toetsing van deze partiële correlaties is niet gebeurd met behulp van de conventionele tabellen voor de toetsing van correlaties. Voor de toetsing van de gevonden partiële correlaties is gebruik gemaakt van de toetsingsgrootheid

$$t = \frac{r_p \sqrt{N - k - 2}}{\sqrt{(1 - r_p^2)}} \text{ die } t\text{-verdeeld is, waarbij: } \begin{array}{l} r_p = \text{de te toetsen partiële correlatie} \\ N = \text{aantal proefpersonen} \\ k = \text{aantal uitgepartialiseerde variabelen} \end{array}$$

De conclusie die uit dit resultaat getrokken zou kunnen worden is dat in het hier gerapporteerde experiment geen directe invloed van motivatie op studieprestatie kan worden aangetoond. Het resultaat is in overeenstemming met gegevens die door Bloom (1976) gerapporteerd worden — Blooms studenten vonden in een aantal studies een gemiddelde correlatie van $r = 0,06$ — en maakt de cognitieve-activiteitshypothese van Fass en Schumacher (1978) onwaarschijnlijk als verklaring voor het gevonden resultaat. Interesse in een bepaald onderwerp lijkt dus geen directe invloed op prestatie te hebben; maar hoe staat het met het omgekeerde? In hoeverre beïnvloedt prestatie de interesse in een bepaald onderwerp? In de inleiding is beargumenteerd dat interesse een bijproduct zou kunnen zijn van prestatie; dat personen die een goede studieprestatie leveren — competent zijn met betrekking tot een bepaald onderwerp — dáárom een grotere interesse gaan vertonen in dat onderwerp. Deze these werd empirisch ondersteund door Bloom (1976) en door resultaten gevonden in de grootscheepse IEA-studies (Husén, 1967; Carroll, 1975). In deze studies werden correlaties van gemiddeld genomen 0,30 gemeld tussen prestatie en gelijktijdig gemeten interesse.

Hoe staat het met de resultaten van het onderhavige experiment in relatie tot de opvatting dat motivatie een bijproduct van prestatie zou zijn? Om deze these te toetsen moeten de correlaties tussen interesse-2 en de prestatiematen bezien worden. Interesse-2 werd min of meer gelijktijdig met prestatie gemeten. De correlatie tussen herkenning en interesse is statistisch significant: $r = 0,29$; $p < 0,025$. Na uitpartialisering van de experimentele manipulatie wordt de partiële correlatie gelijk aan 0,21. Deze verschil niet significant van nul: $t = 1,44$, $p > 0,10$. De correlatie tussen transfer en interesse-2 is gelijk aan 0,18: $t = 1,23$, $p > 0,20$. Al deze resultaten geven aanleiding te veronderstellen dat interesse en prestatie beschouwd moeten worden als twee van elkaar onafhankelijke uitkomsten van een onderwijsleerproces.

Er zijn pogingen gedaan deze conclusie te toetsen aan nieuwe data. Daartoe zijn voor de gegevens verzameld in Experiment 1 correlaties berekend tussen interesse-1 en gelijktijdig daaraan gemeten prestaties met betrekking tot osmose. Die prestaties werden gemeten met behulp van dezelfde toetsen als gebruikt in Experiment 2. De correlaties waren gelijk aan $r_{ilh} = 0,29$, $p < 0,05$ en $r_{ilt} = 0,34$, $p < 0,025$.

Na uitpartialisering van het effect van de experimentele manipulatie waren de partiële correlaties $r_{ilh,e} = 0,19$; $t = 1,30$, $p > 0,10$; en $r_{ilt,e} = 0,22$; $t = 1,51$, $p > 0,10$.

Deze gegevens geven eveneens aanleiding te veronderstellen dat interesse en prestatie in feite onafhankelijke grootheden zijn die beschouwd moeten worden als eigenstandige produkten van onderwijs.

Er is ook nog gekeken naar het verband tussen het eindcijfer voor biologie als maat voor prestatie, en interesse in osmose.

In Experiment 1 was $r_{ilb} = 0,14$ en in Experiment 2 $r_{ilb} = -0,02$.

Beide correlaties verschillen niet significant van nul. De conclusie lijkt dus

gerechtvaardigd dat voor de gekozen onderzoeksopzet geldt dat interesse in een bepaald onderwerp en studieprestaties op dit onderwerp niets met elkaar te maken hebben.

Ten slotte werd nog onderzocht of de intrinsieke motivatie – waarvan geconstateerd was dat hij door de probleemanalyse verhoogd werd – verhoogd blijft, ook na het bestuderen van de tekst. Daartoe werd een variantieanalyse uitgevoerd op de resultaten van de interesse-2-meting, de gecombineerde Guttman-schaal. De resultaten daarvan zijn opgenomen in tabel 7 en 8.

Tabel 7: Gemiddelden en standaarddeviaties van scores op interesse-2 voor Experiment 2.

	$\bar{X}$	SD	N
Probleemanalysegroep	7,52	2,17	23
Controlegroep	8,32	1,41	25
Totaal	7,94	1,84	48

Tabel 8: Variantieanalyse op interesse-2-scores voor Experiment 2.

Bron	Kwadraten-som	df	Kwadraten-gemiddelde	F
Experimentele manipulatie	7,63	1	7,63	2,32
Residu	151,18	46	3,29	
Totaal	158,81	47	3,38	

De probleemanalyse heeft geen effect op interesse-2: $F = 2,32$, $p > 0,13$. Voor zover er sprake is van enige trend in de data kan men zelfs zeggen dat de experimentele groep nu wat minder geïnteresseerd lijkt in osmose dan de controlegroep. De eerder geconstateerde verhoogde intrinsieke motivatie als gevolg van de probleemanalyse blijkt in de loop van het leerproces verdwenen te zijn.

Algemene discussie

De resultaten van twee experimenten waarin intrinsieke motivatie in samenhang met studieprestatie onderzocht werd, kunnen als volgt worden samengevat. Onderwijsmethoden die leerlingen ertoe brengen actief met de leerstof om te gaan – en met name die methoden die zich richten op het expliciteren van onduidelijkheden in de ideeën die leerlingen er over een be-

paald onderwerp op na houden, zoals de controversemethode van Johnson en Johnson (1979), de hypothese-deductie-experiment-methode van Inagaki en Hatano (1977) en de methode van probleemanalyse (Schmidt, 1982a) induceren bij die leerlingen een intrinsiek motivationeel proces, dat zich uit in een verhoogde interesse in dat onderwerp en een verhoogde neiging activiteiten te ontplooiën die erop gericht zijn informatie over dat onderwerp te verzamelen. Dat fenomeen werd gevonden in Experiment 1. Proefpersonen die het bloedcelprobleem geanalyseerd hadden, bleken meer geïnteresseerd te zijn in nadere informatie over osmose (een resultaat dat in Experiment 2 gerepliceerd werd) en vaker geneigd zijn zich op te geven voor het bijwonen van een lezing over het onderwerp. De incongruentietheorie van Berlyne (1978) met betrekking tot intrinsieke motivatie lijkt een goede verklaring te geven voor dit effect: mensen streven naar een optimale congruentie tussen de werkelijkheid en hun eigen kennis van die werkelijkheid. Wordt de discrepantie tussen beide groter dan dat optimale niveau, bijvoorbeeld wanneer mensen met een probleem geconfronteerd worden dat met de aanwezige kennis niet bevredigend kan worden opgelost, dan komt er een intrinsiek gemotiveerd proces op gang — door Berlyne epistemic curiosity genoemd — dat de incongruentie door middel van informatietoevoer reduceert. Het feit dat er een reductie van de incongruentie optreedt onder invloed van de bestudering van probleemrelevante informatie, verklaart waarschijnlijk waarom in Experiment 2 na de tekstverwerkingsfase geen verschillen in intrinsieke motivatie meer geconstateerd kunnen worden tussen experimentele en controlegroep. Beide zijn blijkbaar onder invloed van de tekstbestudering teruggekeerd tot een optimaal niveau van (in)congruentie tussen wat men weet en wat men denkt er nog over te moeten weten.

De klaarblijkelijke fragiliteit van het proces van intrinsieke motivatie, zoals die naar voren komt in de hier besproken data, maakt het er niet eenvoudiger op de opvattingen te steunen van diegenen die menen dat activerende onderwijsmethoden zoals ontdekkend leren en de in dit artikel besproken methoden een *blijvend* effect hebben op de intrinsieke motivatie van de lerenden (Bruner, 1961; Maehr, 1976; Barrows en Tamblyn, 1980). Intrinsieke motivatie lijkt eerder een situationeel kenmerk te zijn dan een stabiel persoonskenmerk, lijkt eerder opgeroepen te worden door aspecten van het onderwijs dan door lerenden in de onderwijssituatie gebracht te worden. Het is natuurlijk niet uitgesloten dat een bepaalde onderwijsbenadering op de lange duur leidt tot een meer of minder hoge 'gegeneraliseerde' intrinsieke motivatie. Tot op heden is echter slechts weinig aangevoerd dat die veronderstelling zou kunnen staven. De gegevens die op dit moment beschikbaar zijn, leiden eerder tot de conclusie dat intrinsieke motivatie een zeer specifieke (dat wil zeggen op bepaalde leerstof betrokken) kennisbehoefte is, die door toevoer van informatie bevredigd wordt en daardoor verdwijnt.

De relatie tussen intrinsieke motivatie, uitgedrukt in de interesse die leer-

lingen voor een bepaald aspect van hun studie hebben, en hun prestaties op dat aspect, blijkt tamelijk gecompliceerd te zijn. Vier modellen werden besproken die ieder ondersteund werden door empirisch materiaal. Voor het model van Fass en Schumacher (1978), dat veronderstelt dat hoog-gemotiveerden leerstof gegeven een standaard hoeveelheid studietijd intensiever cognitief verwerken, en daardoor betere studieprestaties leveren, kon geen bevestiging worden gevonden. In Experiment 2 bleken interesse (gemeten voorafgaande aan de leertaak) en studieprestaties (gemeten na de leertaak) niet met elkaar samen te hangen. Ook de opvatting dat motivatie een affectieve consequentie van competentie zou zijn, zoals verwoord door Deci (1975) en Bloom (1976), vond geen bevestiging. Zowel in Experiment 1 als in Experiment 2 bleken studieprestatie en parallel gemeten interesse in hoge mate ongecorreleerd. Uit dit feit werd afgeleid dat intrinsieke motivatie en studieprestatie beschouwd zouden moeten worden als op zichzelf staande, van elkaar onafhankelijke uitkomsten van een onderwijsleerproces. Dat betekent dat mensen hogelijk geïnteresseerd kunnen zijn in een bepaald onderwerp zonder dat dat leidt tot, of gepaard gaat aan, goede studieresultaten op dat onderwerp. Terwijl er anderen kunnen zijn die hoog scoren op beide variabelen.

Bij dit alles moet wel uitdrukkelijk aangetekend worden dat deze conclusies getrokken worden op grond van experimenten waarin studietijd constant gehouden werd. Het blijft zeer wel mogelijk dat motivatie via beïnvloeding van de studietijd, toch invloed uitoefent op de prestatie (Carroll, 1963). Kersh (in Wittrock, 1966) bijvoorbeeld observeerde dat proefpersonen in zijn discovery-learning-experimenten de geleerde regels na het leren en vóór de prestatiemetingen spontaan vaker oefenden en daardoor waarschijnlijk betere prestaties leverden dan proefpersonen in de controlegroep. Als dat inderdaad zo is, dan bezitten experimenten waarin studietijd constant gehouden wordt waarschijnlijk een lage externe validiteit. Uit deze veronderstelling volgt als het ware onmiddellijk een suggestie voor verder onderzoek. Nagegaan zou moeten worden in hoeverre studietijd varieert als een functie van intrinsieke motivatie.

Nog een laatste opmerking. Onder invloed van de maatschappelijke consequenties die verbonden zijn aan slagen of mislukken op school, heeft kennis in onze samenleving een middelkarakter gekregen. Kennis is niet iets dat interessant is in zichzelf, of nuttig, omdat het je in staat stelt de wereld om je heen te begrijpen, maar een middel om aanzien en welstand te bereiken. Dat leidt ertoe dat lerenden enkel en alleen op grond van de aantrekkelijkheid van het einddoel van een studie bereid zijn jarenlang onderwijsprogramma's te ondergaan die zij ervaren als 'saai, taai en irrelevant' (zoals bijvoorbeeld medische studenten hun studie beschrijven: Bender, 1979). Dat leidt er ook toe dat veel mensen met afkeer terugdenken aan de school en het schoolse leren, en er later niet meer toe gebracht kunnen worden opnieuw te gaan studeren, ook als hun levenssituatie daarom vraagt. Dat alleen al maakt het doen toenemen van intrinsieke interesse in de leer-

stof als zelfstandige onderwijsdoelstelling een nastrevenswaardige zaak.

Literatuur

- Anderson, R.C. The notion of schemata and the educational enterprise: general discussion of the conference. In R.C. Anderson, R.J. Spiro en W.E. Montague (Eds.). *Schooling and the acquisition of knowledge*. Hillsdale: Lawrence Erlbaum, 1977.
- Ausubel, D.P. *Educational Psychology, a cognitive view*. New York: Holt, Rinehart en Winston, 1968.
- Barrows, H.S. en R.M. Tamblyn. *Problem based learning*. New York: Springer, 1980.
- Bender, W. Kritische momenten in de medische studie: enkele recente onderzoeksgegevens. *Medisch Contact*, 1979, 13, 402-403.
- Berlyne, D.E. The arousal and satiation of perceptual curiosity in the rat. *Journal of Comparative and Physiological Psychology*, 1955, 48, 238-246.
- Berlyne, D.E. *Structure and direction in thinking*. New York: Wiley, 1965.
- Berlyne, D.E. What next? Concluding summary. In H.I. Day, D.E. Berlyne en D.E. Hunt (Eds.). *Intrinsic motivation: a new direction in education*. Toronto: Holt, Rinehart and Winston, 1971.
- Berlyne, D.E. Curiosity and learning. *Motivation and Emotion*, 1978, 2, 97-175.
- Bloom, B.S. *Human characteristics and school learning*. New York: McGraw-Hill, 1976.
- Bloom, B.S., J.T. Hastings en G. Madaus. *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill, 1971.
- Bruner, J.S. The act of discovery. *Harvard Educational Review*, 1961, 31, 21-32.
- Carroll, J.B. A model of school learning. *Teachers College Record*, 1963, 64, 723-733.
- Carroll, J.B. *The teaching of French as a foreign language in seven countries: international studies in evaluation*. New York: Wiley, 1975.
- Craik, F.I. en R.S. Lockhart. Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 671-684.
- Day, H.I., D.E. Berlyne en D.E. Hunt. *Intrinsic motivation: a new direction in education*. Toronto: Holt, Rinehart and Winston, 1971.
- DeCharms, R. *Personal causation: the internal affective determinants of behavior*. New York: Academic Press, 1968.
- Deci, E.L. *Intrinsic motivation*. New York: Plenum Press, 1975.
- Deci, E.L. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 1971, 18, 105-115.
- Deci, E.L. en R.M. Ryan. The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Ed.). *Advances in experimental social psychology*, vol 13. New York: Academic Press, 1980.
- Fass, W. en G.M. Schumacher. Effects of motivation, subject activity and readability on the retention of prose materials. *Journal of Educational Psychology*, 1978, 70, 803-807.
- Ferguson, G.A. *Statistical analysis in psychology and education*. New York: McGraw-Hill, 1966.
- Fishbein, M. en I. Ajzen. Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review*, 1974, 81, 59-74.
- Fraser, C.E. *The case method of instruction*. New York: McGraw-Hill, 1931.
- Hebb, D.O. Drives and the conceptual nervous system. *Psychological Review*, 1955, 62, 243-254.
- Hunt, J. McV. Motivation inherent in information processing and action. In O.J. Harvey (Ed.). *Motivation and social interaction*. New York: Ronald, 1963.

- Hunt, J. McV. Intrinsic motivation: information and circumstances. In H.M. Schroder en P. Suedfeld (Eds.). *Personality theory and information processing*. New York: Ronald, 1971.
- Husén, T. (Ed.). *International study of achievement in mathematics: a comparison of twelve countries*. New York: Wiley, 1967.
- Inagaki, K. en G. Hatano. Amplification of cognitive motivation and its effects on epistemic observation. *American Educational Research Journal*, 1977, 14, 485-491.
- Johnson, D.W. en R.T. Johnson. Conflict in the classroom: controversy and learning. *Review of Educational Research*, 1979, 49, 51-69.
- Kagan, J. Motives and development. *Journal of Personality and Social Psychology*, 1972, 22, 51-66.
- Klare, G.R. A second look at the validity of readability formulas. *Journal of Reading Behavior*, 1976, 7, 129-152.
- Lepper, M.R. en D. Greene. Turning play into work: effects of adults surveillance and extrinsic rewards on children's intrinsic motivation. *Journal of Personality and Social Psychology*, 1975, 31, 479-486.
- Lepper, M.R. en D. Greene. Turning play into work: effects of adult surveillance and extrinsic reward. *Journal of Personality and Social Psychology*, 1973, 28, 129-137.
- Lowry, N. en D.W. Johnson. Effects of controversy on epistemic curiosity, achievement and attitudes. *The Journal of Social Psychology*, 1981, 115, 31-43.
- Maehr, M.L. Continuing motivation: an analysis of a seldom considered educational outcome. *Review of Educational Research*, 1976, 46, 443-462.
- McReynolds, P. The three faces of cognitive motivation. In H.I. Day, D.E. Berlyne en D.E. Hunt (Eds.). *Intrinsic motivation: A new direction in education*. Toronto: Holt, Rinehart and Winston, 1971.
- Miller, G.A., E. Galanter en K.H. Pribram. *Plans and the structure of behavior*. New York: Holt, 1960.
- Neame, R.L.B. How to: construct a problem-based course. *Medical Teacher*, 1981, 3, 94-99.
- Reder, L.M. The role of elaboration in the comprehension and retention of prose: a critical review. *Review of Educational Research*, 1980, 5, 5-53.
- Sarafino, E.P. en P.A. DiMattia. Does grading undermine intrinsic interest in a college course? *Journal of Educational Psychology*, 1978, 70, 916-921.
- Schmidt, H.G. Leren met problemen: een inleiding in probleemgestuurd onderwijs. In A.G. Vroon (Red.). *Handboek voor de onderwijspraktijk*. Deventer: Van Loghum Slaterus, 1979.
- Schmidt, H.G. Problem-based learning: rationale and description. *Medical Education*, 1982a, in druk.
- Schmidt, H.G. Activation and restructuring of prior knowledge and their effects on text processing. In A. Flammer en W. Kintsch (Eds.). *Discourse processing*. Amsterdam: North Holland, 1982b, in druk.
- Schmidt, H.G. en P.A.J. Bouhuijs. Effecten van structurering van patiëntenproblemen op leerresultaat en satisfactie van studenten. In *Congresboek Onderwijsresearchdagen*, Amsterdam: Vrije Universiteit, 1977.
- Schmidt, H.G. en P.A.J. Bouhuijs. *Onderwijs in taakgerichte groepen*. Utrecht: Het Spectrum, 1980.
- Smith, K., D.W. Johnson en R.T. Johnson. Can conflict be constructive? Controversy versus concurrence seeking in learning groups. *Journal of Educational Psychology*, 1981, 73, 651-663.
- Ven, A.H.G.S van der. *Inleiding in de schaaltheorie*. Deventer: Van Loghum Slaterus, 1977.
- Wittrock, M.C. The learning by discovery hypothesis. In L.S. Shulman en E.R. Keislar (Eds.). *Learning by discovery: a critical appraisal*. Chicago: Rand McNally, 1966.

EFFECTEN VAN STRUCTURERING VAN PATIENTENPROBLEMEN OP LEERRESULTAAT EN SATISFACTIE BIJ STUDENTEN

H.G. Schmidt en P.A.J. Bouhuijs

Een experiment werd uitgevoerd om na te gaan of, en zo mogelijk: hoe, structurering van de leersituatie in probleemgestuurd onderwijs leerproces en leerresultaten van studenten beïnvloedt.

Een groep proefpersonen werkte aan problemen zonder toegevoegde vragen; dit werd gedefinieerd als de ongestructureerde conditie.

Een andere groep kreeg problemen met toegevoegde vragen voorgelegd; dit was de gestructureerde conditie.

Beide groepen verschilden niet van elkaar in tevredenheid met de eigen werkwijze en die van de onderwijsgroep.

Proefpersonen uit de ongestructureerde conditie bleken echter beter dan die uit de gestructureerde conditie in staat te beoordelen in hoeverre begrippen uit een speciaal voor dat doel geconstrueerde begrippentoets samenhangen met de problemen die bestudeerd waren. Verondersteld werd dat dit verschil tot stand kwam omdat de ongestructureerde conditie significant meer verschillende bronnen raadpleegde bij het zoeken naar probleemrelevante informatie.

Op een enige dagen later afgenomen retentietoets waren geen verschillen meer constateerbaar.

De mate waarin, en de wijze waarop, richting moet worden gegeven aan het leerproces van studenten, vormt een van de centrale moeilijkheden waarvoor een constructeur van probleemgestuurd onderwijs zich gesteld ziet. In die onderwijsbenadering wordt er immers van uitgegaan dat studenten in staat zijn *zelf* leerdoelen te formuleren en die leerdoelen door middel van individuele studie te bereiken. Dat maakt het moeilijk op directe wijze invloed uit te oefenen op wat studenten leren en hoe ze dat doen.

De keuze van de problemen die als stimuli voor het leerproces dienen, en de vormgeving van die problemen, zijn dus van doorslaggevend belang, omdat zij in principe de enige onderwijsmiddelen vertegenwoordigen waarmee een docent richting kan geven aan het leren van studenten.

Deze stand van zaken is problematisch, omdat de docent een verantwoordelijkheid heeft ten opzichte van de samenleving. Hij moet beroepsbeoefenaars opleiden die voldoen aan zekere kwaliteitscriteria. Het is dus zijn taak te voorkomen dat studenten al te veel tijd verdoen met zaken die misschien wel aantrekkelijk zijn, maar in het licht van de te bereiken doel-

stellingen weinig nuttig. Nog afgezien van de vraag of sommige studenten de hen geboden vrijheid niet zullen misbruiken om de kantjes er vanaf te lopen.

Daarbij komt nog dat verwacht mag worden dat een zekere inefficiëntie in de hand gewerkt wordt als studenten zelf hun leerdoelen moeten formuleren en zelf de te bestuderen literatuur bij elkaar moeten zoeken; zelfs als van de veronderstelling wordt uitgegaan dat ze daar vanaf het begin van hun studie toe in staat zijn.

Aan de andere kant zijn er een aantal algemene onderwijskundige doelstellingen, die alleen maar bereikt lijken te kunnen worden als studenten een zekere vrijheid krijgen zélf te bepalen wat er geleerd wordt, hoe dat wordt gedaan, en hoe diepgaand. Bruner (1971) vat die doelstellingen samen onder het kopje 'leren te leren'. In die opvatting zou een belangrijk doel van onderwijs moeten zijn dat studenten leren hun tekorten aan kennis zelf te diagnosticeren en aan te vullen, dat ze leren voor dat doel de bibliotheek te gebruiken, dat ze leren prioriteiten te stellen, en waar mogelijk verschillende bronnen te raadplegen.

Deze doelstellingen worden essentieel geacht voor het universitaire onderwijs, omdat verondersteld kan worden dat eenmaal verworven kennis tamelijk snel veroudert en dat dus voor een optimaal professioneel functioneren permanente zelfstudie noodzakelijk is, ook na afsluiten van het formele onderwijs. In probleemgestuurd onderwijs wordt ervan uitgegaan dat deze vaardigheden het beste kunnen worden aangeleerd door de student zo vroeg mogelijk in hun studie ertoe te brengen verantwoordelijkheid te dragen voor het eigen leren (Neufeld en Barrows, 1974).

Vastgesteld moet overigens worden dat deze zelfwerkzaamheidshypothese nooit direct onderzocht is. Er zijn slechts enige indirecte aanwijzingen dat men met behulp van de probleemgestuurde benadering erin slaagt deze doelstelling te bereiken. Een aanwijzing vormt het onderzoek naar de besteding van studietijd door Maastrichtse medische studenten (Weggeman en Moen, 1982). Daaruit blijkt dat, hoewel het aantal contacturen in het medisch curriculum beperkt is en weinig directe aanwijzingen voor de zelfstudie gegeven worden, studenten in het eerste jaar gemiddeld 30,2 uur aan hun studie besteden en in het vierde jaar gemiddeld 34,5 uur. Weggeman en Moen concluderen dat deze gemiddelde studielast hoog is vergeleken met landelijke gemiddelden, en dat het feit dat docenten niet vaststellen wat studenten moeten doen, er blijkbaar niet toe leidt dat die studenten minder tijd aan hun studie besteden.

Andere bronnen (Schmidt en Moust, 1981; Gijssels en Schmidt, 1982, noot 1) bevestigen dat beeld in grote lijnen maar geven aan dat eerstejaarsstudenten, voordat ze in staat zijn zelf hun studie te plannen, een periode van (soms intense) onzekerheid moeten doormaken over wat er van hen verwacht wordt. Deze onzekerheid komt ook tot uitdrukking in de hoeveelheid tijd besteed aan studie.

Samenvattend kan men zeggen dat de problemenconstructeur in probleemgestuurd onderwijs voor de moeilijke taak staat te varen tussen de Scylla

van de verantwoordelijkheid van de docent en de Charybdis van de vrijheid van de student.

In de praktijk van probleemgestuurd onderwijs in Maastricht en elders heeft dit dilemma aanleiding gegeven tot pogingen van docenten om *op indirecte wijze* invloed uit te oefenen op het leergedrag van hun studenten.

- Sinds 1977 zijn in de meeste blokboeken¹ doelstellingen opgenomen die in termen van leerinhouden globaal aangeven waarmee studenten geacht worden zich bezig te houden.
- In sommige blokken worden tamelijk specifieke literatuursuggesties gedaan. (Sommige blokboeken gaan zelfs vergezeld van een verzameling teksten in de vorm van een 'leesboek'.)
- Sommige blokboeken bevatten lijsten met relevante begrippen waaraan de student kan afchecken of hij zich met de juiste onderwerpen beziggehouden heeft.
- Sommige blokboeken bevatten een verzameling meerkeuzevragen, waarmee de student zijn eigen kennis kan testen (en zo nodig aanvullen).
- Soms organiseert de verantwoordelijke planningsgroep (= groep docenten die samen een blokboek construeren) bijeenkomsten waarop inhoudsdeskundigen (= specialisten op bepaalde vakgebieden) college geven over een of ander onderwerp.
- Een hele tijd is in Maastricht een 'tutorboek' in zwang geweest. Dit boek bevatte een handleiding bij het blokboek waarmee de begeleider zijn onderwijsgroep in het juiste spoor kon houden.
- Meer recent zijn er pogingen gedaan kennisgebieden in kaart te brengen door middel van een zogenaamde 'onderwerpenboom' die vervolgens vertaald werd in een verzameling samenhangende problemen die aan studenten werd voorgelegd (Schmidt, Majoor en Van den Eertwegh, 1980, noot 2).
- In Australië, aan de medische faculteit in Newcastle, geeft men structuur aan het leerproces door klinische problemen op te delen in kleinere eenheden ('cuing material'), die vervolgens successievelijk aan studenten worden aangeboden (Neame, 1980).

(Het zou overigens een verkeerde voorstelling van zaken zijn als uit het voorgaande de indruk ontstaat dat alleen docenten de neiging hebben meer structuur te geven aan het leerproces van studenten. Er zijn zeker ook studenten die daaraan behoefte hebben, en daarvan ook blijk geven.)

In deze bijdrage wordt een andere veelgebruikte structureringsvorm onderzocht; namelijk *het toevoegen van vragen aan een probleem*.

Uit de onderwijskundige literatuur is bekend dat vragen voorafgaande aan de bestudering van een tekst een sturende werking uitoefenen op de wijze waarop die tekst verwerkt wordt. Aan vraagrelevante informatie wordt veel aandacht gegeven; aan informatie die dat niet is, wordt weinig aandacht gegeven (Rothkopf, 1976; Rothkopf en Bisbicos, 1967).

Het is echter niet zeker dat vragen op dezelfde wijze functioneren in een

¹ Een blokboek bevat de problemen waaraan studenten gedurende een onderwijsblok (6 weken) verondersteld worden te werken.

situatie waarin studenten de te bestuderen teksten ook zelf moeten uitzoeken.

In het hier gerapporteerde experiment werkten proefpersonen aan problemen met of zonder toegevoegde vragen. Onderzocht werd wat de invloed van die toegevoegde vragen was op leerproces en leerresultaat.

Het meten van leerresultaten van probleemgestuurd onderwijs stelt de onderzoeker overigens voor bijzondere problemen.

In probleemgestuurd onderwijs worden studenten, zoals gezegd, in belangrijke mate vrijgelaten in de keuze van de te bestuderen leerstof. Wat bestudeerd wordt, is afhankelijk van de uit de discussie over het probleem voortvloeiende leerdoelen, en verder van wat in de bibliotheek beschikbaar is, persoonlijke voorkeuren van studenten voor bepaalde studieboeken, etc. Dat leidt ertoe dat er min of meer toevallige verschillen tussen studenten kunnen optreden, waar het de aard van hun kennis betreft. De aard van hun kennis is namelijk in hoge mate afhankelijk van het toevallige boek dat studenten in handen gehad hebben. Dat maakt het doen van onderwijskundig onderzoek niet eenvoudig, omdat een belangrijk hulpmiddel daarbij, de studietoets, waarschijnlijk niet in staat is optimaal de effecten van de experimentele manipulatie te meten. Zo'n studietoets bestaat meestal uit items die ontleend zijn aan één bepaalde tekst, wat onder andere in de woordkeuze tot uitdrukking komt. Studenten die die tekst toevallig niet bestudeerd hebben (maar wel een even relevante andere tekst) zijn dan vaak enigszins in het nadeel vergeleken met studenten die dat wel gedaan hebben. De vrijheid zelf teksten uit te mogen zoeken is dus een bron van foutenvariantie die experimenteren bemoeilijkt. In het hieronder gerapporteerde experiment is daarom voor het meten van studieprestaties gebruik gemaakt van een benadering die minder dan de studietoets een beroep doet op kennis van specifieke formuleringen. (Een 'klassieke' studietoets werd echter ook gebruikt.)

Methode

Proefpersonen

Proefpersonen waren 47 eerstejaarsstudenten in de geneeskunde aan de Rijksuniversiteit Limburg. Het experiment vond plaats gedurende het reguliere onderwijs.

Materiaal

Het materiaal bestond uit twee problemen, zogenaamde patiëntencasus, twee verzamelingen bij deze problemen behorende vragen, twee satisfactievragenlijsten, een studietijdsbestedingsformulier, twee begrippentoetsen en twee retentietoetsen.

De *patiëntencasus* bestond uit de beschrijving van een aantal verschijnselen waarvoor de studenten een verklaring moesten zoeken in termen van processen, structuren of mechanismen die voor die verschijnselen verantwoordelijk gesteld kunnen worden. Ter illustratie is in tabel 1 de tekst van één van de problemen opgenomen.

Tabel 1: Voorbeeld van een patiëntencasus

De patiënt is een man van 53 jaar, gehuwd en chauffeur van een stadsbus. Er is bekend dat hij een gevoelige maag heeft en chronisch recidiverende bovenbuiksklachten, waarvoor blijkens recent internistisch onderzoek (op verzoek van de patiënt) geen somatisch substraat werd gevonden. Op gezette tijden is hij neurotisch depressief gestemd. Sinds enkele maanden heeft hij pijn in de linker elleboog. De pijn is langzaam toegenomen, waardoor hij momenteel arbeidsongeschikt is voor zijn werk als chauffeur. In de anamnese is geen aanwijzing voor een trauma te vinden. Na rust verminderen de klachten; overigens heeft hij 's avonds meer last dan 's morgens. Het buigen en draaien van de elleboog tegen weerstand in provoceert de bewegingsbeperking (actief zowel als passief). Omschreven drukpijn t.p.v. de epicondylus humeri lateralis links.

Vragen die wel of niet toegevoegd waren, vormden de experimentele manipulatie. In tabel 2 is de verzameling vragen horend bij de patiëntencasus in tabel 1 afgedrukt.

Tabel 2: Vragen horend bij de patiëntencasus in tabel 1

1. Welke spieren hebben hun origo of hun insertie vanuit de epicondylus humeri lateralis?
2. Bij welke bewegingen komt er spanning op de epicondylus humeri lateralis? (Flexie, extensie, pronatie of supinatie?)
3. Wat is een bursitis olecrani?
4. Waardoor ontstaat de pijn bij een tenniselleboog?
5. Welke behandelingswijzen zijn er mogelijk bij een tenniselleboog?
6. In hoeverre kan deze afwijking ontstaan zijn door de beroepsuitoefening?

Met behulp van twee *satisfactievragenlijsten* werd geprobeerd motivationele dimensies van het leerproces te meten. De eerste vragenlijst vroeg proefpersonen de werkwijze van de onderwijsgroep waartoe zij behoorden, te karakteriseren met evaluatieve items als oppervlakkig—diepgaand, efficiënt—inefficiënt en moeizaam—gesmeerd.

De tweede vragenlijst vroeg proefpersonen hun eigen werkwijze in de periode tussen bijeenkomsten van de onderwijsgroep te karakteriseren. Daarbij werden dezelfde 11 items voorgelegd.

Beantwoording vond plaats op een vijf-puntsschaal. Na verwijdering van een item hadden beide vragenlijsten een alfa-betrouwbaarheid van 0,92.

Op het *tijdsbestedingsformulier* konden proefpersonen aangeven welke literatuur bestudeerd was en hoeveel (netto)studietijd daaraan besteed was. Twee *begrippentoetsen* werden samengesteld uit begrippen die in een groot aantal verschillende studieteksten verzameld waren en die meer of minder relevant konden zijn ten aanzien van de processen die aan de problemen ten grondslag lagen. Proefpersonen moesten voor elk van deze begrippen beoordelen in hoeverre ze aan de bestudeerde problemen gerelateerd waren.

Deze beoordeling vond plaats op een drie-puntsschaal.

Hun oordeel werd vervolgens vergeleken met het oordeel van deskundigen. Per item werd de afstand tussen beide oordelen bepaald. De som van deze afstanden was de score die de student op deze begrippentoets haalde.

Voor elk van de problemen was een begrippentoets geconstrueerd bestaande uit 12 items.

Twee studietoetsen fungeerden als *retentietoetsen*. Ze bestonden uit 10 en 11 juist-onjuist-vragen respectievelijk, en vormden de operationalisatie van wat de docent met het probleem in termen van te bereiken leerdoelen voor ogen stond. De goed-min-fout-score werd gecorrigeerd voor het aantal items en getransformeerd in een percentage-score. (Proefpersonen waren in de gelegenheid in geval van niet-weten van het antwoord een vraagteken te omcirkelen.)

Procedure

De patiëntencasus waren tussen de andere problemen in het blokboek opgenomen. Er bestonden dus twee versies van het blokboek. De helft van de proefpersonen (drie onderwijsgroepen) had een blokboek waarin probleem 1 zonder vragen was opgenomen en probleem 2 met vragen, terwijl de andere helft een blokboek had waarin probleem 1 met vragen stond afgedrukt en probleem 2 zonder vragen. De proefpersonen waren aselekt aan de condities van het experiment toegedeeld. Bij de nabespreking van het experiment met studenten bleek dat men er zich niet van bewust was geweest dat andere onderwijsgroepen aan op andere wijze vormgegeven problemen gewerkt hadden. De tutores waren geïnstrueerd zo min mogelijk invloed uit te oefenen op de wijze waarop de groep de problemen aanpakte. Een deel van een onderwijsgroepsbijeenkomst werd besteed aan discussie over het probleem. Vervolgens had iedere student iets meer dan twee dagen beschikbaar voor individuele studie. Tijdens de volgende bijeenkomst werden de satisfactie-vragenlijsten en de begrippentoets ingevuld. Aan 12 studenten (2 per onderwijsgroep; dat wil zeggen: 6 per conditie van het experiment) werd daarnaast gevraagd bij te houden welke literatuur bestudeerd werd en hoeveel tijd daaraan besteed werd. Deze extra activiteit had geen bijzondere invloed op hun studieprestaties.

Deze procedure werd nog eens herhaald bij het werken aan het tweede probleem.

De retentietoetsen werden na afloop van de onderwijseenheid afgenomen; dat wil zeggen 9 dagen na de eerste meting en 4 dagen na de tweede meting. De gekozen proefopzet kan beschouwd worden als een experiment-met-replicatie. Die replicatie was op grond van eisen die door het onderwijs gesteld werden noodzakelijk. De data zijn over beide problemen heen gepoold. Daarmee werd de representativiteit van de resultaten en het onderscheidingsvermogen van de statistische toetsen vergroot (het aantal proefpersonen was nu namelijk 92). In het hierna volgende is dan ook verder sprake van één (samengestelde) begrippentoets en één (samengestelde) retentietoets.

Resultaten en discussie

In tabel 3 en 4 zijn de resultaten op beide satisfactievragenlijsten samengevat.

*Tabel 3: Gemiddelde Satisfactie-Eigen-Werkwijze(SEW)-scores
(met standaarddeviaties)*

	$\bar{X}$	SD	N
Zonder vragen	24,34	2,37	45
Met vragen	25,04	2,59	47
Totaal	24,70	2,47	92

NB: Een lagere score geeft een hogere satisfactie aan.

*Tabel 4: Gemiddelde Satisfactie-Werkwijze-Groep(SWG)-scores
(met standaarddeviaties)*

	$\bar{X}$	SD	N
Zonder vragen	26,28	5,99	45
Met vragen	26,27	2,50	47
Totaal	26,27	4,15	92

NB: Een lagere score geeft een hogere satisfactie aan.

De verschillen tussen de gemiddelden op beide variabelen werden statistisch getoetst: SEW: $F(1,91) = 1,83$; $p > 0,10$; SWG: $F(1,91) = 0,00$; $p > 0,10$. De gemiddelden verschillen niet van elkaar. Men kan dus niet concluderen dat een van beide condities tot een hogere satisfactie leidt.

De correlaties tussen beide variabelen en prestaties verschillen geen van allen significant van nul. De hoogste correlatie is die tussen SEW en de resultaten op de begrippentoets: $r = 0,12$. Dat betekent dat het oordeel over de eigen werkwijze en die van de groep niet gerelateerd is aan prestaties.

Op SWG traden overigens wel significante verschillen op tussen groepen ($p < 0,01$); en op SEW traden verschillen op tussen de eerste en de tweede ronde ($p < 0,05$). Deze effecten duiden erop dat de betreffende meetinstrumenten discrimineren op dimensies die voor die instrumenten relevant zijn en dat niet doen op dimensies die dat niet zijn. De experimentele manipulatie wél versus geen vragen is blijkbaar een dimensie die geen invloed uitoefent op gevoelens van tevredenheid of ontevredenheid.

Tabel 5: Gemiddelde scores op de begrippentoets (met standaarddeviaties)

	$\bar{X}$	SD	N
Zonder vragen	4,94	2,18	45
Met vragen	5,95	1,53	47
Totaal	5,46	2,05	92

NB: Een lagere score geeft meer overeenstemming met het expertoordeel aan.

De gemiddelde prestaties op de begrippentoets verschillen significant van elkaar: $F(1,91) = 6,56$; $p > 0,05$. Dat wil zeggen dat proefpersonen die aan problemen zonder toegevoegde vragen gewerkt hebben, blijkbaar beter in staat zijn te beoordelen in hoeverre een begrip relevant is ten aanzien van een bepaald probleem dan proefpersonen die wel vragen gehad hebben. Men kan zich afvragen hoe dat komt. In de inleiding is al geconstateerd dat vragen blijkens de literatuur een sturende functie vervullen bij de verwerking van informatie. Informatie die relevant is ten aanzien van een vraag wordt gezocht en verwerkt; andere informatie wordt genegeerd. Dat kan ertoe leiden dat het zoekproces naar informatie zoals dat in de hier onderzochte benadering van studenten gevraagd wordt, zich beperkt tot het vinden van antwoorden op gestelde vragen en daarmee stopt. De proefpersonen in de ongestructureerde conditie daarentegen, hebben geen extern criterium beschikbaar dat hen zegt wanneer voldoende informatie verzameld is. Men zou kunnen veronderstellen dat deze proefpersonen een betere prestatie leveren omdat zij langer doorgaan met zoeken naar informatie en meer teksten bestuderen. Deze hypothese kan getest worden aan de hand van de tijdsbestedingsgegevens van 24 proefpersonen (12 per probleem). In tabel 6 zijn deze data samengevat.

Tabel 6: Gemiddelde studietijd, in minuten, en gemiddelde aantal geraadpleegde teksten (met standaarddeviaties)

	Studietijd	Geraadpleegde teksten	N
Zonder vragen	246,67 (94,01)	7,08 (2,15)	12
Met vragen	219,17 (118,22)	4,42 (1,51)	12
Totaal	232,92 (105,40)	5,75 (2,27)	24

De gemiddelde bestede studietijden verschillen niet significant van elkaar: $F(1,23) = 0,40$; $p > 0,10$.

Met betrekking tot het gemiddeld aantal geraadpleegde teksten is er een duidelijk verschil constateerbaar: $F(1,23) = 12,30$; $p < 0,01$. De proefper-

sonen in de ongestructureerde conditie raadplegen significant meer teksten (boeken, artikelen) dan de proefpersonen in de gestructureerde conditie. Het lijkt er dus op dat verschillen in prestatie op de begrippentoets verklaarbaar zijn uit het feit dat proefpersonen die aan een probleem zonder toegevoegde vragen werken meer teksten naslaan en die teksten ieder afzonderlijk minder lang bestuderen. Proefpersonen uit de gestructureerde conditie hebben blijkbaar aan het doornemen van een paar teksten voldoende voor het vinden van antwoorden op de toegevoegde vragen.

Het verschil dat op de begrippentoets is waargenomen, kan in de retentiescores, verzameld enige dagen na het experiment, niet teruggevonden worden. In tabel 7 zijn de gegevens samengevat.

Tabel 7: Gemiddelde retentiescores (met standaarddeviaties)

	$\bar{X}$	SD	N
Zonder vragen	70,99	12,03	47
Met vragen	75,86	14,24	47
Totaal	73,43	14,04	94

$F(1,93) = 3,21$; $p > 0,05$. Voor zover er een trend in de data zit, is die in het voordeel van de proefpersonen uit de gestructureerde conditie. Misschien te wijten aan het feit dat degenen die de toegevoegde vragen schreven, ook de retentietoetsitems construeerden.

Het is merkwaardig dat het effect dat op de begrippentoets en in de aantallen geraadpleegde teksten zo duidelijk constateerbaar was, in de retentietoets niet terugkomt. Misschien is de invloed van structurering van patiëntenproblemen door middel van toegevoegde vragen zo subtiel dat mogelijke effecten met behulp van meerkeuzetoetsen niet gemeten kunnen worden. Het is niet uitgesloten dat dat inderdaad het geval is. Uit onderzoek blijkt dat zelfs tamelijk sterke experimentele manipulaties niet in resultaten op meerkeuzetoetsen kunnen worden opgemerkt (Alba, Alexander, Hasher en Caniglia, 1981; Reder en Anderson, 1980; Willems en Wardenaar, 1982, noot 3). Waarschijnlijk komt dat omdat items in meerkeuzetoetsen zoveel cues bevatten dat zij nauwelijks een beroep doen op strategieën die nodig zijn om informatie in het lange-termijn-geheugen op te zoeken en te reproduceren.

Conclusie

Terugkerend naar de inleiding tot dit artikel mag men concluderen dat de data die in dit experiment verzameld zijn, in ieder geval geen aanleiding kunnen zijn voor zorg van de kant van docenten dat studenten die zelf

hun leerdoelen formuleren slechtere studieprestaties leveren dan studenten in een meer vóórgestructureerde leeromgeving. Men zou zelfs kunnen zeggen dat er aanwijzingen zijn voor het tegendeel; zij het dat die aanwijzingen niet eenduidig zijn. Proefpersonen in de ongestructureerde conditie leverden een betere prestatie op de begrippentoets en bestudeerden meer verschillende teksten. Zij besteedden echter niet significant meer tijd aan de onderwerpen en leverden geen betere prestatie op een retentietoets.

Literatuur

- Alba, J.W., S.G. Alexander, L. Hasher en K. Caniglia. The role of context in the encoding of information. *Journal of Experimental Psychology: Human Learning and Memory*, 1981, 7, 283-292.
- Bruner, J.S. *Toward a theory of instruction*. Londen: Oxford University Press, 1971.
- Neame, R.L.B. Innovation in medical education: the Newcastle curriculum. In *Proceedings of the Seminar on the applicability of problem-based, community-oriented medical education in Egypt*. Alexandria, 1980.
- Neufeld, V.R. en H.S. Barrows. The 'McMaster philosophy': an approach to medical education. *Journal of Medical Education*, 1974, 49, 1040-1050.
- Reder, L.H. en J.R. Anderson. A comparison of texts and their summaries: memorial consequences. *Journal of Verbal Learning and Verbal Behavior*, 1980, 19, 121-134.
- Rothkopf, E.Z. Writing to teach and reading to learn: a perspective on the psychology of written instruction. In N.L. Gage (Ed.). *The psychology of teaching methods*. Chicago: National Society for the Study of Education, 1976.
- Rothkopf, E.Z. en E.E. Bisbicos. Selective facilitative effects of interspersed questions on learning from written materials. *Journal of Educational Psychology*, 1967, 58, 56-61.
- Schmidt, H.G. en J.H.C. Moust. Studiebeleving van Maastrichtse medische studenten. *Medisch Contact*, 1981, 49, 1515-1518.
- Weggeman, M. en J. Moen. Een onderzoek naar de studiebelasting van studenten aan de faculteit der geneeskunde. Rijksuniversiteit Limburg. In H.G. Schmidt (Red.). *Probleemgestuurd onderwijs*. Stichting voor Onderzoek van het Onderwijs. Harlingen: Flevodruk, 1982.

Verwijzingsnoten

1. Gijsselaers, W. en H. Schmidt. *Resultaten programmaevaluatie 1981-1982*. Maastricht: Capaciteitsgroep Onderwijsontwikkeling en Onderwijsresearch, 1982.
2. Schmidt, H., G. Majoor en E. van den Eertwegh. *Studeren aan de Rijksuniversiteit Limburg, blokboek 1.1*. Maastricht, 1980.
3. Willems, J. en E. Wardenaar. *Studietaken als instructiemiddel: een explorerend onderzoek*. Nijmegen: Instituut voor Onderzoek van het Wetenschappelijk Onderwijs, 1982.

DE RELATIE TUSSEN STUDIEJAAR EN TENTAMENRESULTAAT

H.G. Schmidt, M.L. de Volder en W.H. Gijssels

Een positieve correlatie tussen studiejaar en studieprestatie werd gevonden in gegevens, betrekking hebbend op de studieresultaten van studenten aan de medische faculteit in Maastricht. In dit artikel worden enige pogingen gedaan verklaringen te vinden voor dit verschijnsel. Getracht wordt aanneemelijk te maken dat de aan- of afwezigheid van relevante voorkennis bij studenten, en hun studiemotivatie, een rol spelen bij het ontstaan van dit verschijnsel.

Crombag (1981) rapporteert een merkwaardig fenomeen. Aan de Faculteit der Rechtsgeleerdheid van de Leidse universiteit werd enige jaren geleden een nieuw curriculum ingevoerd waarbij verschillende cursussen naar eerdere studiejaar verplaatst werden. Zo werd een tweedejaarscursus Staatskunde naar de propedeuse verplaatst en een cursus Burgerlijk Recht van het derde naar het tweede jaar. Crombag constateerde dat een dergelijke — op het eerste gezicht eenvoudige — verschuiving aanzienlijke consequenties heeft voor het percentage geslaagden bij zo'n cursus. Het percentage geslaagden daalde met gemiddeld genomen ongeveer 14 procent na verplaatsing van een cursus naar een eerder studiejaar, hoewel onderwijs, tentamens en toegepaste zak-slaag-normen ongewijzigd gebleven waren.

Crombag zelf geeft twee verklaringen voor het feit dat ouderejaarsstudenten op dezelfde tentamens betere prestaties leveren dan jongerejaarsstudenten. Hij verwijst ten eerste naar de mogelijkheid dat oudere studenten meer studeerervaring hebben, die hen helpt meer van hetzelfde onderwerp mee te nemen dan jongere studenten. Met andere woorden: de *studievaardigheden* zijn bij ouderejaarsstudenten beter ontwikkeld.

De tweede mogelijke verklaring die hij geeft, heeft te maken met de *selectie* die in onderwijsprogramma's plaatsvindt. Deze selectie leidt ertoe dat alleen de goede studenten overblijven. Het ligt voor de hand dat een cohort waaruit de slechte studenten verdwenen zijn, betere prestaties levert dan een cohort waaruit de slechten nog niet verdwenen zijn. De door Crombag gerapporteerde data geven voor deze verklaring ook wel enige evidentie. De cursus Staatskunde, verplaatst van het tweede naar het eerste jaar, ziet het percentage geslaagden met maar liefst 23% dalen. Met betrekking tot cursussen, verplaatst van het derde naar het tweede jaar, worden veel lagere dalingspercentages gerapporteerd; gemiddeld genomen ongeveer 12%. Dit verschil zou verklaarbaar kunnen zijn uit het feit dat het eerste jaar veel meer afvallers kent dan het tweede jaar.

Het doel van deze bijdrage is aannemelijk te maken dat er met de relatie tussen studiejaar en tentamenresultaat meer aan de hand zou kunnen zijn dan een spurieuze correlatie als gevolg van selectie. De Volder en Schmidt (1982) vonden namelijk — toevalligerwijze; het was niet het doel van dat onderzoek — een significante correlatie van 0,35 tussen studiejaar en tentamenresultaat. Dat resultaat is nauwelijks terug te voeren op selectie van studenten, omdat het daarbij ging om data verkregen in het onderwijsprogramma van de medische faculteit in Maastricht. Dit programma kent slechts een klein percentage afvallers. Wijnen (1982) noemt een uitvalpercentage van 2% over een periode van zes jaar. Dit resultaat is des te verrassender omdat het in dit onderzoek ging om een vergelijking van prestaties over *inhoudelijk verschillende* tentamens. Blijkbaar leveren ouderejaarsstudenten, onafhankelijk van de toevallige thematiek en inhoud van een cursus, betere prestaties dan studenten in eerdere studiejaar.

De gegevens, gebruikt in het onderzoek van De Volder en Schmidt (1982), waren zogenaamde goed-min-fout-scores. Dat wil zeggen dat de analyse gedaan werd op studieresultaten die verkregen waren door voor iedere student het aantal foute antwoorden op een meerkeuzetoets af te trekken van het aantal goede antwoorden. Daar de student in de gelegenheid gesteld wordt te antwoorden met omcirkeling van een vraagteken wanneer hij een antwoord niet weet, wordt deze wijze van berekening van de totaalscore beschouwd als een methode om raden tegen te gaan.

Het is niet uitgesloten dat de gevonden relatie tussen tentamenresultaat en studiejaar een artefact is van de gevolgde scoringsprocedure. Het is bijvoorbeeld niet onmogelijk dat ouderejaars niet zozeer betere prestaties leveren (een hogere goed-score halen), maar beter in staat zijn fouten te vermijden. Dat zou betekenen dat het door De Volder en Schmidt gevonden verschijnsel eerder het gevolg is van een zich wijzigende antwoordstrategie dan van een of ander nog onopgehelderd proces.

In deze bijdrage wordt verslag gedaan van een poging tot replicatie en opheldering van het verschijnsel. Daartoe zijn gegevens verzameld in het academiejaar volgend op het academiejaar waaruit De Volder en Schmidt hun gegevens haalden. Het onderwijsprogramma had sinds het eerste onderzoek slechts op ondergeschikte punten veranderingen ondergaan.

Procedure

In het academiejaar 1980-1981 werden gegevens met betrekking tot resultaten op studietoetsen verzameld voor de eerste vier jaar van de medische studie aan de Rijksuniversiteit Limburg. In elk studiejaar wordt gedurende zes weken (een blok geheten) in kleine groepen aan een bepaald thema gewerkt. Thema's zijn bijvoorbeeld: atherosclerose, de oudere mens, moeheid, pijn op de borst, leefwijzen, bloedverlies.

Na afloop van elk blok wordt in het kader van de formatieve evaluatie de zogenaamde bloktoets aan studenten voorgelegd. Met uitzondering van dit

formatieve karakter, heeft de bloktoets alle eigenschappen van een klassieke studietoets. De leerstof die tijdens een bepaald programmaonderdeel aan bod gekomen is, wordt verwerkt in een toets bestaande uit 100 à 200 vragen met meervoudige antwoordkeuze (in dit geval: juist – onjuist – vraagteken). *Essentieel* aan een dergelijke programma-gebonden toets is ten eerste de zogenaamde relevantie van de toets: meten de vragen datgene wat men ermee wil meten: namelijk de behandelde leerstof; en ten tweede de moeilijkheid van de toets: zijn de vragen qua moeilijkheid afgestemd op het niveau van de studenten voor wie de toets bestemd is (Dousma en Horsten, 1980).

In totaal werden van het academiejaar 1980-1981 18 bloktoetsen in de analyse betrokken: 6 uit het eerste studiejaar, 5 uit het tweede jaar, 4 uit het derde jaar en 3 uit het vierde jaar. Het aantal bloktoetsen loopt per studiejaar iets terug omdat het aantal perioden dat studenten keuzeonderwijs volgen elk jaar toeneemt. Van één blok waren de gegevens niet beschikbaar. De mediaanbetrouwbaarheid van de toetsen was gelijk aan 0,82. Van niet alle toetsen waren betrouwbaarheidsgegevens beschikbaar. De aantallen studenten die de toetsen invulden, varieerden per keer in geringe mate.

Per toets werden de gemiddelde goed-score en de gemiddelde goed-min-fout-score getransformeerd in percentagescores om vergelijkingen tussen toetsen mogelijk te maken.

Resultaten en discussie

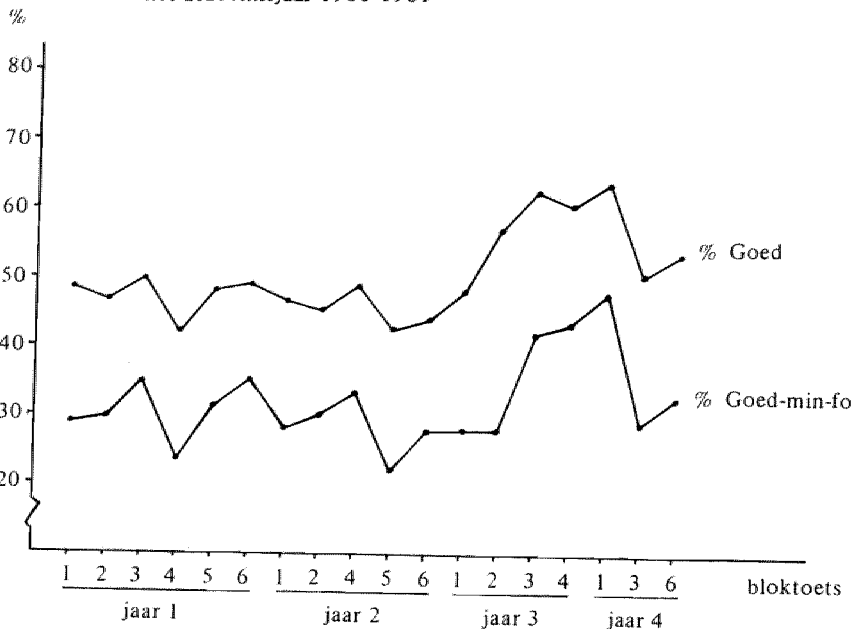
In figuur 1 zijn de gemiddelde scores op alle bloktoetsen in kaart gebracht. Dat geldt voor zowel goed-scores als voor goed-min-fout-scores.

In figuur 2 zijn bloktoetsscores na middeling over studiejaar ingetekend. (Voor beide figuren: zie volgende pagina.)

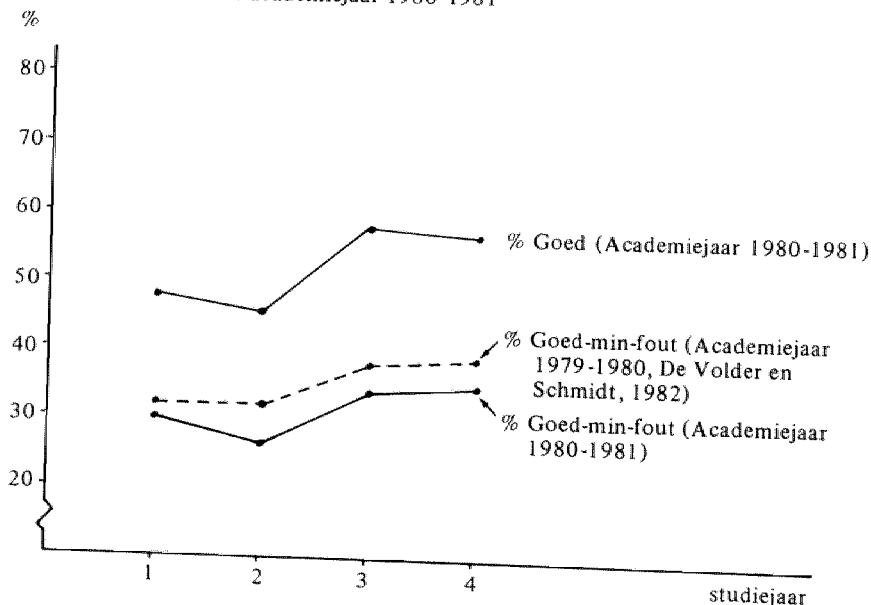
Het eerste dat naar aanleiding van deze resultaten op te merken valt, is dat het verschijnsel zoals dat door Crombag (1981) en De Volder en Schmidt (1982) geconstateerd is, ook in deze data terug te vinden is. Ouderejaarsstudenten leveren gemiddeld genomen betere prestaties op studietoetsen dan jongerejaarsstudenten, zelfs als die toetsen inhoudelijk van elkaar verschillen.

Men kan zich afvragen wat de verklaring voor dit verschijnsel kan zijn. De selectieve werking van het programma waarbij alleen de goede studenten overblijven – een mogelijke verklaring voor de door Crombag (1981) geconstateerde effecten – speelt in de hierboven gepresenteerde data geen rol omdat zoals gezegd in het curriculum van de medische faculteit maar op zeer beperkte schaal selectie plaatsvindt. Ook kan het verschijnsel niet teruggevoerd worden op een artefact van de gebruikte scoringsprocedure, zoals voor de data van De Volder en Schmidt (1982) nog gevreesd werd. De goed-score volgt namelijk in grote lijnen de goed-min-fout-score. Een

Figuur 1: Gemiddelde percentuele scores op bloктоetsen in vier studiejaar voor het academiejaar 1980-1981



Figuur 2: Gemiddelde percentuele prestatiescores geaggregeerd binnen studiejaar voor het academiejaar 1980-1981



zich wijzigende antwoordstrategie van de kant van de studenten kan dus niet voor het effect verantwoordelijk worden gesteld.

Merkwaardig is dat de prestatieverbetering zich sprongsgewijs voordoet. Tussen tweede en derde studiejaar stijgt de prestatie duidelijk, terwijl verschillen tussen jaar 1 en 2 aan de ene kant en jaar 3 en 4 aan de andere kant te verwaarlozen lijken. De Volder en Schmidt (1982) stelden hetzelfde verschijnsel vast (zie ook figuur 2). Het is niet uitgesloten dat deze sprong op de een of andere manier verband houdt met karakteristieken van het programma. De eerste twee jaar van de medische studie leggen wat sterker de nadruk op verwerking van basale (vooral biologische) kennis, terwijl in jaar 3 en 4 kennis aan de orde komt die meer-directe relevantie heeft voor de beoefening van de gezondheidszorg. Op grond van deze constatering dringt zich een motivationele verklaring voor het verschijnsel op. Men zou kunnen veronderstellen dat studenten in de latere fasen van hun studie onderwerpen bestuderen die door hen als relevanter voor de beroepspraktijk worden waargenomen dan de onderwerpen die in het begin van de studie aan de orde komen. In medische curricula is dat zeker het geval. In de eerste jaren komen de zogenaamde basisvakken (anatomie, fysiologie, biochemie, etc.) aan de orde, die betrekking hebben op de normaal functionerende mens, terwijl in het derde en het vierde jaar de nadruk ligt op de klinische vakken. (In Maastricht wordt dit onderscheid overigens minder dan elders gehanteerd in de programmaopbouw.) Het is dan ook verleidelijk de stijgende prestatie toe te schrijven aan toenemende studiemotivatie.

Deze hypothese kan — zij het op indirecte wijze — getoetst worden, omdat naast de resultaten op studietoetsen ook andere programmaevaluatiegegevens verzameld worden. Dat gebeurt onder andere door middel van een vragenlijst die na afloop van elk blok aan de studenten wordt voorgelegd. Die vragenlijst heeft betrekking op een groot aantal verschillende aspecten van het onderwijsprogramma. Onder andere wordt hen gevraagd naar plezier in het werk, gepercipiëerde relevantie van de leerstof, en studietijd. Deze variabelen kunnen worden beschouwd als indicatoren voor studiemotivatie (Carroll, 1963; Bloom, 1976). Ze werden als volgt geoperationaliseerd:

	volledig oneens			volledig eens		
— Over het geheel genomen heb ik de afgelopen periode plezierig gewerkt.	1	2	3	4	5	
— De onderwerpen die in dit blok aan de orde zijn gekomen heb ik in het kader van mijn studie als relevant ervaren.	1	2	3	4	5	
— Het programma vergde teveel studietijd.	1	2	3	4	5	

Daarnaast werd studenten nog gevraagd hoeveel tijd ze gemiddeld genomen per week aan individuele studie besteed hadden.

Als motivatiehypothese ter verklaring van het gevonden verschijnsel van de stijgende prestatiescores enige plausibiliteit heeft, mag men positieve correlaties verwachten tussen de hierboven genoemde variabelen enerzijds en studieprestaties anderzijds. In tabel 1 zijn correlaties tussen deze variabelen opgenomen.

Tabel 1: Correlaties¹ tussen indicatoren van studiemotivatie en studieprestaties

	Studieprestaties
Plezier in het werk	0,30
Relevantie van leerstof	0,18
Programma vergde teveel studietijd	0,29
Studietijd in hele uren	0,43

Met uitzondering van de relevantievariabele zijn alle variabelen statistisch significant op het 5% -niveau. De correlatie tussen de schatting van de studietijd en studieprestatie is zelfs verrassend hoog: studietijd verklaart meer dan 18% van de variantie in studieprestaties.

Dit alles duidt erop dat motivationele factoren een rol spelen bij het eerder geconstateerde effect.

Er is echter nog een andere potentiële kandidaat voor de verklaring van het verschijnsel.

Als we ervan uitgaan dat docenten niet doelbewust de moeilijkheidsgraad van toetsen variëren, en dat dus de met het studiejaar toenemende prestatie geen functie is van een of ander ingrijpen van de kant van de onderwijsgevers, dan zou het wel of niet hebben van relevante *voorkennis* een rol kunnen spelen. Studenten worden in de eerste studiejaar vaak geconfronteerd met een overvloed aan gegevens, concepten, theorieën en opvattingen, die volstrekt nieuw voor hen zijn en waarvoor het hen aan relevante voorkennis ontbreekt. Daarnaast ontberen zij meestal een goed overzicht met betrekking tot de structuur van het vakgebied dat zij zich eigen moe-

¹ Deze correlaties zijn berekend op andere gegevens dan de hier gepresenteerde, namelijk de studieresultaten over het academiejaar 1981-1982. Dat vindt zijn oorzaak in het feit dat de vragenlijst met behulp waarvan de informatie over plezier in het werk, relevantie van de stof en studietijdgegevens verzameld is, pas dit academiejaar voor het eerst afgenomen werd. De studieresultaten over '81-'82 vertonen, voor zover beschikbaar, overigens dezelfde trend als die van de beide vorige academiejaar.

De correlaties werden noodgedwongen berekend op gemiddelde scores verkregen van onderwijsgroepen (een onderwijsgroep bestaat uit 8 à 10 studenten). Noodgedwongen, omdat de programmaevaluatievragenlijst anoniem wordt afgenomen. Studenten geven slechts het nummer van hun groep op. N = 136.

ten maken. Het is bekend dat het bezit van relevante voorkennis een belangrijke faciliterende invloed uitoefent op het cognitief verwerken van nieuwe informatie (Bransford en Johnson, 1972; Spilich e.a., 1979). Men zou dus kunnen veronderstellen dat ouderejaarsstudenten in de loop van hun studie een adequate voorkennis hebben verworven en daarom in staat zijn leerstof gemakkelijker te verwerken dan jongerejaarsstudenten, gegeven een standaard hoeveelheid studietijd van zes weken, de lengte van een blok.

Ook deze hypothese kan indirect getoetst worden aan de hand van programmaevaluatiegegevens. Aan studenten wordt na elk blok gevraagd aan te geven in welke mate de leerstof aansloot bij hun voorkennis. Daarnaast wordt hen gevraagd een oordeel uit te spreken over de moeilijkheidsgraad van de stof. Ten slotte moeten ze aangeven in hoeverre de doelstellingen van het blok hen duidelijk waren. Als de hiervoor geformuleerde overwegingen met betrekking tot de rol van voorkennis niet van redelijkheid ontbloomt zijn, mag men positieve relaties verwachten tussen deze variabelen en studieprestaties. In tabel 2 worden de betreffende correlaties gerapporteerd.

Tabel 2: Correlaties tussen indicatoren van voorkennis en studieprestaties

	Studieprestaties
Blok sloot aan bij voorkennis	0,52
Leerstof was relatief moeilijk	-0,01
Doelstellingen blok waren duidelijk	0,45

De correlatie tussen de moeilijkheidsgraad van de stof en studieprestatie verschilt niet significant van nul, de beide andere variabelen correleren relatief hoog met studieprestatie. De correlatie tussen de vraag naar de voorkennis en studieprestatie is zelfs de hoogste van allemaal. Er lijken dus aanwijzingen te zijn dat ook de aan- of afwezigheid van adequate voorkennis een rol speelt bij het optreden van het effect dat studieprestaties stijgen met studiejaar. (De lezer moet zich bij deze conclusie wel realiseren dat ze gebaseerd is op een zwakke test: niet de feitelijke voorkennis van studenten is onderzocht. Men is slechts gevraagd een *oordeel* te geven over hun voorkennis. Die twee zaken hoeven, zoals bekend, niet overeen te komen.)

Om op een nog wat andere wijze een beeld te geven van het belang van de besproken onderwijsvariabelen werd een multiple regressie van deze variabelen op studieprestaties uitgevoerd. De variabelen werden stapsgewijs in de analyse betrokken, die met de hoogste correlatie het eerst, de laagste het laatst. (De variabelen wier correlaties niet significant van nul verschilden, werden buiten beschouwing gelaten.) In tabel 3 zijn resultaten van deze analyse opgenomen.

Tabel 3: Multiple regressie van voorkennis- en motivatievariabelen op studieprestaties

Variabele	R	Bèta
Voorkennis	0,52	0,38
Studietijd in uren	0,62	0,27
Doelstellingen duidelijk	0,65	0,17
Plezier in werk	0,65	0,09
Teveel studietijd	0,66	0,04

R representeert de correlatie tussen studieprestaties en een lineaire combinatie van de betrokken variabelen. Hij kan geïnterpreteerd worden in termen van een percentage verklaarde variantie. Als alle variabelen in de analyse betrokken worden, kan 43% van de variantie in studieprestaties uit deze variabelen verklaard worden. Uit de snel afnemende grootte van de bèta-gewichten kan afgeleid worden dat de twee laatstgenoemde variabelen weinig eigen bijdragen meer leveren aan variatie in studieprestaties.

Conclusie

Twee conclusies lijken gerechtvaardigd.

De eerste is dat nader onderzoek naar wat gemakshalve het Crombag-effect genoemd zou kunnen worden, gewenst is. Met name zou het interessant zijn na te gaan in hoeverre het optreedt in andere curricula met een betrekkelijk laag uitvalspercentage. Als het hier gaat om een algemeen optredend fenomeen zou dat naar ons oordeel consequenties moeten hebben voor de selectiepraktijken in lagere studiejaar in het wetenschappelijk onderwijs. Dit fenomeen suggereert immers dat docenten er blijkbaar niet erg goed in slagen de moeilijkheidsgraad te schatten van de toetsen die zij construeren, omdat ze onvoldoende zicht hebben op de mate waarin de aangeboden leerstof, respectievelijk toets, aansluit op de voorkennis van studenten. In een situatie waarin studietoetsen voor selectiedoeleinden gebruikt worden, kan dit betekenen dat jongerejaarsstudenten moeten afhaken op grond van studieresultaten verkregen met behulp van toetsen die een te hoge moeilijkheidsgraad hebben voor die groep. Het Crombag-effect heeft een, ons inziens ongewild, negatief effect op het rendement van het wetenschappelijk onderwijs.

De tweede conclusie is dat het waarschijnlijk loont in het onderwijs meer expliciet aandacht te besteden aan maatregelen die de aansluiting van nieuwe informatie aan de voorkennis bevorderen en aan maatregelen die motivatie van studenten verhogen.

We zijn ons ervan bewust dat dat gemakkelijker is gezegd dan gedaan.

Literatuur

- Bloom, B.S. *Human characteristics and school learning*. New York: McGraw-Hill, 1976.
- Bransford, J.D. en M.K. Johnson. Contextual prerequisites for understanding: some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior*, 1972, 11, 717-726.
- Carroll, J.B. A model of school learning. *Teachers College Record*, 1963, 64, 723-733.
- Crombag, H.F.M. Studiejaar en tentamenresultaat. *Tijdschrift voor Onderwijsresearch*, 1981, 6, 247-248.
- Dousma, T. en A. Horsten. *Tentamineren*. Utrecht: Het Spectrum, 1980.
- Spilich, G.J., G.T. Vesonder, H.L. Chiesi en J.F. Voss. Text processing of domain-related information for individuals with high and low domain knowledge. *Journal of Verbal Learning and Verbal Behavior*, 1979, 18, 275-290.
- Volder, M.L. de en H.G. Schmidt. Tutor: procesbegeleider of inhoudsdeskundige? In H.G. Schmidt (red.). *Probleemgestuurd Onderwijs*. Stichting voor Onderzoek van het Onderwijs. Harlingen: Flevodruk, 1982.
- Wijnen, W.H.F.W. Toespraak gehouden bij het neerleggen van het Rectoraat, Maastricht, januari 1982.

SAMENVATTING

De centrale these van dit proefschrift was dat de analyse van een probleem een activerende en herstructurerende invloed uitoefent op voorkennis die deelnemers aan die probleemanalyse beschikbaar hebben, en dat dit proces het begrijpen, onthouden en gebruiken van informatie uit tekst vergemakkelijkt.

Een probleem werd daartoe in hoofdstuk 1 gedefinieerd als een verzameling verschijnselen of gebeurtenissen uit de werkelijkheid; en de taak van studenten als het bedenken van verklaringen voor die verschijnselen of gebeurtenissen.

In hoofdstuk 2 en 3 werden de resultaten besproken van twee experimenten waarin bovenstaande hypothese onderzocht werd. In het eerste experiment kregen proefpersonen een beschrijving van het gedrag van een rode bloedcel in water en in een zoutoplossing voorgelegd, met het verzoek dit gedrag te verklaren. In Experiment 2 analyseerden proefpersonen ditzelfde probleem, maar kregen vervolgens een tekst over osmose en diffusie voorgelegd.

De resultaten van het eerste experiment kunnen als volgt worden samengevat.

Deelname aan een probleemanalyse leidt tot aanzienlijke activatie en herstructurering van bestaande cognitieve structuren. De probleemanalysegroep produceerde ruim twee maal meer proposities over osmose en diffusie dan een controlegroep. Veertig procent van de in de recall gevonden proposities was door de proefpersonen eerder zelf in de probleemanalyse onder woorden gebracht; dit deel werd het *geactiveerde* deel van de voorkennis genoemd. Twintig procent was eerder door anderen geproduceerd en door de persoon in de eigen cognities geïntegreerd; dit werd het *geherstructureerde* deel genoemd.

Geconcludeerd werd dat deze operationalisatie van activatie versus herstructurering enigszins onbevredigend is, omdat kan worden aangenomen dat zij tot een onderschatting leidt van het aandeel van herstructurering in de probleemanalyse. Individueel nadenken over het probleem zal ook al een zekere herstructurering van cognitieve structuren teweegbrengen die tot uitdrukking zal komen in wat personen zelf produceren.

Een voorstel werd gedaan tot (gedeeltelijke) replicatie van Experiment 1 waarbij de invloed van het probleem op herstructurering van het eigen kennisbestand zou kunnen worden onderzocht door proefpersonen *individueel* aan een probleem te laten werken.

De probleemanalysegroep produceerde drieëneenhalf keer meer *verklarende* proposities dan de controlegroep. Dat duidde erop dat de probleemanalysegroep zich door de confrontatie met het probleem een dieper inzicht in het osmotisch proces verworven had. Dat diepere inzicht kwam tot uit-

drukking in het feit dat de probleemanalysegroep, zoals in hoofdstuk 3 bleek, een betere prestatie leverde op een transfertoets, geconstrueerd om te meten in hoeverre proefpersonen in staat waren kennis met betrekking tot osmose en diffusie ook werkelijk toe te passen in een nieuwe situatie. En ten slotte waren er aanwijzingen dat de probleemanalysegroep ook beter in staat was juiste en onjuiste beweringen over osmose en verwante onderwerpen als zodanig te herkennen.

In Experiment 2 bleek dat activatie en herstructurering van voorkennis door middel van probleemanalyse een niet geringe invloed op de verwerking van een tekst uitoefende: 25% van de variantie in de recall van de tekst werd verklaard uit de experimentele manipulatie. Het effect was waarneembaar over de hele tekst, maar het sterkst met betrekking tot de centrale ideeën uit die tekst. Proefpersonen uit beide condities van het experiment gebruikten de structuur van de tekst als middel om informatie uit die tekst in het lange-termijn-geheugen te organiseren.

Proefpersonen uit de probleemanalysegroep waren beter in staat juiste proposities met betrekking tot de informatie in de tekst te onderscheiden van onjuiste proposities, blijkens hun resultaten op een herkenningstoets, en konden de verworven kennis ook beter toepassen bij het oplossen van vraagstukken in een transfertoets.

Het effect van probleemanalyse was niet alleen robuust maar ook tamelijk stabiel. Bij een recall na 14 dagen was het effect nog duidelijk waarneembaar, zij het dan dat de experimentele groep meer verval vertoonde dan de controlegroep.

Geconcludeerd werd dat activatie en herstructurering van bestaande cognitieve structuren een algemeen facilitatief effect hebben op de codering van informatie uit tekst, waarschijnlijk omdat voorkennis tijdens het verwerkingsproces gemakkelijker toegankelijk is en daardoor de capaciteit van het werkgeheugen vergroot. De mogelijkheid meer informatie per tijdseenheid te verwerken bevordert de hechtheid van de relaties tussen proposities die de cognitieve representatie van de verworven informatie vormen. Deze is op haar beurt verantwoordelijk voor de gevonden superieure recall, herkenning en transfer.

Deze conclusie werd bereikt door uitschakeling van twee alternatieve verklaringen: de mogelijkheid dat activatie van voorkennis leidt tot een andere organisatie van de cognitieve representatie, en de mogelijkheid dat activatie leidt tot meer elaboreren en daarom de aanmaak van meer verschillende retrievalpaden bevordert. Deze beide alternatieve theorieën moesten op grond van de data verworpen worden.

In hoofdstuk 4 werden gegevens uit beide experimenten gerapporteerd, die betrekking hadden op motivationele effecten van probleemanalyse. Onderzoek naar intrinsieke motivatie werd van belang geacht omdat die vorm van motivatie een alternatieve, niet-cognitieve, verklaring zou kunnen zijn voor de in hoofdstuk 2 en 3 gerapporteerde cognitieve effecten.

In beide experimenten was een klein, doch significant effect van probleem-analyse constateerbaar. De probleemanalysegroep bleek meer interesse te hebben in nadere informatie over osmose — en gaf zich ook vaker op voor een bijeenkomst over dat onderwerp — dan de controlegroep.

Dit effect werd verklaard met behulp van Berlyne's onzekerheidsreductie-theorie van intrinsieke motivatie, die zegt dat discrepanties tussen de aard van de werkelijkheid en de kennis die mensen daarvan hebben een intrinsiek gemotiveerde behoefte aan kennis ('epistemic curiosity') oproepen, die door middel van informatieverwerking bevredigd wordt. Het feit dat verschillen in intrinsieke motivatie na de verwerking van tekst verdwenen waren, werd gezien als een ondersteuning van Berlyne's theorie.

De aard van de verzamelde motivatiegegevens — er werden zowel voor als na bestudering van de tekst metingen gedaan — maakte het mogelijk enkele modellen te testen met betrekking tot de relatie tussen intrinsieke motivatie en prestatie. Zo werd onderzocht of intrinsieke motivatie de cognitieve activiteit verhoogt; of dat het beschouwd moest worden als affectieve consequentie van (eerder geleverde) prestatie. Geen van de onderzochte modellen had in het licht van de gevonden (partiële) correlaties veel plausibiliteit en geconcludeerd werd dat intrinsieke motivatie en prestatie in belangrijke mate van elkaar onafhankelijke uitkomsten van een onderwijs-leerproces zijn.

Deze conclusie werd gerelativeerd door de constatering dat één motivatie-prestatie-model buiten beschouwing blijven moest, namelijk Carrolls opvatting dat motivatie prestatie beïnvloedt via beïnvloeding van de studietijd. (Zie ook de resultaten van onderzoek gerapporteerd in hoofdstuk 6.) Deze hypothese kon niet onderzocht worden omdat studietijd in Experiment 2 constant gehouden moest worden, ten einde te verhinderen dat plafondefecten in de cognitieve metingen zouden ontstaan.

Een replicatie van Experiment 2 is noodzakelijk om deze kwestie uit te zoeken. Daarbij zou studietijd vrij moeten worden gelaten. Een dergelijke replicatie zou aanbevelingswaardig zijn, niet alleen — en zelfs niet in de eerste plaats — om Carrolls model te testen, maar ook om de externe validiteit van de overige uitkomsten van beide experimenten te onderzoeken. Het normale onderwijs kenmerkt zich juist door het feit dat studietijd (tot op zekere hoogte) vrij is.

In hoofdstuk 5 werden resultaten gerapporteerd van een experiment waarin het vraagstuk centraal stond in welke mate problemen een nadere structurering door middel van toegevoegde vragen nodig hebben om als effectieve stimulans voor de zelfstudie van studenten te kunnen dienen.

Proefpersonen die niet gestudeerd hadden aan de hand van toegevoegde vragen, bleken meer teksten te hebben geraadpleegd, en meer overeen te stemmen met experts bij de beoordeling van de vraag in hoeverre gegeven begrippen gerelateerd waren aan de geanalyseerde problemen. Zij besteedden echter niet meer tijd aan de studie, noch waren verschillen waarneembaar op een retentietoets die na ongeveer een week werd afgenomen. Ook

verschillen beide groepen niet in tevredenheid over hun werkwijze. Uit deze niet eenduidig te interpreteren resultaten werd de negatieve conclusie getrokken dat werken aan een probleem zonder richtinggevende toegevoegde vragen in ieder geval niet tot *slechtere* studieprestaties aanleiding geeft.

In hoofdstuk 6 ten slotte werden relaties tussen voorkennis en (intrinsieke) motivatie enerzijds en studieprestaties anderzijds vanuit een ander perspectief onderzocht. Aanleiding was een per toeval gevonden positieve correlatie tussen studiejaar en studieprestaties: ouderejaarsstudenten van de medische faculteit in Maastricht leveren betere studieprestaties dan jongerejaarsstudenten, zelfs min of meer onafhankelijk van de te bestuderen leerstof. Verondersteld werd dat dit effect het resultaat is van verschillen in voorkennis en motivatie tussen jongerejaars- en ouderejaarsstudenten, die ertoe leiden dat de laatsten betere studieprestaties leveren.

Aan studenten werd per blok onder andere gevraagd of ze voldoende voorkennis hadden ten aanzien van de bestudeerde leerstof, of de doelstellingen hen duidelijk waren geweest, in hoeverre ze met plezier gewerkt hadden, en hoeveel tijd ze aan hun studie besteed hadden. Deze metacognitieve indicatoren van voorkennis en motivatie verklaarden samen 43% van de variantie in studieprestaties.

De resultaten van het onderzoek dat in dit proefschrift gebundeld werd, tonen eens te meer aan hoe doorslaggevend de rol van voorkennis is bij het begrijpen, onthouden en toepassen van nieuwe informatie.

Zij tonen daarmee eveneens aan dat pogingen tot differentiatie en individualisering van het onderwijs te komen geen modieuze stokpaardjes van onderwijsvernieuwers zijn, maar een noodzaak die wortelt in het functioneren van het cognitieve systeem zelf. Geen twee mensen zijn aan elkaar gelijk, en hun uniciteit vraagt om onderwijs dat optimaal tegemoet komt aan ieders voorkennis en ervaring.

Een tweede implicatie van de resultaten is dat voorkennis in het onderwijs blijkbaar niet als een onveranderbaar gegeven beschouwd hoeft te worden, een feitelijkheid die slechts als zodanig geaccepteerd kan worden. Voorkennis blijkt actief gemanipuleerd te kunnen worden op een wijze die de kwaliteit van het leerproces ten goede komt. Probleemgestuurd onderwijs representeert slechts één van de methoden die dat mogelijk maken. Ongetwijfeld zijn andere mogelijkheden denkbaar. Een actief zoeken naar nieuwe onderwijsvormen die zich richten op activatie en herstructurering van voorkennis als eerste stap in het onderwijsleerproces zal zeker vruchten afwerpen. Ten slotte ondersteunen de resultaten enige van de vooronderstellingen die ten grondslag liggen aan het onderwijs, zoals dat aan de Rijksuniversiteit Limburg tot ontwikkeling gebracht wordt.

SUMMARY

The central thesis of this dissertation was that the analysis of a problem activates and restructures the prior knowledge of those participating in this analysis, and that this process facilitates the understanding, retention and use of information from text.

In chapter 1, a problem was defined as a set of phenomena or events from reality, and the students' task was to explain these phenomena or events in terms of underlying processes, principles or mechanisms.

In the chapters 2 and 3 results were discussed of two experiments designed to test the above hypothesis. Subjects participating in the first experiment were given a description of the behavior of a red blood cell in water and in an aqueous salt solution and asked to explain this behavior. Participants in the second experiment analyzed the same problem, and subsequently processed a text on osmosis and diffusion.

The results of the first experiment can be summarized as follows.

Participation in the analysis of a problem leads to substantial activation and restructuring of existing cognitive structures. The problem-analysis group produced more than twice as many propositions about osmosis and diffusion as a control group.

Forty per cent of the propositions found in the recall had been produced before by the subjects themselves in the course of the problem-analysis process. This was called the *activated* part of their prior knowledge. Twenty per cent was produced before by others and had been integrated by the individual into his own cognitions. This part was referred to as the *restructured* part of their prior knowledge.

It was concluded that this operationalization of activation versus restructuring is not entirely satisfactory because it presumably underestimates the part of restructuring of prior knowledge in the problem-analysis. Probably even individual analysis of the problem will cause cognitive restructuring, which will find expression in the propositions, produced by the subjects during the analysis.

It was proposed that the influence of the problem on the restructuring of someone's knowledge structures should be investigated by having subjects tackle the problem *individually*.

The problem-analysis group produced three and a half times as many *explanatory* propositions as the control group. The interpretation of this phenomenon was that in the confrontation with the problem the problem-analysis group had acquired a deeper insight into the osmotic process. As can be deduced from the results presented in chapter 3, this deeper understanding was reflected in a better performance of the problem-analysis group on a *transfer* test. This test had been constructed to measure subjects' ability to apply knowledge about osmosis and diffusion in a new situation.

Finally, there were indications of this group's superiority in recognizing true and false statements about osmosis and related subjects.

Experiment 2 showed that activation and restructuring of prior knowledge by means of problem-analysis had a robust effect on the processing of a text: 25% of the variance in text recall was explained by the experimental manipulation. The effect was perceivable throughout the text, but appeared strongest in its core ideas.

Subjects in either of the experimental conditions used the text structure as a vehicle for organizing the information from the text in long-term memory.

The results of a recognition test showed that the members of the problem-analysis group were better able to distinguish true propositions about the text from false ones. These subjects were also better able to apply knowledge acquired to solve problems contained in a transfer test.

The effect of problem-analysis not only appeared to be rather robust but rather stable as well. In a recall after 14 days, it was still clearly perceivable, be it that the problem-analysis group showed a greater decay than the control group.

It was concluded that activation and reconstruction of existing cognitive structures have a general facilitative effect on the encoding of information from text, probably because prior knowledge is more readily accessible during encoding and thus enlarges the capacity of working-memory. The opportunity of processing more information per time unit promotes the strength of the relations between the propositions constituting the cognitive representation of the information acquired. The latter, in its turn, is responsible for the established superior recall, recognition and transfer. This conclusion was reached by eliminating two alternative explanations: the possibility that activation of prior knowledge leads to a different organization of the cognitive representation, and the possibility that activation results in increased elaboration and in that way stimulates the development of multiple redundant retrieval paths. For these two alternative theories no support could be found in the data.

Chapter 4 reported on data from both experiments with respect to motivational effects of problem analysis.

Investigating motivation was considered important especially because intrinsic motivation might be an alternative, non-cognitive explanation of the cognitive effects reported in the chapters 2 and 3.

A small, but significant, effect of problem analysis has been established in either experiment: the problem-analysis group appeared to be more interested in further information about osmosis and more often wished to attend a lecture on that topic than the control group. These results were explained with the aid of Berlyne's uncertainty-reduction theory of intrinsic motivation, which states that discrepancies between the nature of reality and people's knowledge of it generate an intrinsically motivated need for knowledge ('epistemic curiosity'), which is satisfied by the pro-

cessing of relevant information. The fact that differences in intrinsic motivation had disappeared after the processing of the text was considered supporting Berlyne's theory.

The nature of the intrinsic motivation data — measuring was carried out both before and after text processing — made it possible to test some models of the relationship between intrinsic motivation and achievement. It was examined whether intrinsic motivation increased cognitive activity or whether it should be looked upon as an affective consequence of (previous) achievement. In the light of the (partial) correlations found, none of the tested models showed much plausibility and it was concluded that intrinsic motivation and achievement are to a high degree mutually independent results of a learning process.

The above conclusion is limited by the fact that one motivation-achievement model had to be left unconsidered, namely Carroll's notion that achievement is indirectly affected by motivation because motivation influences study time (see, for instance, the results reported in chapter 6). This hypothesis could not be tested because study time had to be kept constant in Experiment 2 in order to avoid ceiling effects in the cognitive measures.

Replication of Experiment 2 is necessary to investigate the matter more adequately. On that occasion, study time should not be limited. Such replication would be useful not only — and not even in the first place — to test Carroll's model, but also to examine the external validity of the other results of the two experiments. Normal instruction is characterized by the very fact that (to a certain extent) study time is left free.

In chapter 5 an experiment was discussed that focused on the question to what extent problems require structuring by means of adjunct questions in order to serve as an effective stimulus to students' self-study. Subjects who had not studied directed by the adjunct questions appeared to have consulted a larger number of texts, and to show a greater agreement with experts in judging to what extent concepts were related to the analyzed problems. However, they had not spent more time on studying, nor were any differences perceivable on a retention test taken one week later. Also, the two groups did not differ in satisfaction about their strategies either.

From these, somewhat ambiguous, results the negative conclusion was drawn that working on a problem without directional adjunct questions at least does not result in *inferior* achievement.

Finally, in chapter 6, the relation between prior knowledge and (intrinsic) motivation on the one hand and achievement on the other was investigated from a different point of view. Motive for this research was the accidental discovery of a positive correlation between academic year and academic achievement: senior students at the medical faculty in Maastricht show

better achievement than their juniors, independent from the subject studied.

It was assumed that this phenomenon stems from differences in prior knowledge and motivation between junior and senior students, leading to superiority in achievement of the latter group.

At the end of each course students were asked, among other things, if they had had sufficient prior knowledge with respect to the subject-matter studied, if the learning objectives of the course had been clear to them, to what extent they had liked their study activities, and how much time they had spent on their study. These meta-cognitive indicators of prior knowledge and motivation together accounted for 43% of the variance in academic achievement.

The results of the research reported in this dissertation once again show the decisive role of prior knowledge in the understanding, retention and application of new information. At the same time they make clear that efforts aiming at differentiation and individualization of instruction are not fashionable hobbies of educational reformers, but a necessity rooted in the functioning of the cognitive system itself. No two people are identical, and their uniqueness calls for instruction which is optimally adapted to their prior knowledge and experience.

A second implication of the results is that, apparently, in instruction prior knowledge need not be considered an unchangeable condition; a fact that can only be accepted as such. It appears that prior knowledge can be actively manipulated in a way that benefits the quality of the learning process. Problem-based learning represents only one of the possible approaches. Others are conceivable, no doubt. Active search for new ways of activation and restructuring of prior knowledge as a first step in the learning process may yield fruitful results.

Finally, the results of this work support some of the presuppositions underlying the educational approach being developed at Rijksuniversiteit Limburg.

CURRICULUM VITAE

De auteur van deze dissertatie werd op 2 december 1947 in Heerlen geboren. In die stad voltooide hij een gymnasium-bèta opleiding. Van 1968 tot 1974 studeerde hij psychologie in Utrecht. Sinds 1974 is hij als onderwijskundige verbonden aan de capaciteitsgroep Onderwijsontwikkeling en Onderwijsresearch van de Rijksuniversiteit Limburg. In de periode 1980-1982 was hij voorzitter van die groep.